JMIR mHealth and uHealth

Impact Factor (2022): 5.0
Volume 6 (2018), Issue 12 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

Contents

Reviews

Original Papers

Corrigenda and Addenda

Review

Use of Mobile Devices to Help Cancer Patients Meet Their Information Needs in Non-Inpatient Settings: Systematic Review

Rebecca Richards[1], PhD; Paul Kinnersley[2], MD; Kate Brain[1], PhD; Grace McCutchan[1], PhD; John Staffurth[3], MD; Fiona Wood[1], PhD

[1]Division of Population Medicine, Cardiff University, Cardiff, United Kingdom

[2]Centre for Medical Education, Cardiff University, Cardiff, United Kingdom

[3]Section of Oncology, Palliative Care Medicine, Cardiff University, Cardiff, United Kingdom

Corresponding Author:
Rebecca Richards, PhD
Division of Population Medicine
Cardiff University
5th Floor Neuadd Meirionnydd
Heath Park
Cardiff, CF14 4YS
United Kingdom
Phone: 44 02920687219
Email: richardsb3@cardiff.ac.uk

Abstract

Background: The shift from inpatient to outpatient cancer care means that patients are now required to manage their condition at home, away from regular supervision by clinicians. Subsequently, research has consistently reported that many patients with cancer have unmet information needs during their illness. Mobile devices, such as mobile phones and tablet computers, provide an opportunity to deliver information to patients remotely. To date, no systematic reviews have evaluated how mobile devices have been used specifically to help patients meet to their information needs.

Objective: A systematic review was conducted to identify studies that describe the use of mobile interventions to enable patients with cancer meet their cancer-related information needs in non-inpatient settings, and to describe the effects and feasibility of these interventions.

Methods: MEDLINE, Embase, and PsycINFO databases were searched up until January 2017. Search terms related to "mobile devices," "information needs," and "cancer" were used. There were no restrictions on study type in order to be as inclusive as possible. Study participants were patients with cancer undergoing treatment. Interventions had to be delivered by a mobile or handheld device, attempt to meet patients' cancer-related information needs, and be for use in non-inpatient settings. Critical Appraisal Skills Programme checklists were used to assess the methodological quality of included studies. A narrative synthesis was performed and findings were organized by common themes found across studies.

Results: The initial search yielded 1020 results. We included 23 articles describing 20 studies. Interventions aimed to improve the monitoring and management of treatment-related symptoms (17/20, 85%), directly increase patients' knowledge related to their condition (2/20, 10%), and improve communication of symptoms to clinicians in consultations (1/20, 5%). Studies focused on adult (17/20; age range 24-87 years) and adolescent (3/20; age range 8-18 years) patients. Sample sizes ranged from 4-125, with 13 studies having 25 participants or fewer. Most studies were conducted in the United Kingdom (12/20, 52%) or United States (7/20, 30%). Of the 23 articles included, 12 were of medium quality, 9 of poor quality, and 2 of good quality. Overall, interventions were reported to be acceptable and perceived as useful and easy to use. Few technical problems were encountered. Adherence was generally consistent and high (periods ranged from 5 days to 6 months). However, there was considerable variation in use of intervention components within and between studies. Reported benefits of the interventions included improved symptom management, patient empowerment, and improved clinician-patient communication, although mixed findings were reported for patients' health-related quality of life and anxiety.

Conclusions: The current review highlighted that mobile interventions for patients with cancer are only meeting treatment or symptom-related information needs. There were no interventions designed to meet patients' full range of cancer-related information needs, from information on psychological support to how to manage finances during cancer, and the long-term effects of treatment.

More comprehensive interventions are required for patients to meet their information needs when managing their condition in non-inpatient settings. Controlled evaluations are needed to further determine the effectiveness of these types of intervention.

(JMIR Mhealth Uhealth 2018;6(12):e10026) doi:10.2196/10026

KEYWORDS

cell phone; smartphone; computers, handheld; cancer; neoplasms; patients; information dissemination; consumer health information

Introduction

It is estimated that one in two people in Great Britain will develop some form of cancer during their lifetime [1]. In 2017 in the United Kingdom, 359,000 new cases of cancer were diagnosed and the rate of incidence is increasing [2]. However, UK survival rates have doubled in the last 40 years and so, for many patients, cancer is a chronic condition they live with for many years [2]. Subsequently, there has been a shift from inpatient to outpatient and community cancer care, where patients are required to manage their condition at home, away from regular supervision by clinicians. This change in care requires patients to take a more active role in their treatment and survivorship. Patients are often faced with an uncertain future, unfamiliar tests and procedures, complex decisions about treatment options, treatment-related side effects, and lifestyle changes. To take a more active role in their care, and to cope with and manage these changes to daily life, patients require relevant information [3]. Research has established that patients with cancer have a wide range of information needs throughout their illness. Studies suggest that patients generally want information on the extent of the disease, likelihood of cure and prognosis, available treatments, side effects of treatment, self-care, and return to normal life [4-6]. Other, less urgent, information needs include the impact of cancer and treatment on social activities, family and friends, mental well-being, and sexual activity, and the risk of family and friends getting cancer [4-6]. A need is described as a desire to receive support with an experienced problem [7], and so an information need can be described as the more specific desire for informational support. It is important to note that an information need is separate from other types of needs, such as emotional or practical needs. However, information related to other types of illness-related needs can enable patients to meet these other needs. For example, access to information on services that provide psychological support enables patients to contact those services and meet their emotional needs. In this paper, the term "illness-related information needs" refers to any type of illness-related information needed by a patient, such as information related to the disease itself, treatment, psychological support services, practical support, and so on.

While many people with cancer want as much information as possible about their condition and related issues [8], studies across the United States and Europe have reported very high rates of unmet information needs [4,9]. As well as limiting patients' ability to participate in their care, there is evidence that unmet information needs are associated with a lower quality of life, losing a sense of control over one's life, increased anxiety and depression, and dissatisfaction with care [10-13]. The introduction of "smart" technology has provided a new platform for delivering information-based interventions to patients. Smart devices, such as smartphones and tablet computers, are called "smart" due to their advanced capabilities in comparison to older devices. For example, old generation mobile phones served the sole purpose of sending and receiving communications in the form of text messages and voice calls, whereas the new generation of devices has dramatically enhanced power and capabilities, as well as an increasing list of software apps. In addition to customized apps, new mobile phones and tablet computers are typically equipped with a touchscreen interface, internet access, digital cameras, music players, global positioning systems (GPS) systems, and much more. Tablet computers typically offer a larger touchscreen interface compared to mobile phones. Most mobile phones that are made and sold today can be described as smartphones, as even the cheapest, less advanced mobile phones available offer the same types of functions as the most expensive and advanced smartphones on the market. The more expensive smartphones and tablet computers are also made affordable by low monthly payment plans.

Apps that are built for smart devices can make use of their enhanced capabilities. Many companies have created apps so that it is easy for consumers to find and use their services, and it is now commonplace for people to use apps daily for communication with family and friends, banking, shopping, emailing, gaming, or consulting the news and weather [14]. Due to the many advantages of smart technology, approximately 93% of adults in the United Kingdom now personally own or use a mobile phone, of whom 71% specify that they own a smartphone and over two thirds own or have access to a tablet computer [15]. Importantly, similar statistics of ownership and use have been reported in cancer patient populations [16,17]. For example, one survey of 210 patients with breast cancer reported that 97% (204/210) of patients owned a mobile phone, of which 69% (145/210) specified a smartphone, and 83% (174/210) reported using their mobile phone several times a day, in comparison to a computer by 52% (109/210) [17]. Over half of these patients used their mobile phones for "smart" activities, such as accessing websites (53%, 111/210), emailing (51%, 107/210), or planning or scheduling (49%, 103/210). As studies highlight the increasing use of smart devices surpassing that of conventional computers and laptops, it is important to deliver interventions using the platforms that are preferred by patients [17]. Furthermore, interventions delivered via smart devices have the potential to benefit cancer care due to the wide reach to patients at the point of need and lower cost compared to traditional health care interventions, as well as enabling access to tailored health care to those in resource-poor settings or those facing barriers to accessing traditional health care [18,19]. Subsequently, the UK government has encouraged the integration of interventions delivered by mobile technology into

traditional health care services since the early 2000s [20]. Furthermore, key reviews over the last few years, such as National Health Service (NHS) Five Year Forward [21] and the Wachter review [22], have highlighted the importance of, and urgent push for, digitization in the NHS, in order for it to continue to provide a high level of health care at an affordable cost.

Over the last decade, interventions have been developed and delivered via a range of smart devices, including smartphones and tablet computers, as well as older mobile devices, such as old generation mobile phones, personal digital assistants (PDAs), and other handheld devices that have enhanced capabilities, such as internet access and real-time data transmission. This range of devices is referred to as "mobile" devices throughout this paper, as in the relevant body of literature, as they have been primarily designed to be used when on the move and can be stored away easily on one's person due to their compact size. Due to the many advantages of mobile devices, there has been prolific development of "mobile" interventions over the last decade to facilitate patients' self-management of chronic conditions, such as diabetes, heart disease, and asthma, where patients are at home, without the supervision of a health care professional [23]. Studies have found that these interventions may improve patients' biological markers of disease, quality of life, communication with clinicians and family, and adherence to medication, while reducing health service costs [23-25]. Following the early indicators of the effectiveness of this type of intervention for other chronic conditions, there has been development of mobile interventions to support patients with cancer.

Several existing systematic and scoping reviews have explored the general use of mobile devices for patients with cancer [26-31]. Findings from these reviews show that interventions delivered via mobile devices have been developed for a range of purposes, including the prevention, detection, and management of cancer. However, most interventions have been designed to support patients during the treatment phase, with fewer interventions developed to assist prevention, diagnosis, follow-up, and survivorship. There has not yet been a review that identifies how interventions delivered via mobile devices have been specifically used to enable patients with cancer to meet their illness-related information needs in non-inpatient settings. This paper therefore presents a systematic review and critical appraisal of studies describing the use of interventions delivered via mobile devices that are designed to enable patients with cancer to meet their illness-related information needs in non-inpatient settings. Specifically, we assessed the effects and feasibility of this type of intervention. This review focused on mobile devices due to the growing number of patients that own this type of technology and the advantages of mobile devices in comparison to older types of technology, such as accessibility (eg, cost), portability, and enhanced capabilities.

Methods

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines for the conduct of systematic reviews [32]. The review was registered on the PROSPERO (International Prospective Register of Systematic Reviews) to prevent duplication (registration number: CRD42014010614). At all stages of the search, data extraction, and quality appraisal, 10% of studies were independently double-checked for consistency by another researcher. Discrepancies were resolved through discussion.

Identification and Screening

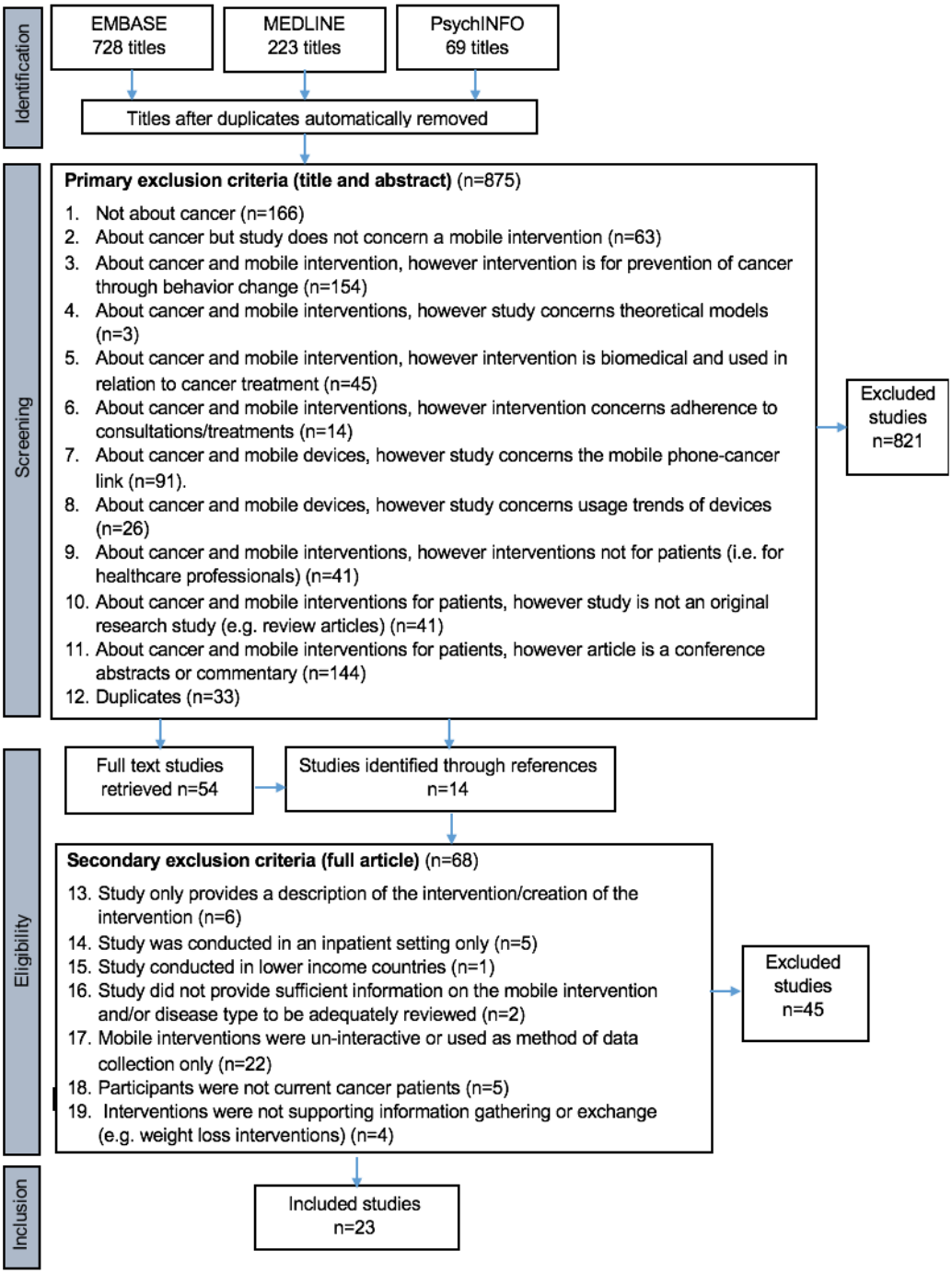
A systematic search of titles and abstracts was conducted in MEDLINE (1946-2017), Embase (1947-2017), and PsycINFO (1806-2017) databases up to January 2017. Search terms focused on three concepts critical to the review question: "mobile devices," "information needs," and "cancer" (Multimedia Appendix 1). Terms relating to the same concept were combined using the Boolean operator "OR," and different concepts were combined using the operator "AND." Duplicates were electronically removed using the Ovid de-duplicate function prior to review of abstracts. Titles and abstracts of citations were screened for appropriate studies. References of included articles were searched for further studies.

The aim of this review was to assess data on the effects and feasibility of this type of intervention, provided by empirical studies. Prior to the search, it was therefore decided that gray literature would not be searched as these studies are not peer-reviewed and are unlikely to contain empirical data. Identification of studies included a 4-stage process of identification, screening, eligibility assessment, and inclusion [32]. To be as inclusive as possible, there were no restrictions on study methodology or date of publication. However, searches were limited to include only human studies and those written in English. Included studies were required to meet the following criteria: (1) interventions were delivered by a mobile or handheld device (eg, mobile phone, PDA), (2) interventions attempted to meet patients' illness-related information needs, (3) primary participants were patients with cancer who were undergoing treatment, and (4) interventions were for use in non-inpatient settings, or non-inpatient *and* inpatient settings. Only those participants who currently had cancer were included in this review as cancer survivors may have different information needs to those who are currently undergoing treatment for cancer. Additionally, only interventions that were used to support patients in non-inpatient settings were included, as this is where patients are now primarily managed for most of their time during their illness.

Eligibility and Inclusion

Searches during the identification stage generated 1020 citations. A total of 54 articles were considered appropriate for eligibility screening, and an additional 14 articles were identified through references. The full texts of these 68 articles were screened using the inclusion criteria, which resulted in the exclusion of a further 45 articles. Reasons for exclusion of articles are documented in the PRISMA flowchart (Figure 1). As a result, 23 articles were included in the review.

Figure 1. PRIMSA flowchart.



Data Extraction and Synthesis

Data were extracted into a template under the following headings: research identification (authors, year of publication, country of study sample, study population), intervention (intervention type, mobile device type), research methods (study design, method, data analysis), outcome measures, principal findings, and quality appraisal. Due to a lack of suitable data, a meta-analysis was not conducted. A narrative synthesis was performed and the findings were organized by common themes found across studies [33].

Quality Appraisal

Included studies were assessed for methodological quality using the Critical Appraisal Skills Programme checklists for quantitative and qualitative research [34]. The quality of each study was assessed according to each domain included in the checklists, including methodology, design, recruitment, data collection, data analysis, ethical issues, reporting of findings, and contribution to research. The overall quality of the studies was categorized as good, medium, or poor. The checklists each consisted of 10 sections of appraisal questions, with one point assigned for satisfying the criteria for each section. However, half a point was awarded for a section if researchers deemed some of the criteria to be satisfied. A total score of 1-5 was

considered "poor" quality, 6-7.5 was considered "medium" quality, and 8-10 was considered "good" quality.

Results

Description of Included Studies

A total of 20 studies were described by the 23 included articles (Table 1 [35-57]). Within these 20 studies, 14 different interventions were identified. The Advanced Symptom Management System was used in six studies (described by nine of the 23 articles), and the Cancer Care Home Telehealth intervention was used in two studies (described by two of the 23 articles). The remaining 12 articles described 12 separate intervention studies. Of the 23 articles, there were 13 early-phase feasibility studies, one full randomized controlled trial (RCT), three pilot RCTs, three process evaluations, one matched-case control study, a secondary qualitative analysis of data generated by an RCT included in this review, and an analysis of software-logged data from a feasibility study included in this review. Sample sizes of patients ranged from 4 to 125, with 13 studies having 25 participants or fewer. Of the 23 articles included, 12 were of medium quality, nine of poor quality, and two of good quality (Multimedia Appendix 2).

Sample Characteristics

Patients with a wide range of cancer types were included in studies. A total of 17 studies were of adult patients, and three studies were of children or adolescent patients. Ages of adult patients ranged from 24-87 years, and ages of child/adolescent participants ranged from 8-18 years. Nineteen studies included non-inpatient participants only. Nine studies provided participants with a mobile device on entry to the study, a further four studies provided devices for participants but participants needed to have a landline phone in order to participate, two studies required participants to own a mobile device, and five studies failed to report whether participants needed to own a mobile device to participate or a device was provided for the study period. It is also worth noting that one study that provided a mobile device for participants included only those who were "able and willing" to use a mobile device and another study excluded participants if they had poor proficiency with the device.

Description of the Interventions

Types of Mobile Devices

Ten interventions were run on mobile phones; nine of which specifically used smartphones. One intervention that required participants to use their own mobile phone for the study included both smartphones and non-smartphones. Four interventions were run on tablets, and two were run on a PDA (a palmtop computer that functions as a personal organizer but also provides access to the internet). A further four interventions were run on handheld devices that were attached to the participants' phone line. Studies that used a handheld device did not report the functions of this type of mobile device; however, these devices are typically the most limited device type in terms of functions. Studies published from 2013 onwards used more advanced smartphones and tablet computers that are commonly used today, such as iPhones and iPads.

Intervention Characteristics

Two interventions were primarily designed to directly increase patients' knowledge of their upcoming surgical operations and coping with cancer-related pain, respectively. One further intervention study primarily aimed to improve patients' communication of symptoms to clinicians in consultations, thereby facilitating information exchange. The primary aim of the remaining 17 intervention studies was to improve the monitoring and management of treatment-related symptoms. These interventions provided treatment-related self-care information following patients' symptom reports or included a system where clinicians would be alerted to contact patients and exchange symptom-related information in order to manage severe symptoms. One of these 17 interventions also provided cognitive and behavioral skills training in non-pharmacological pain management strategies. Study periods ranged from 5 days to 6 months; however, some study periods may have been longer due to the duration of participants' treatment, which was not reported.

Themes

Findings from the narrative synthesis were organized into two main themes: (1) acceptability of the interventions, which included the subthemes of perceived usefulness, perceived ease of use, and adherence to interventions, and (2) benefits of the interventions, which included the subthemes of symptom management, patient empowerment, reassurance and reduced anxiety, patient-clinician communication, and health-related quality of life (HRQOL; Multimedia Appendix 2).

Acceptability

Perceived Usefulness

The mobile interventions were perceived as useful by most patients, particularly the self-care advice provided in response to symptom reports [36,41-44,46,49,54,56,57]. Qualitative interviews with patients who took part in an RCT reported that the information provided them with expectations for their treatment, reminded them to watch for symptoms, and suggested helpful home remedies [43]. Qualitative interviews from another RCT showed that patients were positive about the real-time, fast response of the clinician-alerting facility [49]. However, interviews from a feasibility study found that some patients felt that the depth of the self-care information was insufficient and repetitive [44], and two further feasibility studies revealed variation in use of the self-care advice/information pages [47,50,54]. One study reported that while over half of patients (62%, 37/60) found a mobile phone, symptom-monitoring intervention useful, patients with lower education and chemotherapy-naïve patients rated the intervention significantly more useful than those with higher education (75%, 45/60 vs 35%, 21/60) or those who had received chemotherapy before (82%, 49/60 vs 53%, 32/60) [57].

Table 1. Characteristics of included studies.

Study	Study population	Intervention	Methods	Outcome measures
Aldiss et al, 2010 [35]	4 adolescent patients. Non-Hodgkins lymphoma and osteosarcoma. Age range 13-15 years. United Kingdom	PDA, symptom-monitoring for one cycle of chemotherapy (2 weeks). Mobile device provided	Mixed methods, pilot RCT. Semistructured questionnaires, interviews. Narrative summary of findings	Patients' perceptions of the intervention (effects of the intervention, acceptability)
Besse et al, 2016 [36]	9 adult patients. gastrointestinal, lung, pancreatic, urogenital cancers, osteosarcoma, unknown/ other cancers. Mean age 58 years. Netherlands	Mobile phone, pain monitoring for 4 weeks. Access to own mobile device required	Quantitative, feasibility study. Questionnaires. Paired *t* tests	Pain, quality of life, satisfaction with the intervention
Chumbler et al, 2007 [37]	125 adult patients. Lung, head and neck, colorectal, other cancers. Mean age 63 years. United States	Handheld device, symptom-monitoring for 6 months. Access to home phone line required	Quantitative, matched-case control study. Electronic medical records. Multivariate regression	Number of preventable service uses (ie, unplanned clinical visits), and cancer-related service uses (ie, expected clinical visits) over 6-month period
Chumbler et al, 2007 [38]	48 adult patients. Lung, head and neck, colorectal, other cancers. Mean age 64 years. United States	Handheld device, symptom-monitoring for 6 months. Access to home phone line required	Quantitative, feasibility study. Questionnaires, medical records. Descriptive statistics, linear mixed regression	Patients' cooperation with the intervention (adherence) and health-related quality of life during cancer treatment
Dawes et al, 2015 [39]	20 adult patients, 18 of which had colorectal cancers. Median age 58 years. United States	Tablet computer, symptom monitoring for 6-24 days, depending on time between operation and clinic visit. Mobile device provided (participants excluded for poor proficiency)	Mixed methods, feasibility study. Questionnaires. Descriptive statistics, qualitative data was summarized narratively	Adherence, patient perceptions of the intervention (effects of the intervention)
Foley et al, 2016 [40]	39 adult patients. Breast cancer. Median age in intervention group 54 years. Ireland	Tablet, information provision prior to surgery. 1 week. Mobile device provided	Quantitative, pilot RCT. Questionnaires. Mann-Whitney tests, Fischer's Exact tests	Anxiety and depression, mental adjustment to cancer and satisfaction with information received
Forbat et al, 2009 [41]	12 adult patients from intervention arm of Kearney et al. Colorectal and breast cancer. Mean age 50 years, age range 38-66 years. United Kingdom	Mobile phone, symptom-monitoring for 4 weeks of chemotherapy (12-16 weeks). Provision of device unknown	Qualitative, secondary analysis. Semistructured interviews. Foucauldian approach with focus on surveillance and power	Patients' perceptions of the intervention (effects of the intervention)
Fortier et al, 2016 [42]	12 adolescent patients. Leukemia, tumors of the central nervous system. Mean age 12 years. United States	Tablet, pain monitoring for 10 days. Mobile device provided	Quantitative, feasibility study. Questionnaires. Descriptive statistics. One-sample Wilcoxon signed rank tests were performed to determine whether the observed median was equal to the middle value of the scale for each test	Patient perceptions of the intervention (satisfaction, perceived usefulness), symptom assessment, pain assessment, pain-related coping strategies

Study	Study population	Intervention	Methods	Outcome measures
Head et al, 2011 [43]	44 adult patients. Head and neck cancers. Mean age 59 years. United States	Handheld device, symptom-monitoring for the duration of treatment, average 70 days (around 10 weeks). Access to home telephone line required	Mixed methods, process evaluation (from an RCT). Interviews, phone questionnaires. Descriptive statistics, correlation analysis, descriptive qualitative analysis	Feasibility (median and modal use, nurse-initiated contacts), satisfaction with the intervention, and long-term impact of the intervention. Narrative responses and a poststudy survey provided additional data examining feasibility and satisfaction with the intervention. While outcomes of the clinical trial are not the subject of this article, the results of quality of life and symptom burden measures for the treatment group were reported.
Kearney et al, 2006 [44]	15 adult patients. Lung and colorectal cancer. Age range 24-77 years. United Kingdom	Handheld device, symptom-monitoring for two cycles of chemotherapy (approximately 6-8 weeks). Access to home phone line required	Mixed methods, feasibility study. Semistructured questionnaires, semistructured interviews, software log of activity (reported in McGee and Gray). Descriptive statistics, thematic content analysis	Patients' perceptions of the intervention (effects of the intervention)
Kearney et al, 2009 [45]	112 adult patients. Breast, lung, or colorectal cancer. Mean age 56 years. United Kingdom	Mobile phone, symptom-monitoring for 4 weeks of chemotherapy (12-16 weeks). Provision of device unknown	Quantitative, RCT. Logistic regression	Incidence, severity, and distress of 6 chemotherapy-related symptoms (nausea, vomiting, fatigue, mucositis, hand/foot syndrome, diarrhea)
Maguire et al, 2005 [46]	10 adult patients. Breast and lung cancer. Age range 44-74 years. United Kingdom	Mobile phone, symptom-monitoring for 2 weeks. Provision of device unknown	Mixed methods, process evaluation (from pilot RCT). Semistructured questionnaires, semistructured interviews. Descriptive statistics, thematic content analysis	Patients' perceptions of the intervention (effects of intervention, acceptability)
Maguire et al, 2015 [47]	16 adult patients. Lung cancer. Mean age 64 years. United Kingdom	Mobile phone, symptom monitoring for duration of radiotherapy treatment plus 1-month posttreatment. Provision of device unknown	Mixed-methods, feasibility study. Semistructured questionnaires, semistructured interviews. Descriptive statistics, *t* tests, Mann-Whitney U tests, 1-way ANOVA tests, Kruskal-Wallis tests, Fisher Exact tests, Wilcoxon signed ranks tests, McNemar tests, thematic analysis	Patients' perceptions of the intervention (feasibility, acceptability) anxiety levels, self-care self-efficacy, well-being, quality of life, physical symptom distress
McCall et al, 2008 [48]	21 adult patients receiving palliative care. Breast, prostate, oral, respiratory, gastrointestinal/colorectal, gynecology, myeloma, unknown primary cancers. Mean age 64 years, age range 40-87 years. United Kingdom	Mobile phone, symptom-monitoring for 30 days. Provision of device unknown	Mixed methods, feasibility study. Questionnaires, semistructured interviews. Descriptive statistics, thematic analysis	Patients' perceptions of the intervention (effects of intervention, acceptability)
McCann et al, 2009 [49]	53 adult patients from the intervention arm of Kearney et al. Breast, lung, or colorectal cancer. Mean age approximately 55 years. United Kingdom	Mobile phone, symptom-monitoring for 4 weeks of chemotherapy (12-16 weeks). Provision of device unknown	Mixed methods, process evaluation. Semistructured questionnaires, semistructured interviews. Descriptive statistics, thematic content analysis	Patients' perceptions of the intervention (effects of intervention, acceptability)

Study	Study population	Intervention	Methods	Outcome measures
McGee et al, 2016 [50]	15 adult patients. Lung and colorectal cancer. Age range 24-77 years. United Kingdom	Handheld device, symptom-monitoring for 2 cycles of chemotherapy (approximately 6-8 weeks). Access to home phone line required	Software log of activity, descriptive statistics	Software-logged activity; modem events, questionnaire events, and information access events
Post et al, 2013 [51]	60 adult patients. Breast cancer. Mean age 51 years. United States	PDA, symptom communication with clinicians, for 160 days (around 5 months). Provision of device unknown	Mixed methods, pilot RCT. Questionnaires, interviews. Descriptive statistics, random-effects linear regression, qualitative analysis	Pain, fatigue, and depression symptoms, patients' health-related quality of life and communication self-efficacy. Patients' perceptions of the intervention (effects of the intervention)
Somers et al, 2015 [52]	25 adult patients. Breast, lung, colorectal, prostate cancers. Mean age 53 years. United States	Tablet, pain coping skills. Four sessions (30-45 minutes). Mobile device provided	Mixed methods, feasibility study. Questionnaires, qualitative data collection method not specified. Descriptive statistics, paired sample *t* tests	Patients' perceptions (effects of the interventions, acceptability); pain severity, physical functioning, physical symptoms, psychological distress, self-efficacy for pain management, pain catastrophizing
Stinson et al, 2013 [53]	14 adolescent patients. Acute lymphocytic leukemia, acute myeloid leukemia, Ewing sarcoma, non-Hodgkin's lymphoma, osteosarcoma, rhabdomyosarcoma, other. Mean age 13 years. Canada	Mobile phone, pain-related symptom-monitoring for 2 weeks. Mobile device provided	Mixed methods, feasibility study. Semistructured questionnaires. Descriptive statistics, *t* tests	Patients' perceptions of the intervention (acceptability) and feasibility (adherence)
Sundberg et al, 2015 [54]	9 adult patients. Prostate cancer. Mean age 69 years. Sweden	Mobile phone, symptom monitoring for 2 weeks. Mobile device provided	Mixed methods, feasibility study. Focus group, interviews. Descriptive statistics, content analysis	Software logged data (symptom alerts) patient perceptions of the intervention (acceptability)
Weaver et al, 2007 [55]	6 adult patients. Colon cancer. Age range 54-76 years, median age 64 years. United Kingdom	Mobile phone, symptom-monitoring for two cycles of chemotherapy (approximately 6-8 weeks). Mobile device provided	Mixed methods, feasibility study. Informal interviews. Descriptive statistics, narrative summary of results due to informal nature of interviews	Feasibility (symptom alerts, reasons for alerts, adherence). Patients' perceptions of the intervention (effects of intervention, acceptability)
Weaver et al, 2014 [56]	26 adult patients. Breast, colorectal cancers. Mean age 57 years. United Kingdom	Mobile phone, symptom monitoring for approximately 5 cycles of chemotherapy. Mobile device provided (participants need to be able to use device)	Mixed methods, feasibility study. Questionnaires, interviews. Descriptive statistics, thematic analysis	Feasibility (symptom alerts generated, reasons for alerts, advice given). Patients' perceptions of the intervention (effects of the intervention)
Yap et al, 2013 [57]	68 adult patients. Breast, GI, head & neck, lung, lymphoma, ovarian, cervical, bladder cancers. Median age 50 years. Singapore	Mobile phone, symptom-monitoring for 5 days. Access to own mobile device required	Mixed methods, feasibility study. Semistructured telephone questionnaires. Descriptive statistics, Pearson chi square and Fisher exact tests, qualitative analysis	Feasibility (adherence), number of pharmacists' interventions, patients' perceptions of the intervention (usefulness, acceptability)

Perceived Ease of Use

Almost all patients reported that they found the mobile interventions easy to use, regardless of age, cancer type, and experience with technology [36,43,44,46-48,53-55]. For example, one study reported that all 44 patients from the intervention arm of an RCT reported a handheld device to be very easy (85%, 37/44) or easy (15%, 7/44) to use [43]. Similarly, a feasibility study reported that although 66% (12/18) of patients had little prior computer experience, at poststudy all 11 patients who had received the intervention reported that they felt comfortable using the handheld device [44]. A similar study including a sample of 13 patients receiving palliative care reported that patients lacked confidence and experience in using technology, particularly the internet and PDAs [48]. Poststudy, all patients reported that they felt very comfortable (6/13) or comfortable (7/13) using the mobile phone intervention. However, 5 patients required help from family to complete the electronic questionnaire due to poor physical health. Interviews and questionnaire findings from an RCT and feasibility study

suggested that daily use of a mobile phone intervention did not impact on patients' daily routines or privacy and was not perceived as burdensome or too time-consuming [36,49]. Most patients experienced no or very few technical problems with their mobile devices. Those who did tended to encounter problems with internet connection or practical problems with the device itself [46,48-51,55,56].

Adherence to Mobile Interventions

Studies generally reported high adherence rates to the mobile interventions, regardless of the length of the study [36,38,39,43,51-53,55-57]. A pilot RCT of 44 patients reported that patients used a handheld device consistently for an average of 10 weeks [51]. Similar results were reported in another pilot RCT of 60 patients who used a PDA for approximately 22 weeks, where 83% (49/60) of patients completed symptom inventories and 90% (54/60) watched communication videos when instructed [51]. A feasibility study with the longest study period included in this review (up to 6 months) reported that the mean adherence of 48 patients to daily dialogues with a care coordinator using a handheld device was 84%, with a decrease in adherence as treatment progressed [38]. One study suggested that adherence might be affected by the type of device used or experience with this type of technology, as adherence was significantly higher among smartphone users compared to basic mobile phones users (87%, 52/60 vs 47%, 28/60) [57]. The most common reasons reported for nonadherence to interventions were hospitalization, forgetfulness, and technical problems [43,51].

Benefits of the Interventions

Symptom Management

Most patients perceived the mobile interventions to be helpful in monitoring their treatment-related symptoms. Additionally, studies highlighted that mobile interventions can capture patient information and outcomes that are not captured via conventional reporting, such as questionnaires [39,42,44,46,48,52,54,56]. However, an RCT of 112 breast, lung, and colorectal cancer patients showed mixed results [45]. Authors hypothesized that a real-time, symptom monitoring intervention would facilitate better measurement of six chemotherapy-related symptoms, resulting in more timely interventions. Although two out of six monitored symptoms were significantly different between groups, there were conflicting findings of significantly lower reports of fatigue and significantly higher reports of hand/foot syndrome in the intervention versus control group. There was some evidence to suggest that symptom-monitoring interventions have the potential to reduce the unnecessary use of health care services by improving symptom management [36,37,56]. For example, a matched case-control study of 125 patients investigated the effects of a handheld device intervention by measuring patients' unexpected and expected use of cancer-related services over 6 months [37]. Findings showed that the intervention group had significantly lower use of unexpected care services and significantly higher use of most expected care services. However contrastingly, patients in the intervention group had significantly fewer expected clinic visits compared to controls. Authors suggested this contrasting result was possibly due to patients resolving issues with the care coordinator prior to an expected clinic visit thereby reducing the need for the visit.

Most of the symptom-monitoring intervention studies further reported that patients perceived that the interventions had led to improved symptom management [39,43,46,47,49,52,56]. A process evaluation from an RCT of 44 patients found that 52% (23/44) reported that they were much better, and 44% (19/44) somewhat better, at managing their condition as a result of a handheld, symptom-monitoring intervention [43]. A more recent feasibility study reported that participants showed significantly decreased pain severity, physical symptoms, psychological distress, and pain catastrophizing following a tablet-run pain-coping skills intervention [52]. Similarly, a feasibility study of a mobile phone intervention [36] reported that the mean pain score of participants from the start to end of a feasibility study decreased nonsignificantly, but when measured using the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30), the mean pain score decreased significantly from 56 to 35. Furthermore, two studies reported that patients were admitted to hospital as a result of a real-time symptom monitoring intervention, which resulted in proactive management of those patients' symptoms [36,56].

Patient Empowerment

Some studies suggested that remote monitoring of symptoms empowered patients to participate in their care and better manage their condition due to increased knowledge of their condition and symptom management strategies provided by the mobile interventions [39,42-44,56]. In qualitative interviews with 11 lung and colorectal cancer patients, patients explained that this type of intervention had increased their understanding of their symptom-related problems and consequently, their confidence in their abilities to manage symptoms [44]. Furthermore, six patients who used a mobile phone, symptom-monitoring intervention reported that they felt more involved and responsible for their care [55]. More recent studies supported these results [52,56]. A feasibility study of a mobile phone intervention reported that patients felt more in control of their care and had increased confidence to self-manage their condition at home as a result of the intervention [56]. Similarly, a feasibility study of a tablet device intervention showed that 95% (20/25) of patients reported that the intervention helped them understand the experience of pain and 76% (19/25) of participants felt the intervention had taught them skills that improved their pain coping. However, an observed increase in pain self-efficacy following the pain-related coping skills intervention was not significant [52]. Finally, a similar feasibility study of a tablet device intervention [42] reported on the perceived usefulness of pain management strategies used by children, including self-talk, heat application, and social support and suggested that this type of intervention provided patients with the opportunity to increase their self-efficacy in coping with pain during treatment.

Reassurance and Reduced Anxiety

Most of the studies reported that patients perceived clinicians' surveillance of, and response to, their symptoms as reassuring. There were some mixed findings, however, for the effects of

information on levels of anxiety [40,41,44,46-49,54-56]. Qualitative interviews with 12 patients from a process evaluation of an RCT of a mobile phone, symptom-monitoring intervention reported that patients felt secure in the knowledge that clinicians were being alerted about their symptoms [49]. Results from a secondary analysis of these interviews suggested that patients viewed their surveillance as liberating, freeing them of the worry of deciding to contact clinicians themselves [41]. Similar perceptions were reported by patients in a smaller pilot RCT, where patients felt the mobile phone intervention allowed them to relax [46]. In contrast, a feasibility study of a mobile symptom monitoring intervention reported no change in anxiety levels [47] and one study suggested that information interventions may increase patients' anxiety [40]. A pilot RCT study of a tablet-based information provision intervention found that there was a significant increase in pre-operative fatalism in the intervention group and anxiety was significantly lower in the control group at 7 days postoperation [40]. This study suggests that increasing patients' knowledge of treatment could potentially increase rather than reduce their anxiety. However, authors reported that some women were anxious about using a tablet computer that they were unfamiliar with and this may have increased their anxiety [40]. Additionally, the follow-up period was short at 7 days after surgery.

Patient-Clinician Communication

Many patients perceived that communication with clinicians had improved or that their relationship with clinicians had strengthened as a result of the interventions [35,39,41,43,46,47,55]. A poststudy questionnaire of 44 patients from an RCT of a handheld, symptom-monitoring intervention found that 65% (29/44) of patients were more satisfied with the communication with their clinicians [43]. A secondary qualitative analysis of patient interviews from an RCT of a mobile phone, symptom-monitoring intervention reported that patients felt the intervention gave them easier access to cancer specialists, as well as increasing the amount of communication with clinicians [41]. Authors suggested that easier access to clinicians may change the dynamic of the traditional hierarchical models of health care to a more patient-centered model, as clinicians are more responsive to the patients' reports and needs. Furthermore, two feasibility studies found that as the intervention prompted clinicians to contact the patients, patients' uncertainty about whether to contact their clinicians when needed was reduced and they felt less "bothersome" to their clinicians [47,55].

Health-Related Quality of Life

Studies reported mixed findings of the interventions on patients' HRQOL [36,38,43,47,51]. An RCT of 44 patients using a handheld device during treatment periods, which required patients to report symptoms 3-5 times daily, reported significant positive correlations between usage of the intervention and physical well-being and emotional well-being scores during treatment [43]. A feasibility study of 48 patients using a handheld device to answer daily symptom questions from a care coordinator found a clinically significant improvement of 6.3 points in patients' HRQOL between baseline and 6 months [38]. This study suggested that a symptom-monitoring intervention could reassure patients who are anxious during treatment, thereby maintaining their HRQOL. In contrast, although one feasibility study reported a nonsignificant increase in quality of life following a pain-monitoring intervention [36], one feasibility study reported no change in well-being [47]; however, both studies had small sample sizes. Negative findings were also reported in a pilot RCT study of 60 patients using a PDA, where patients reported symptoms weekly during treatment periods and viewed videos on how to communicate their symptoms to their clinicians prior to their consultations [51]. This study found that patients' HRQOL was not significantly different between groups. Furthermore, the pre-post treatment decrease in HRQOL was generally greater among the intervention group. Authors suggested that this result might be due to the intervention drawing attention to the symptoms experienced by patients in the intervention group [51]. However, due to the methodological differences between studies, such as study design, measurement of HRQOL, and intervention intensity (eg, intervention functions, interaction with patient and duration of intervention), meaningful comparison of these studies is not possible, though it is possible that intervention intensity is partly responsible for these mixed findings.

Discussion

Principal Results

To our knowledge, this is the first systematic review to identify and critically appraise studies that describe the use of mobile interventions designed to enable patients with cancer to meet their illness-related information needs in non-inpatient settings. The primary aim of most intervention studies included in this review was to improve the monitoring and management of patients' treatment-related symptoms, which included the provision of self-care information and interactive information exchange with clinicians. Although these interventions attempted to educate patients in some way, the information and skills provided were solely related to their treatment. There were no interventions that primarily aimed to meet patients' full range of illness-related information needs by increasing their understanding of their condition and other important, related issues. Overall, findings from this review indicated that patients reported this type of technology and intervention to be acceptable, regardless of age, experience with technology, cancer type, or stage of cancer. Patients perceived the mobile interventions to be useful, particularly the self-care advice and the fast response from clinicians. Additionally, there was evidence to suggest that patients with lower education or chemotherapy-naïve patients could benefit most. Patients also reported that they found the mobile interventions easy to use and nonintrusive on their daily routine, with few technical problems encountered. Adherence to interventions was generally high; however, there was considerable variation in usage of the different intervention components within and between studies. Reported benefits of the interventions included improved symptom management, patient empowerment, and improved clinician-patient communication; however, mixed findings were reported for patients' anxiety and HRQOL.

Findings in the Context of Other Literature

Many mobile interventions have been developed to support patients remotely with a range of chronic conditions, such as diabetes and heart disease. Findings of this review mirror what previous literature has found—mobile technology is an acceptable platform to deliver interventions to patients with chronic conditions, regardless of the patients' type of disease, age, gender, and experience with technology [23-25]. The finding that few technical problems were experienced in this review contrasts previous literature, where many patients cited technical difficulties as a barrier to use and satisfaction with the intervention [58-60]. This contrast may be due to the fact that many interventions for other conditions, such as diabetes and heart disease, require additional technological devices to monitor symptoms (eg, glucose monitor, blood pressure monitor), which would increase the likelihood of technical errors.

Adherence rates to mobile interventions included in this review were generally high throughout the study periods, which were up to 6 months. However, engagement appeared to decrease over the course of the intervention. These patterns mirror those of studies of mobile interventions for other chronic conditions, which included study periods of 12 months [60]. Despite generally high rates of adherence for this type of intervention, there appears to be considerable variation in usage of the different intervention components within and between studies, such as the self-care advice pages. It is important that future studies better describe interventions by coding intervention functions in order to determine the components that are responsible for positive outcomes and enable more systematic evaluations [61].

Patients recognized the benefits of real-time symptom monitoring interventions, such as increased knowledge and confidence to participate in self-care, which appeared to result in improved management of symptoms. Additionally, the capability of this technology to capture patient-reported outcomes in real-time may be of clinical importance as it promotes timely intervention [60,61]. This could reduce the number of preventable hospitalizations, as suggested by some studies included in this review. Previous studies of mobile symptom-monitoring or adherence interventions have shown similar findings, including improvements to symptoms, such as an increased blood glucose control, increased self-management behaviors, such as increased adherence to treatment, and fewer hospital admissions [23-25,60].

In this review, patients reported that communication with their clinician had improved as a result of the interventions and they found clinicians' monitoring of their symptoms to be reassuring. Similar findings have been reported in studies of symptom-monitoring interventions for other chronic conditions, where patients described feelings of security, felt that they had not been forgotten, and were receiving good care outside of hospital and clinics [62,63]. Mobile interventions offer an inexpensive way to bridge the gap between patients and clinicians and increase their contact at a time when patients require more support following a shift from inpatient to outpatient cancer care.

Findings of this review reported mixed findings on the impact of mobile interventions on patients' anxiety and HRQOL; however, few studies included in this review measured these outcomes. For some patients, having more knowledge on their condition might reduce their anxiety due to the development of realistic expectations of the future and preparedness for treatment-related side effects, resulting in a better experience. Conversely, information might also increase patients' anxiety by drawing attention to their condition, unknown symptoms, or risks of treatment. The few studies that have measured the impact of mobile devices on patients' quality of life or emotional distress for other chronic conditions have also reported mixed findings [64,65]. However, some studies have highlighted the potential of smartphones to specifically increase patients' awareness of stress and emotional well-being, by recording moods during both health and illness, and deliver therapeutic interventions accordingly, which has led to reduced anxiety [65,66]. Mobile interventions can provide an opportunity to increase patients' access to psychological support and deliver psychological interventions remotely at a time when patients are vulnerable.

Quality of Included Studies

The large number of early-phase studies in this field means that many studies included in this review used an uncontrolled design. The current evidence for the effectiveness and feasibility of mobile interventions to support patients with cancer is therefore limited. Although these studies highlighted the potential benefits of such interventions, RCTs are needed to support the findings of this review. Additionally, most studies included in this review were critically appraised as poor or medium quality, which further limits the conclusions that can be drawn from these studies. Limitations of some studies included small sample sizes, samples limited to single cancer types, underreporting of response rates and details of participants who were lost to follow-up, and short study periods. Other limitations included the failure of studies to explore the opinions of patients with negative views and the economic costs of these types of intervention. Additionally, some studies included only participants who had access to their own device or were already able to competently use a mobile device. This inclusion criterion may have biased findings, as those who participated in these studies may have had more favorable perceptions of mobile interventions than those who were unable to participate. Finally, many studies relied on self-reported data, which may have been affected by recall or the Hawthorne effect [67], where participants may have changed their behavior due to knowingly being observed.

Strengths and Limitations of this Review

The Measurement Tool to Assess Systematic Reviews (AMSTAR) checklist was used to assess the quality of this systematic review. Strengths of this review include an a priori design, 10% of studies at each stage of the search, data extraction and quality appraisal was checked for consistency by another researcher, multiple databases and references of included studies were searched, study characteristics were reported, and the studies were critically appraised on their quality, which was considered when drawing conclusions.

However, this review has several limitations. A meta-analysis was not conducted as included studies did not have suitable data to aggregate; however, a narrative synthesis was considered a suitable alternative method to explore the findings of these studies. Other limitations include poor indexing of studies, which may have limited the number of studies included in this review, and some potential studies were found through searching references of included studies. Finally, this review did not report on the perceptions and experiences of health care professionals who participated in some studies as this was beyond the scope of the review.

Implications for Policy and Practice

This review has several implications. First, it established that a wide range of patients with cancer perceived mobile devices to be an acceptable medium to receive interventions remotely. Second, this type of intervention appears to have the potential to provide a range of benefits for patients, clinicians, and the health care service. Specifically, findings of this review suggest that symptom-monitoring interventions that provide treatment-related information to patients have the potential to improve patients' self-management of their condition and provide clinicians with a better understanding of patients' symptom experiences, while improving the patient-clinician relationship. This may lead to earlier detection of treatment-related side effects and timely intervention, which could reduce costs for the health care system. This type of intervention also has the potential to sustain or improve patients' well-being during a time when they typically experience a decrease in well-being. Importantly, this review established that, to date, mobile interventions for patients with cancer have attempted to meet only a single type of information need (eg, treatment-related symptom information, coping skills), which has typically been achieved indirectly.

This review has also identified that more comprehensive interventions are required for patients currently receiving treatment for them to meet their full range of illness-related information needs in non-inpatient settings, where they are now spending most of their time away from the direct supervision of their clinicians. The literature has established that the type of illness-related information required by patients with cancer varies within and between patients with cancer and any unmet information needs will likely depend on the information provided by their health care team. It is therefore unlikely that a single intervention can include this large amount of information in a single intervention and tailor it to an individuals' condition and location for related services. However, there already exists a huge number of useful and reputable cancer-related information resources and services throughout the United Kingdom, such as information websites, telephone helplines, support groups, and financial services, which are developed and run by reputable cancer charities and health organizations. Intervention developers could incorporate and organize existing services within interventions to arm patients with the tools they need to obtain relevant information.

Most interventions identified in this review required continued monitoring and interaction from clinicians; however, involving clinicians places unrealistic demands on an already stretched health care service. Few mobile interventions have been developed to be used independently by patients. Development of such an intervention would support the initiatives of UK governments and health organizations to empower patients to take a more active role in their care by increasing support for patients in non-inpatient settings and harnessing the power of technology to do so [21,22].

Conclusions

This is the first systematic review to identify how mobile devices have previously been used to help patients with cancer to meet their illness-related information needs in non-inpatient settings. So far, the majority of mobile interventions have been designed to enable clinicians' surveillance of patients remotely in the form of symptom-monitoring interventions. Despite promising findings, these interventions have sought only to increase patients' knowledge of their treatment-related side effects and coping strategies. More comprehensive interventions are required for patients who are currently receiving treatment in order to meet their full range of illness-related information needs when managing their condition in non-inpatient settings. Given the variation of information needs within and between patients, it may be useful for intervention developers to incorporate existing cancer-related information resources and services into interventions to enable patients to obtain their desired information. Nevertheless, mobile devices appear to be an acceptable platform to deliver interventions remotely to patients with cancer. This review also highlights the early stage of the research that is being conducted in this area, which limits the conclusions that can be drawn. Following on from early-phase feasibility studies, RCTs are needed to support the findings of this review, further determine the effectiveness of this type of intervention to improve patient outcomes, and support the transfer of interventions into standard practice.

Acknowledgments

This study was funded by Tenovus Cancer Care. There was no involvement from the funder in this review.

Authors' Contributions

RR, PK, KB, GM, JS, and FW were responsible for the concept, design, and conduct of the study. RR was responsible for collection of data and manuscript preparation. GM was responsible for double checking at all stages of the search. FW, PK, and KB extensively reviewed and edited the manuscript drafts. All authors were involved in interpretation of results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 18KB - mhealth_v6i12e10026_app1.pdf]

Multimedia Appendix 2

Findings and quality appraisal of included studies.

[PDF File (Adobe PDF File), 35KB - mhealth_v6i12e10026_app2.pdf]

References

1. Ahmad AS, Ormiston-Smith N, Sasieni PD. Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. Br J Cancer 2015 Mar 03;112(5):943-947 [FREE Full text] [doi: 10.1038/bjc.2014.606] [Medline: 25647015]
2. Cancer Research UK. Cancer incidence statistics. 2017. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence [accessed 2018-09-14] [WebCite Cache ID 72Qaiko1u]
3. Hibbard JH. Patient activation and the use of information to support informed health decisions. Patient Educ Couns 2017 Jan;100(1):5-7. [doi: 10.1016/j.pec.2016.07.006] [Medline: 27432014]
4. Faller H, Koch U, Brähler E, Härter M, Keller M, Schulz H, et al. Satisfaction with information and unmet information needs in men and women with cancer. J Cancer Surviv 2016 Feb;10(1):62-70. [doi: 10.1007/s11764-015-0451-1] [Medline: 25956402]
5. Kassianos AP, Raats MM, Gage H. An Exploratory Study on the Information Needs of Prostate Cancer Patients and Their Partners. Health Psychol Res 2016 Jun 23;4(1):4786 [FREE Full text] [doi: 10.4081/hpr.2016.4786] [Medline: 27403460]
6. Richards R, McNoe B, Iosua E, Reeder A, Egan R, Marsh L, et al. Cancer Information Seeking Among Adult New Zealanders: a National Cross-Sectional Study. J Cancer Educ 2018 Jun;33(3):610-614. [doi: 10.1007/s13187-016-1136-9] [Medline: 27853983]
7. Osse BH, Vernooij-Dassen MJ, de Vree BP, Schadé E, Grol RP. Assessment of the need for palliative care as perceived by individual cancer patients and their families: a review of instruments for improving patient participation in palliative care. Cancer 2000 Feb 15;88(4):900-911. [Medline: 10679661]
8. Ahamad AW, Wallner PE, Salenius S, Ross R, Fernandez E. What do patients really want to know? JCO 2017 May 20;35(15_suppl):e18261-e18261. [doi: 10.1200/JCO.2017.35.15_suppl.e18261]
9. Kent EE, Arora NK, Rowland JH, Bellizzi KM, Forsythe LP, Hamilton AS, et al. Health information needs and health-related quality of life in a diverse population of long-term cancer survivors. Patient Educ Couns 2012 Nov;89(2):345-352 [FREE Full text] [doi: 10.1016/j.pec.2012.08.014] [Medline: 23021856]
10. Blödt S, Kaiser M, Adam Y, Adami S, Schultze M, Müller-Nordhorn J, et al. Understanding the role of health information in patients' experiences: secondary analysis of qualitative narrative interviews with people diagnosed with cancer in Germany. BMJ Open 2018 Dec 12;8(3):e019576 [FREE Full text] [doi: 10.1136/bmjopen-2017-019576] [Medline: 29530909]
11. Faller H, Strahl A, Richard M, Niehues C, Meng K. The prospective relationship between satisfaction with information and symptoms of depression and anxiety in breast cancer: A structural equation modeling analysis. Psychooncology 2017 Nov;26(11):1741-1748. [doi: 10.1002/pon.4358] [Medline: 28024096]
12. Husson O, Thong MSY, Mols F, Oerlemans S, Kaptein AA, van de Poll-Franse LV. Illness perceptions in cancer survivors: what is the role of information provision? Psychooncology 2013 Mar;22(3):490-498. [doi: 10.1002/pon.3042] [Medline: 22307579]
13. Wiener CH, Cassisi JE, Paulson D, Husson O, Gupta RA. Information support, illness perceptions, and distress in survivors of differentiated thyroid cancer. J Health Psychol 2017 Feb 01:1359105317692143. [doi: 10.1177/1359105317692143] [Medline: 28810403]
14. Ofcom. Landline and and mobile statistics Fact sheet. 2016. Fast facts URL: https://www.ofcom.org.uk/about-ofcom/latest/media/facts [accessed 2018-09-14] [WebCite Cache ID 72QVN3Zzl]
15. Deloitte. There's no place like phone. Consumer usage patterns in the era of peak smartphone: UK cut.: Deloitte; 2016. URL: http://www.deloitte.co.uk/mobileuk2016/assets/pdf/Deloitte-Mobile-Consumer-2016-There-is-no-place-like-phone.pdf [WebCite Cache ID 72QVY5UhU]
16. Girault A, Ferrua M, Lalloué B, Sicotte C, Fourcade A, Yatim F, et al. Internet-based technologies to improve cancer care coordination: current use and attitudes among cancer patients. Eur J Cancer 2015 Mar;51(4):551-557. [doi: 10.1016/j.ejca.2014.12.001] [Medline: 25661828]

17. O'Brien C, Kelly J, Lehane EA, Livingstone V, Cotter B, Butt A, et al. Validation and Assessment of a Technology Familiarity Score in Patients Attending a Symptomatic Breast Clinic. World J Surg 2015 Oct;39(10):2441-2449. [doi: 10.1007/s00268-015-3134-1] [Medline: 26178658]
18. Nasi G, Cucciniello M, Guerrazzi C. The performance of mHealth in cancer supportive care: a research agenda. J Med Internet Res 2015 Feb 13;17(1):e9 [FREE Full text] [doi: 10.2196/jmir.3764] [Medline: 25720295]
19. Nasi G, Cucciniello M, Guerrazzi C. The role of mobile technologies in health care processes: the case of cancer supportive care. J Med Internet Res 2015 Feb 12;17(2):e26 [FREE Full text] [doi: 10.2196/jmir.3757] [Medline: 25679446]
20. Liddell A, Adshead S, Burgess E. Technology in the NHS: Transforming the patient's experience of care.: King's Fund; 2008. URL: https://www.kingsfund.org.uk/sites/default/files/Technology-in-the-NHS-Transforming-patients-experience-of-care-Liddell-Adshead-and-Burgess-Kings-Fund-October-2008_0.pdf [WebCite Cache ID 72QVmPto0]
21. NHS. Next steps on the NHS five year forward view. 2017. URL: https://www.england.nhs.uk/wp-content/uploads/2017/03/NEXT-STEPS-ON-THE-NHS-FIVE-YEAR-FORWARD-VIEW.pdf [accessed 2018-09-14] [WebCite Cache ID 72QVwMJhd]
22. Wachter P. Making IT work: Harnessing the power of information technology to improve care in England. Report of the National Advisory Group on Health Information Technology in England. UK: Department of Health; 2016. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/550866/Wachter_Review_Accessible.pdf [WebCite Cache ID 72QW2fahv]
23. 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]
24. Athilingam P, Jenkins B. Mobile Phone Apps to Support Heart Failure Self-Care Management: Integrative Review. JMIR Cardio 2018 May 02;2(1):e10057 [FREE Full text] [doi: 10.2196/10057]
25. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
26. Bender JL, Yue RYK, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013 Dec 23;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
27. Davis SW, Oakley-Girvan I. mHealth Education Applications Along the Cancer Continuum. J Cancer Educ 2015 Jun;30(2):388-394. [doi: 10.1007/s13187-014-0761-4] [Medline: 25482319]
28. Hesse BW, Beckjord E, Rutten LJF, Fagerlin A, Cameron LD. Cancer communication and informatics research across the cancer continuum. Am Psychol 2015;70(2):198-210. [doi: 10.1037/a0036852] [Medline: 25730725]
29. Nasi G, Cucciniello M, Guerrazzi C. The role of mobile technologies in health care processes: the case of cancer supportive care. J Med Internet Res 2015 Feb 12;17(2):e26 [FREE Full text] [doi: 10.2196/jmir.3757] [Medline: 25679446]
30. Odeh B, Kayyali R, Nabhani-Gebara S, Philip N. Optimizing cancer care through mobile health. Support Care Cancer 2015 Jul;23(7):2183-2188. [doi: 10.1007/s00520-015-2627-7] [Medline: 25649121]
31. Pandey A, Hasan S, Dubey D, Sarangi S. Smartphone apps as a source of cancer information: changing trends in health information-seeking behavior. J Cancer Educ 2013 Mar;28(1):138-142. [doi: 10.1007/s13187-012-0446-9] [Medline: 23275239]
32. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
33. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version.: Institute for Health Research; 2006. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.3100&rep=rep1&type=pdf [WebCite Cache ID 72QWKNtR6]
34. CASP UK. 2014. Critical appraisal skills programme (CASP): making sense of the evidence URL: http://www.casp-uk.net/casp-tools-checklists [accessed 2017-12-15] [WebCite Cache ID 6vjIk4uKl]
35. Aldiss S, Taylor R, Soanes L, Maguire R, Sage M, Kearney N, et al. Working in collaboration with young people and health professionals. A staged approach to the implementation of a randomised controlled trial. Journal of Research in Nursing 2010 Sep 17;16(6):561-576. [doi: 10.1177/1744987110380803]
36. Besse KTC, Faber-te BND, Janssen GHP, Vernooij-Dassen M, Vissers KCP, Engels Y. Pain Assessment with Short Message Service and Interactive Voice Response in Outpatients with Cancer and Pain: A Feasibility Study. Pain Pract 2016 Mar;16(3):320-326. [doi: 10.1111/papr.12278] [Medline: 25581306]
37. Chumbler NR, Kobb R, Harris L, Richardson LC, Darkins A, Sberna M, et al. Healthcare utilization among veterans undergoing chemotherapy: the impact of a cancer care coordination/home-telehealth program. J Ambul Care Manage 2007;30(4):308-317. [doi: 10.1097/01.JAC.0000290399.43543.2e] [Medline: 17873662]

38. Chumbler NR, Mkanta WN, Richardson LC, Harris L, Darkins A, Kobb R, et al. Remote patient-provider communication and quality of life: empirical test of a dialogic model of cancer care. J Telemed Telecare 2007;13(1):20-25. [doi: 10.1258/135763307779701112] [Medline: 17288654]
39. Dawes AJ, Reardon S, Chen VL, Kaiser W, Russell MM, Ko CY, et al. Wireless Technology to Track Surgical Patients after Discharge: A Pilot Study. Am Surg 2015 Oct;81(10):1061-1066. [Medline: 26463309]
40. Foley NM, O'Connell EP, Lehane EA, 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]
41. Forbat L, Maguire R, McCann L, Illingworth N, Kearney N. The use of technology in cancer care: applying Foucault's ideas to explore the changing dynamics of power in health care. J Adv Nurs 2009 Feb;65(2):306-315. [doi: 10.1111/j.1365-2648.2008.04870.x] [Medline: 19191934]
42. Fortier MA, Chung WW, Martinez A, Gago-Masague S, Sender L. Pain buddy: A novel use of m-health in the management of children's cancer pain. Comput Biol Med 2016 Dec 01;76:202-214 [FREE Full text] [doi: 10.1016/j.compbiomed.2016.07.012] [Medline: 27479493]
43. Head BA, Keeney C, Studts JL, Khayat M, Bumpous J, Pfeifer M. Feasibility and Acceptance of a Telehealth Intervention to Promote Symptom Management during Treatment for Head and Neck Cancer. J Support Oncol 2011 Jan 01;9(1):e1-e11 [FREE Full text] [doi: 10.1016/j.suponc.2010.12.006] [Medline: 21499540]
44. Kearney N, Kidd L, Miller M, Sage M, Khorrami J, McGee M, et al. Utilising handheld computers to monitor and support patients receiving chemotherapy: results of a UK-based feasibility study. Support Care Cancer 2006 Jul;14(7):742-752. [doi: 10.1007/s00520-005-0002-9] [Medline: 16525792]
45. Kearney N, McCann L, Norrie J, Taylor L, Gray P, McGee-Lennon M, et al. Evaluation of a mobile phone-based, advanced symptom management system (ASyMS) in the management of chemotherapy-related toxicity. Support Care Cancer 2009 Apr;17(4):437-444. [doi: 10.1007/s00520-008-0515-0] [Medline: 18953579]
46. Maguire R, Miller M, Sage M, Norrie J, McCann L, Taylor L, et al. Results of a UK based pilot study of a mobile phone based advanced symptom management system (ASyMS) in the remote monitoring of chemotherapy related toxicity. Clinical Effectiveness in Nursing 2005 Sep;9(3-4):202-210. [doi: 10.1016/j.cein.2006.08.013]
47. Maguire R, Ream E, Richardson A, Connaghan J, Johnston B, Kotronoulas G, et al. Development of a novel remote patient monitoring system: the advanced symptom management system for radiotherapy to improve the symptom experience of patients with lung cancer receiving radiotherapy. Cancer Nurs 2015;38(2):E37-E47. [doi: 10.1097/NCC.0000000000000150] [Medline: 24836956]
48. McCall K, Keen J, Farrer K, Maguire R, McCann L, Johnston B, et al. Perceptions of the use of a remote monitoring system in patients receiving palliative care at home. Int J Palliat Nurs 2008 Sep;14(9):426-431. [doi: 10.12968/ijpn.2008.14.9.31121] [Medline: 19060793]
49. McCann L, Maguire R, Miller M, Kearney N. Patients' perceptions and experiences of using a mobile phone-based advanced symptom management system (ASyMS) to monitor and manage chemotherapy related toxicity. Eur J Cancer Care (Engl) 2009 Mar;18(2):156-164. [doi: 10.1111/j.1365-2354.2008.00938.x] [Medline: 19267731]
50. McGee MR, Gray P. A handheld chemotherapy symptom management system: results from a preliminary outpatient field trial. Health Informatics J 2016 Jul 25;11(4):243-258. [doi: 10.1177/1460458205055686]
51. Post DM, Shapiro CL, Cegala DJ, David P, Katz ML, Krok JL, et al. Improving symptom communication through personal digital assistants: the CHAT (Communicating Health Assisted by Technology) project. J Natl Cancer Inst Monogr 2013 Dec;2013(47):153-161 [FREE Full text] [doi: 10.1093/jncimonographs/lgt027] [Medline: 24395985]
52. Somers TJ, Abernethy AP, Edmond SN, Kelleher SA, Wren AA, Samsa GP, et al. A Pilot Study of a Mobile Health Pain Coping Skills Training Protocol for Patients With Persistent Cancer Pain. J Pain Symptom Manage 2015 Oct;50(4):553-558 [FREE Full text] [doi: 10.1016/j.jpainsymman.2015.04.013] [Medline: 26025279]
53. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013 Mar 08;15(3):e51 [FREE Full text] [doi: 10.2196/jmir.2350] [Medline: 23475457]
54. Sundberg K, Eklöf AL, Blomberg K, Isaksson A, Wengström Y. Feasibility of an interactive ICT-platform for early assessment and management of patient-reported symptoms during radiotherapy for prostate cancer. Eur J Oncol Nurs 2015 Oct;19(5):523-528. [doi: 10.1016/j.ejon.2015.02.013] [Medline: 25813529]
55. Weaver A, Young AM, Rowntree J, Townsend N, Pearson S, Smith J, et al. Application of mobile phone technology for managing chemotherapy-associated side-effects. Ann Oncol 2007 Nov;18(11):1887-1892. [doi: 10.1093/annonc/mdm354] [Medline: 17921245]
56. Weaver A, Love SB, Larsen M, Shanyinde M, Waters R, Grainger L, et al. A pilot study: dose adaptation of capecitabine using mobile phone toxicity monitoring - supporting patients in their homes. Support Care Cancer 2014 Oct;22(10):2677-2685. [doi: 10.1007/s00520-014-2224-1] [Medline: 24771299]
57. Yap KY, Low HX, Koh KS, Un M, Shih V, Chan A. Feasibility and acceptance of a pharmacist-run tele-oncology service for chemotherapy-induced nausea and vomiting in ambulatory cancer patients. Telemed J E Health 2013 May;19(5):387-395. [doi: 10.1089/tmj.2012.0136] [Medline: 23531095]

58. 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]
59. Cleland J, Caldow J, Ryan D. A qualitative study of the attitudes of patients and staff to the use of mobile phone technology for recording and gathering asthma data. J Telemed Telecare 2007;13(2):85-89. [doi: 10.1258/135763307780096230] [Medline: 17359572]
60. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
61. 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]
62. 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]
63. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. [doi: 10.1186/2194-0827-1-1]
64. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. J Med Internet Res 2013 Nov 13;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
65. Kristjánsdóttir OB, Fors EA, Eide E, Finset A, Stensrud TL, van DS, et al. A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain: randomized controlled trial. J Med Internet Res 2013 Jan 07;15(1):e5 [FREE Full text] [doi: 10.2196/jmir.2249] [Medline: 23291270]
66. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. J Med Internet Res 2010 Apr 30;12(2):e10 [FREE Full text] [doi: 10.2196/jmir.1371] [Medline: 20439251]
67. Merrett F. Reflections on the Hawthorne Effect. Educational Psychology 2006 Jan;26(1):143-146. [doi: 10.1080/01443410500341080]

Abbreviations

HRQOL: health-related quality of life
PDA: personal digital assistant
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 02.02.18; peer-reviewed by L McCann, S Davis; comments to author 01.03.18; revised version received 20.06.18; accepted 10.07.18; published 14.12.18.

Please cite as:
Richards R, Kinnersley P, Brain K, McCutchan G, Staffurth J, Wood F
Use of Mobile Devices to Help Cancer Patients Meet Their Information Needs in Non-Inpatient Settings: Systematic Review
JMIR Mhealth Uhealth 2018;6(12):e10026
URL: https://mhealth.jmir.org/2018/12/e10026/
doi:10.2196/10026
PMID:30552082

Review

Mobile-Based Oral Chemotherapy Adherence–Enhancing Interventions: Scoping Review

Xiomara Skrabal Ross[1,2*], BPsych (Hons); Kate M Gunn[1*], BPsych (Hons), MClinPsych, PhD; Pandora Patterson[3,4*], BPsych (Hons), PhD; Ian Olver[1*], MBBS, MD, PhD, FRACP, FAChPM, AFRACMA, FAHMS

[1]Cancer Research Institute, University of South Australia, Adelaide, Australia

[2]School of Health Sciences, University of South Australia, Adelaide, Australia

[3]Cancer Nursing Research Unit, University of Sydney, Sydney, Australia

[4]Research, Evaluation, and Social Policy Team, CanTeen Australia, Sydney, Australia

[*]all authors contributed equally

Corresponding Author:
Xiomara Skrabal Ross, BPsych (Hons)
Cancer Research Institute
University of South Australia
GPO Box 2471
Adelaide, 5001
Australia
Phone: 61 0410900975
Email: xioyen@gmail.com

Abstract

Background: Adherence to oral chemotherapy is crucial to maximize treatment outcomes and avoid health complications in cancer patients. Mobile phones are widely available worldwide, and evidence that this technology can be successfully employed to increase medication adherence for the treatment of other chronic diseases (eg, diabetes) is well established. However, the extent to which there is evidence that mobile phone–based interventions improve adherence to oral chemotherapy is unknown.

Objective: This scoping review aims to explore what is known about mobile phone–delivered interventions designed to enhance adherence to oral chemotherapy, to examine the reported findings on the utility of these interventions in increasing oral chemotherapy adherence, and to identify opportunities for development of future interventions.

Methods: This study followed Arksey and O'Malley's scoping review methodological framework.

Results: The review search yielded 5 studies reporting on 4 interventions with adults (aged >18 years) diagnosed with diverse cancer types. All interventions were considered acceptable, useful, and feasible. The following themes were evident: text messages and mobile apps were the main methods of delivering these interventions, the 2 most commonly employed oral chemotherapy adherence–enhancing strategies were management and reporting of drug-related symptoms and reminders to take medication, the importance of stakeholders' engagement in intervention design, and the overall positive perceptions of delivery features. Areas for future research identified by this review include the need for further studies to evaluate the impact of mobile phone–delivered interventions on adherence to oral chemotherapy as well as the relevance for future studies to incorporate design frameworks and economic evaluations and to explore the moderator effect of high anxiety, poor baseline adherence, and longer time taking prescribed drug on adherence to oral chemotherapy.

Conclusions: Despite the increasing body of evidence on the use of mobile phones to deliver medication adherence–enhancing interventions in chronic diseases, literature on the oral chemotherapy context is lacking. This review showed that existing interventions are highly acceptable and useful to cancer patients. The engagement of stakeholders as well as the use of a design framework are important elements in the development of mobile phone–delivered interventions that can be translated into oncology settings.

(JMIR Mhealth Uhealth 2018;6(12):e11724) doi:10.2196/11724

KEYWORDS

medication adherence; antineoplastic agents; neoplasms; cell phone; text messaging; mobile apps; review; mHealth

Introduction

Background

The widespread increase in the use of chemotherapy delivered via the oral route is transforming oncology. However, self-administration of oral chemotherapy encompasses a number of challenges for patients and health professionals to ensure adequate management of adherence and toxicities [1]. Nonadherence can reduce treatment efficacy and lead to dangerous health complications, including death [2]. The rates of adherence to oral antineoplastic agents can be as low as 46% [3]. Despite this fact, most health institutions do not practice standardized patient monitoring procedures for adherence [4].

Adherence is defined as the extent to which a persons' medication-intake behavior corresponds with the agreed recommendations of the clinician [5]. Adequate oral medication adherence is also important for the optimal treatment of other chronic conditions (eg, diabetes and HIV). Due to the long-term nature of these diseases, adherence and monitoring are required over long periods, which can be problematic. Technology is increasingly being used to help chronically ill patients adhere to their treatment regimens [6]. Mobile phones are a technological platform that allows delivery of behavioral interventions, assessments, and real-time data collection [7] and, importantly, can also facilitate access to support patients who, due to their remote geographical location or limited mobility, cannot access face-to-face services. Mobile text messages (short message service, SMS) and mobile apps are 2 types of mobile phone–based technology that are most commonly used to support patients with chronic diseases [6].

Worldwide availability of mobile phones is extensive and ownership of these technologies will continue to grow. As a result, there is great potential to use mobile technology to improve health care delivery. In 2016, 62.9% of the world population (4.65/7.40 billion) owned a mobile phone and this figure is set to increase to 67% (5.16/7.71 billion) in 2019 [8]. The introduction of smartphones means that mobile phones are no longer limited to being a tool for calls and text messages but also allow internet connectivity. Mobile devices (including smartphones and tablets) are currently the main source of internet connection. In 2017, approximately three-fourths of the worldwide internet access occurred through mobile devices [9]. In Australia, in 2016, 84% of the population (approximately 16 million people) owned a smartphone. The only places in the world where uptake of smartphones is greater are South Korea, the Netherlands, and Norway [10].

Previous research has already established that interventions delivered via mobile phones can significantly improve medication adherence for people with arterial hypertension [11], heart failure [12], and diabetes [13,14]. Moreover, acceptability and usefulness of mobile phone–delivered interventions are known to be high among chronically ill patients [6]. However, it is important to note that the efficacy of adherence-enhancing interventions is determined by the quality of the strategies delivered by a form of technology. Gains should not be simply attributed to the type of technology employed.

Although the reach, popularity, and many technological features of mobile phones now mean they are ideal platforms to provide health care support to cancer patients undergoing oral chemotherapy, because the widespread use of oral chemotherapy drugs is relatively new, the extent to which evidence is available to support this strategy is unknown.

Objectives

To address the emerging issue of oral chemotherapy nonadherence, this scoping review aims to:

- explore what is known about oral mobile phone–delivered interventions designed to enhance adherence to oral chemotherapy,
- examine the reported findings on the utility of mobile phone–delivered interventions in increasing adherence to oral chemotherapy, and
- identify opportunities for future development of oral chemotherapy adherence–enhancing interventions via mobile phone.

Methods

Overview of Methods

The scoping review methodological framework used in this review was outlined by Arksey and O'Malley [15]. This approach is ideal to understand research fields that are in early stages because it allows the rapid mapping of key concepts, sources, and evidence available, leading to identification of gaps in the existing literature [15]. This method aims to produce broad results from all relevant literature instead of trying to answer highly focused questions from specific study designs, as is the case in systematic reviews. Consistent with Arksey and O'Malley's framework [15], this study presents a narrative review of literature based on an analytic framework (thematic analysis) [16] and does not seek to assess the quality of studies, including risk of bias or generalizability of findings.

Identification of Relevant Studies

A structured database search was conducted in April 2018 with the following databases: MEDLINE, EMBASE, EMcare, and PsycINFO using terms related to 5 key areas: mobile phones, adherence, intervention, oral chemotherapy, endocrine therapy, and cancer. Subject headings and keywords used in databases and independent reviews were collated for each of the 5 key areas using the "OR" function and groups 1 to 5 were connected with the "AND" function. Examples of keywords included in each area were (1) mobile phone, text messaging, mHealth, mobile app; (2) patient compliance, medication compliance, medication adherence; (3) program, pilot, study, review, randomized controlled trial; (4) oral chemotherapy, antineoplastic agents, oral anti-cancer, endocrine therapy; and (5) neoplasm, tumour, cancer. Subject headings (eg, Medical Subject Headings) were employed. English language limits were applied, but no date restrictions were used for this search. Grey literature was searched through the ProQuest Dissertations and Theses database, limited to doctoral theses between January 2013 and March 2018 due to the high number of irrelevant results obtained with unlimited searching. To extend the results, an independent search in a Web journal took place. Reference

lists of relevant articles were also reviewed for references that may have been missed when conducting the database research.

Conference proceedings were included to make this review as broad and informative as possible. Titles and abstracts of retrieved documents were screened against inclusion criteria, followed by full text review of relevant studies.

Selection of Studies

To include relevant studies and associated content that contributed to meeting the objectives of this review, the following inclusion criteria were used to guide the screening process: (1) research-based studies on interventions that aim to increase adherence to oral chemotherapy or endocrine therapy, (2) targets cancer patients taking oral chemotherapy or endocrine therapy, (3) use of mobile phones as a main tool to deliver the intervention, and (4) articles written in English. In this study, adherence was defined as taking oral chemotherapy in accordance with the dose and frequency prescribed by the clinician.

Data Charting

Full text articles were assessed to extract relevant information, and this was transferred into an Excel spreadsheet. Information charted in this process was as follows: (1) authors, (2) study purpose, (3) research design (eg, qualitative and randomized controlled trial [RCT]), (4) participants (age, cancer diagnosis, oral chemotherapy or endocrine treatment, and country), (5) mobile phone features (eg, text messages and apps), (6) intervention (duration and key components of intervention), and (7) main findings (summary of the most relevant findings, including recommendations for future studies). A second reviewer assessed 50% (5/10) of articles to ensure validity of information extraction.

Thematic Analysis and Reporting of Results

Following the methodological framework from Arksey and O'Malley [15], thematic analysis [16] was used to identify what is already known about mobile phone–based oral chemotherapy adherence–enhancing interventions, the utility of interventions in improving oral chemotherapy adherence, and the opportunities for future development in this area. A second reviewer assessed all codes and a complete agreement was achieved. Due to the research designs used in the reviewed studies (most were nonexperimental), results are reported in category groups conformed by common themes in the reviewed literature. Categories are aligned with the aims of this research.

Results

Overview of Results

During the initial database search, 43 articles were retrieved. After removal of duplications, 29 unique publications were identified. Titles and abstracts were scanned for relevance against the inclusion criteria and those that did not match (eg, focus on another illness or lack of focus on improving oral chemotherapy adherence) were removed, leaving 10 articles for full text review. After full text review, 6 articles were excluded. An additional publication was added following an independent review in an online journal. No articles fitting the inclusion criteria were identified through the grey literature search in the ProQuest Dissertations and Theses database or through search in articles' reference lists (Figure 1 shows the detailed process of screening and inclusion). Ultimately, 5 articles [17-21] reporting 4 interventions were included in this review, 2 of which were abstracts from conference proceedings and 3 were peer-reviewed journal articles. Moreover, 1 study [19] describes the development of an intervention that was later tested in another one of the included articles [18]. Multimedia Appendix 1 shows the information extracted.

Search results showed that research on mobile phone–based interventions to increase adherence to oral chemotherapy began in 2015. The research designs of the studies were qualitative (3 out of 5 studies) [19-21] and experimental RCTs (2 out of 5 studies) [17,18]. All but 1 study explored the feasibility and acceptability of interventions, and 2 out of 5 studies used an RCT design to evaluate the effect of mobile phone interventions on adherence to oral chemotherapy [17,18].

Moreover, 3 out of the 5 studies had small samples (5-32 participants) [19-21]; the remaining 2 studies used an RCT design with larger samples (80 and 181 participants) to measure oral chemotherapy adherence as a primary outcome [17,18].

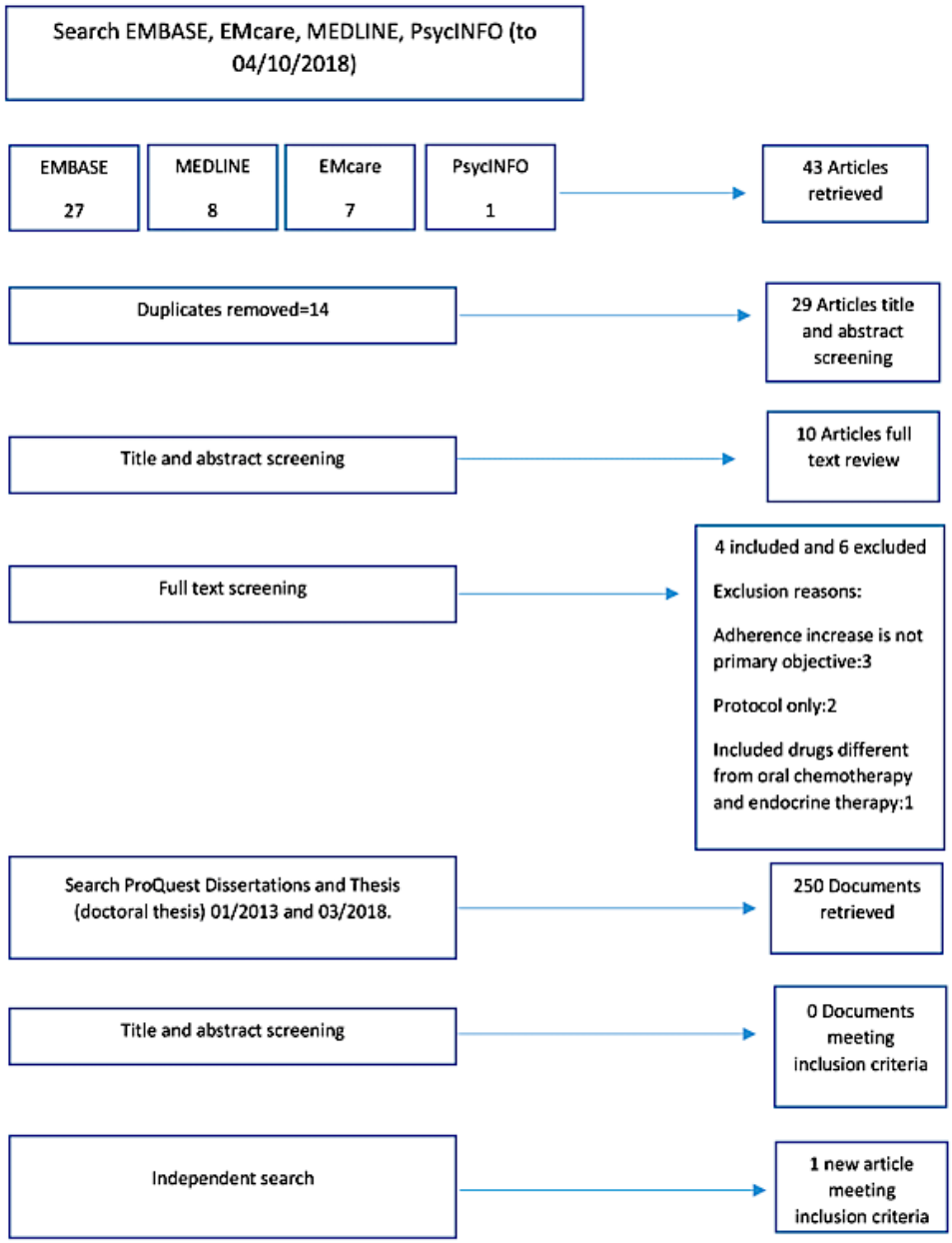
Of the 5 studies, 3 included participants with diverse types of cancer [17-19], 1 focused on patients with chronic myeloid leukemia [20], and 1 focused on breast cancer patients [21]. All the studies focused on adult participants (aged >18 years).

Strategies and Features of Intervention Delivery

Various methods of delivery and adherence-enhancing strategies were employed in the reviewed studies. In addition, 2 distinct methods were used to deliver the interventions: SMS text messages and mobile apps. The SMSs were sent as medication intake reminders and frequency was daily or twice daily according to individual medication intake schedules [17,20]. The SMS reminders were bidirectional (participants' response "Y" or "N" was expected) to collect data on the frequency of taking medication. The content of the SMS (message bank, wording, and theoretical framework) was reported by Spoelstra et al [17] and informed by social cognitive theory. The SMS content in Pereira-Salgado et al's study was not described [20]. For future research to understand the process of how SMS reminders influence adherence, it is important that these details (eg, content, frequency, and sender) are reported. Overall, participants showed high rates of response to SMS reminders. In 2 studies, they replied 87.13% (1036/1189) of the times [17] and only up to 22% (17/80) of the SMS were not replied when sent [20]. This not only shows that participants received and read the SMS but also actively participated in reporting medication-taking events.

The SMS reminders in the study by Pereira-Salgado et al were part of an intervention that also included a Web-based app. It recorded participants' side effects and severity and provided real-time self-management advice [20]. This intervention also incorporated nurse-led phone support consultations. This last strategy was not necessarily delivered via mobile phone.

Figure 1. Flowchart of study selection process and included studies.



Mobile apps were used as delivery features in 2 interventions [18,21]. They all included reminders to take medication. Daily reminders (alarms) were set by users according to their preferences in 1 study [19]. Brett et al's study did not report on any characteristics of medication reminders (eg, frequency and type) [21].

Other strategies used in mobile app interventions were real-time side effect management advice, cancer-specific information, medication-specific information, and chat forums (Table 1 shows features and strategies of each of the interventions). Automatic generation of reports accessible by research teams was another technical element in 1 mobile app and a combined Web app and SMS intervention. Reports contained data on symptom severity [19] or a combination of medication-taking and symptom severity [20].

The potential influence of oral chemotherapy-induced side effects on adherence to oral chemotherapy was recognized widely in the reviewed studies. Strategies to improve symptom management were incorporated in three-fourths of the interventions. These strategies included symptom reports with real-time reception of tailored feedback [18,21] and symptom recording in a side effect mobile app diary [21]. Symptom burden and severity was also 1 of the primary outcomes in most interventions (3 out of 4 studies) [17,18,20].

Table 1. Strategies and features of interventions.

Strategy or feature	Spoelstra et al, 2016 [17]	Greer et al, 2017 [18]	Pereira-Salgado et al, 2017 [20]	Brett et al, 2018 [21]
Strategy				
Medication reminders	✓	✓	✓	✓
Symptom management information	—[a]	✓	✓	✓
Cancer-specific information	—	✓	—	✓
Individual nurse phone support	—	—	✓	—
Medication-specific information	—	✓	—	✓
Recording of side effects	—	✓	✓	✓
Chat forum	—	—	—	✓
Feature				
SMS[b] text messages	✓	—	✓	—
Mobile app	—	✓	—	✓
Web-based app	—	—	✓	—

[a]Strategy or feature not present.

[b]SMS: short message service.

Acceptability, Usability, and Feasibility of Interventions

Overall, all interventions reviewed were found to be acceptable and useful, and their implementation was feasible. The definition of acceptability differed across studies. Spoelstra et al's study defined acceptability as the percentage of patients who agreed to take part in the study and the percentage of patients who completed the study [17]. This SMS medication reminder intervention had high rates of acceptability among cancer patients, with 75.7% (78/103) of eligible potential participants consenting to participate and 86% (42/49) of the initial participants completing the entire intervention.

Preliminary acceptability of ChemOtheRapy Assistant (CORA) mobile app [19] was described as the result of a 5-stage developmental process that led to intervention improvement. Stakeholders (eg, patients, and oncology clinicians) participated in qualitative research to provide information that led to the design of an acceptable final version of the mobile app.

Pereira-Salgado et al's REMIND system study assessed intervention acceptability by interviewing participants (patients and nurses) after using the system [20]. The intervention was highly acceptable to participants in terms of its content, timing, and perceived utility of each component (SMS medication reminders, symptom management advice, and nurse phone support). Usability was assessed separately in this study with participants expressing appreciation for the ease of use of text reminders and the weekly symptoms component, with the exception of a small number of participants who expressed delays in receipt of the SMS. The nurses also found the intervention useful but suggested changes to the number of report emails and layout of the intervention manual (ie, inclusion of tabs to facilitate the search for specific content).

The definitions of acceptability and usefulness in Brett et al's study were not provided [21]. This conference proceeding indicated that participants considered individual components of the mobile app intervention (information section, links to evidence around adjuvant endocrine therapy, side effects diary, and repeat prescription reminders) acceptable and useful. However, perceptions of the usefulness of the chat forum were mixed. Participants also suggested more information on side effect management strategies.

Measures of feasibility also varied across studies. Measures included the number of SMS delivered and returned [17] and possibility to implement and integrate the intervention into a clinical setting [20]. Brett et al's study, which aimed to explore the feasibility of a mobile app intervention, reported results in terms of high acceptability, usefulness, and usability of the intervention [21]. The CORA exploratory study [19] did not report on the intervention feasibility. However, its later implementation in clinical settings with 181 oral chemotherapy users can be considered confirmation of feasibility [18].

Stakeholders' Engagement in the Design of Interventions

Stakeholders' engagement was evidenced in the design of all the reviewed interventions, through the exploration of end users' perceptions of the acceptability and usefulness of mobile phone interventions. An example of this is the design of the CORA mobile app, which was informed at every stage by groups of stakeholders: patients and caregivers; oncology clinicians; cancer practice administrators; and representatives of the care system, community, and society [19]. Engagement of these groups allowed the intervention to be shaped by users' opinions on the need for oral chemotherapy self-management support; the role of the intervention in supporting oral chemotherapy

self-management, acceptability, and usability; as well as exploring possible implementation barriers [19].

Cancer patients also participated in early design phases of Pereira-Salgado et al's REMIND system for patients with chronic myeloid leukemia [20] and Brett et al's mobile app for women taking adjuvant endocrine therapy after breast cancer [21]. The first study also included oncology clinicians and explored their perceptions of the nature, extent, and reasons for nonadherence to tyrosine kinase inhibitors. The second study examined patients' preferences on the content of a mobile app [21]. Findings from both studies informed the strategies incorporated into mobile phone interventions, in line with patients' needs and preferences.

Spoelstra et al's study [17] assessed, among other variables, participants' acceptability and satisfaction with the intervention, and both were rated as high by study participants. These findings support the possibility of cancer patients to incorporate this text message intervention into their daily lives.

Design Framework Informing Development of Mobile Phone Interventions

Overall, 2 of the 4 interventions [19,20] reported the use of a design framework as a guide during the development process. In the design of the CORA mobile app [19], the investigators incorporated Whittaker et al's framework [22], which sets a process that involves steps to develop and test mobile phone–based health interventions. In doing this, the investigators based the design on a theoretical model, conducted formative research, pretested the intervention with stakeholders, and piloted the app with 5 participants enrolled in the next phase of the research, which was an RCT. Results from the experimental phase were reported in a separate study [18], and qualitative follow-up was also intended to be measured in the same study. However, due to the nature of the article (conference proceeding), detailed information on this was not provided.

Fishbein et al's CORA exploratory study [19] compared their intervention development process (a posteriori) with recommendations highlighted by Darlow and Wen's review [23], which recommends the adoption of 8 practices in the development of mobile phone interventions. In addition to the steps described above, user testing was conducted via qualitative methods, adequate time needed to test technology was anticipated, stakeholders were engaged in all steps of the intervention design, usability of the app to ensure the technology was simple and intuitive was assessed, the intervention's promotion of a sense of competence over patients' own care was explored, health professionals were consulted to ensure the use of the mobile app was not a burden to them, and the results of development and testing phase were published.

Schofield and Chambers's framework [24] specifies 7 features for the development of effective, clinically feasible, and sustainable interventions: (1) targeting cancer type and stage, (2) tailoring to unique individual needs, (3) promoting self-management, (4) efficient intervention delivery, (5) ensuring evidence-based and theoretical grounding, (6) specifying protocol training and adherence, and (7) confirming stakeholder acceptability. All the previously described steps were followed in the design of Pereira-Salgado et al's REMIND system study [20].

The importance of using a theoretical grounding in the design of mobile phone interventions was highlighted by this review. Murray et al's conceptual model [25], which provides a description of multidimensional factors affecting medication adherence, was used to inform the strategies used by the CORA mobile app [19]. Self-determination theory informed the use of motivational interviewing as part of the nurse phone support strategy in Pereira-Salgado et al's REMIND system study [20]. Self-determination is a theory of motivation that emphasizes the importance of supporting individuals' natural tendencies to exhibit healthy behaviors [26].

Social cognitive theory [27], more specifically self-efficacy, guided the content design of the SMS messages in Spoelstra et al's intervention [17]. According to the authors, messages were written using motivational content to stimulate the participants' engagement with SMS and behavior change. In Pereira-Salgado et al's REMIND system study [20], motivational interviewing provided as part of the nurse support was designed to stimulate participants' self-assessment of the problem as well as to help provide them with the information, resources, and skills needed to achieve oral chemotherapy adherence. The authors indicated that the nurse phone support strategy appeared to increase self-efficacy according to the analysis of participants' interviews, but this was only assessed qualitatively.

Utility of Mobile Phone Interventions in Increasing Adherence to Oral Chemotherapy

Due to research designs employed by the reviewed studies and the aims of this scoping review, numerical comparisons are not offered. This section provides a narrative approach to describe findings on the observed utility of mobile phone interventions in improving adherence to oral chemotherapy.

The 2 experimental studies in this review (Spoelstra et al's study and Greer's et al's CORA experimental study) [17,18] did not find statistically significant differences between the experimental and control groups. However, findings point toward patient and treatment variables (high levels of anxiety, poor baseline adherence, and length of treatment), which may moderate the effect of interventions on oral chemotherapy adherence. Participants who reported adherence problems at baseline showed better adherence after using the app than the standard care group (as measured by Medication Event Monitoring System) [18]. This study also found that participants with high levels of anxiety in the experimental group showed better adherence to oral chemotherapy than the standard care group at the end of the study (measured by Morisky Medication Adherence Scale). Spoelstra et al's study [17] found that participants in the experimental group showed better adherence than the control group in later weeks of the study (measured by SMS reply self-report).

Pereira-Salgado et al's REMIND system study described participants' perceived utility of the intervention in increasing adherence to tyrosine kinase inhibitors [20]. Most participants reported that reception and response to SMS reminders stimulated their medication adherence due to accountability (eg,

reinforcing habits at the beginning of treatment or drug intake support during time of routine change).

Although qualitative studies [19-21] were not designed to evaluate the effect of interventions on medication adherence, they constitute a necessary step in the development of acceptable, usable, and relevant interventions, which were also found useful to participants in supporting their oral chemotherapy intake.

Issues and Limitations Related to the Use of Mobile Technology

Failure to receive up to 40% of SMS on time was experienced by 2 out of the 9 participants who completed Pereira-Salgado et al's study because of slow networks in rural areas [20]. Technological difficulties and being without their mobile phone (eg, left at home and losing phone) were reported by some participants. These barriers seem difficult to overcome and should be taken into consideration at the time of designing interventions using mobile phones.

One limitation found in the use of the CORA mobile app [19] was the need to send symptom reports to clinicians via email instead of using the electronic health record system due to regulations. This method did not guarantee that clinicians would open the report emails when sent. Another limitation of the app was the support of only iPhone and Android phones, excluding other operating systems. The authors recognized the potential to include other smartphone operating systems to reach a broader population of smartphone users.

Brett et al's study did not describe limitations and issues related to the use of their mobile app [21]. This study was a conference proceeding, which can explain reduced information about the topic.

Discussion

Principal Findings

This scoping review brings together the available evidence on adherence-enhancing interventions delivered via mobile phone in the context of oral chemotherapy. A total of 5 studies describing 4 interventions met the inclusion criteria. This low figure may be because the widespread use of oral chemotherapy is a relatively new medical advancement, and the extended access to mobile phones, especially smartphones, is also a recent phenomenon, which can also explain the young data of studies in this area.

Consistent with trends in other chronic diseases [6], this review shows that the 2 main features used to deliver mobile phone interventions aiming to increase oral chemotherapy adherence are SMS and mobile apps. Regardless of the technology feature employed, all interventions explored were highly acceptable, useful, and feasible to be implemented in clinical settings.

Despite the variety of adherence-enhancing strategies in the interventions, 2 strategies were common to most studies: drug-related symptom management advice and reporting and medication-intake reminders. This approach is compatible with evidence on drug-related symptoms and forgetfulness as the 2 main barriers to oral chemotherapy adherence [33]. It is important to notice that although all the reviewed interventions primarily addressed adherence barriers related to the patient (forgetfulness, knowledge of therapy, and condition) and the therapy (side effects), only half of those interventions took into account health care team and system-related barriers (communication with treating team, monitoring of adherence, and side effects). Strategies to address these barriers consisted of reports on the presence and severity of side effects and adherence frequently sent to treating teams to stimulate prompt communication and adequate monitoring of oral chemotherapy treatment as required [18,20].

The relevance of patients' involvement in the design, implementation, and evaluation of health research has been widely recognized [34]. The use of a participatory research model allows generation of more significant research questions, alignment of intervention goals with end users' needs, increase in the acceptability and usability of health interventions, and enhancement of translation of findings into real-life settings. This was acknowledged by most studies in this review whose intervention strategies were shaped by stakeholders' perceptions of barriers to adherence, need for self-management support, or their preferences on components of the interventions. It was also generally recognized that interventions need to be found to be acceptable, useful, and usable to stakeholders before moving toward experimental research phases.

This review showed that there were no established processes for the development of mobile phone health interventions. Some researchers did not use or at least did not report the utilization of a mobile phone intervention design framework, including theoretical grounding. Following a framework to design mobile phone–based adherence-enhancing interventions in the oral chemotherapy context supports the development of acceptable interventions that are of intuitive and relevant use to cancer patients. Overall, the use of design frameworks can help to adequately plan the resources needed in each stage of the design as well as to canalize these assets into tools that can be successfully implemented in oncology settings.

Mobile phone health intervention design frameworks in this review also highlight the need to develop interventions based on a theoretical approach. This is crucial as technology alone cannot be seen as a strategy to increase medication adherence. Although most reviewed studies reported a theoretical framework informing their design, some inconsistencies were found in the explanation of the theoretical elements of the interventions, for example, the use of the term "motivation" alone to describe self-efficacy-informed SMS content or intervention strategies. According to Bandura's self-efficacy theory [27], individuals' levels of motivation are heavily based on their beliefs in their capacity to display behaviors that will impact events affecting their lives. Therefore, motivation alone may not be enough to explain the influence of self-efficacy-based interventions on medication adherence. Moreover, in the context of self-efficacy theory, patients' success in adhering to oral chemotherapy is the most effective source informing patients of their ability to follow their drug prescriptions. It is crucial for self-efficacy-based interventions to describe the process through which self-efficacy (as a

construct influenced by multiple elements) is expected to support cancer patients to achieve adherence to oral chemotherapy. Without this explanation, motivation remains an isolated variable that cannot be linked to self-efficacy.

Overall, general perceptions of mobile phone technologies in this review were positive. As an example, SMS had high rates of delivery and response success, presenting this mobile phone feature as one that is able to be successfully implemented in clinical settings. However, the use of SMS and other mobile phone features encompass challenges that are not easy to overcome. Patients who live remotely, with poor internet or phone coverage, are prone to miss medication reminders or experience issues accessing mobile apps. Patients may also lose their mobile phones or leave them at home, missing the opportunity to benefit from real-time interventions at times. At the time of reporting results related to adherence-enhancing interventions delivered via mobile phones, it is important for authors to describe strengths, limitations, and barriers in the use of mobile phone technology. This will help to inform future researchers on the obstacles and advantages of delivery features when designing interventions of this type.

A strength of this review is its novelty as it is the first study to examine the current state of knowledge about oral chemotherapy adherence–enhancing interventions delivered via mobile phones and to identify opportunities for future research in the area. Another strength of this study is the use of a methodological framework for scoping reviews, which increases consistency and structure of the search process and reporting of findings. In addition, reliability of the search strategy was increased by involving a research librarian in the process.

Limitations

Scoping reviews include a variety of study types to answer broad research questions by mapping available evidence and identifying knowledge gaps. The purpose of scoping reviews is not to ask highly focused research questions or to assess the quality of the reviewed literature, as is the case in systematic reviews. Due to the variety of research designs in the reviewed studies, quantitative analyses on available data were not possible.

The search strategy in this study was limited to research published in English, which may have led to the omission of other sources of information.

Furthermore, this review was able to incorporate and analyze only those studies available at the time of the search that fit the inclusion criteria. Our search yielded results showing 1 study protocol that despite meeting most of the inclusion criteria, was not research-based at the protocol stage [35].

Opportunities for Future Research

This review provides evidence of the scarcity on studies that evaluate the effect of mobile phone interventions on adherence to oral chemotherapy. Mobile phone interventions in this review were highly acceptable and useful to oral chemotherapy users. Therefore, there is a need for future research to take the next steps into experimental studies to generate evidence-based knowledge that has the potential to be translated into oncology settings.

In other chronic diseases, the use of SMS medication reminders has proven to be effective in increasing medication adherence [28]. It would be useful if future studies carefully described the key elements of SMS reminders used in interventions (eg, content, frequency, and sender) so that researchers are able to determine which elements are most likely to have an impact.

It is possible for SMS reminder interventions to incorporate content grounded on evidence-based theoretical models that encourage behavior change. In addition, due to internet accessibility, use of smartphones enables text message interventions to deliver not only medication reminders but also larger contents of information addressing additional barriers to oral chemotherapy (eg, education).

Despite a general failure of studies to report cost-effectiveness analysis of mobile phone adherence tools [6], the design of these types of interventions may involve elevated costs in time, human, and financial resources. Due to this consideration and the need for adherence-enhancing tools to be translatable to real oncology settings, future research could benefit from following a mobile phone health intervention design framework and the inclusion of economic analysis.

According to the World Health Organization, patients in developing countries face a number of health care barriers (eg, short staffed hospitals, lack of patient access to care, and long waiting times to see a doctor) [29], which may increase the chances of oral chemotherapy nonadherence. It is estimated that in developing countries in 2015, one-third of people owned a smartphone, and this figure is set to increase to approach the ownership rates in developed countries in the next few years [30]. Studies in this review focus exclusively on patients living in developed countries. It would also be useful to explore the impact of such interventions in developing countries.

Although the scope of this review was not limited to adults, the body of literature included in this study only targeted cancer patients older than 18 years. Evidence shows that adolescents and young adults are at higher risk of oral chemotherapy nonadherence than younger and older users [31]. The use of mobile phones among adolescents and young adults is even higher than that among adults. In Australia, in 2015, 9 in 10 teenagers (aged 14-17 years) owned a mobile phone [32]. Therefore, it is important that future research also addresses the nonadherence of younger oral chemotherapy users via mobile phone–delivered interventions.

More studies on the moderator effect of anxiety, poor baseline adherence, and length of treatment would be beneficial to understand the role those variables play on oral chemotherapy adherence interventions delivered via mobile phone.

Conclusions

This review shows the lack of research on oral chemotherapy adherence–enhancing interventions delivered via mobile phone. Available interventions, delivered via SMS and mobile apps, are highly acceptable and useful to oral chemotherapy users, and nonadherence in this group is a serious issue. These findings

support the need for the development and evaluation of mobile phone tools to assist cancer patients to follow their oral chemotherapy prescriptions. This review also highlighted the importance of stakeholders' involvement and the use of a design framework in the development of mobile phone–based interventions aiming to support oral chemotherapy intake to increase translatability into real oncology practices. Given the increasing use of oral chemotherapy and the widespread availability of mobile phones worldwide, further research in this field is expected to rapidly increase in the near future.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Included studies.

[PDF File (Adobe PDF File), 280KB - mhealth_v6i12e11724_app1.pdf]

References

1. Weingart SN, Brown E, Bach PB, Eng K, Johnson SA, Kuzel TM, et al. NCCN task force report: oral chemotherapy. J Natl Compr Canc Netw 2008 Mar;6(Suppl 3):S1-14. [Medline: 18377852]
2. Makubate B, Donnan PT, Dewar JA, Thompson AM, McCowan C. Cohort study of adherence to adjuvant endocrine therapy, breast cancer recurrence and mortality. Br J Cancer 2013 Apr 16;108(7):1515-1524 [FREE Full text] [doi: 10.1038/bjc.2013.116] [Medline: 23519057]
3. Greer JA, Amoyal N, Nisotel L, Fishbein JN, MacDonald J, Stagl J, et al. A systematic review of adherence to oral antineoplastic therapies. Oncologist 2016 Mar;21(3):354-376. [doi: 10.1634/theoncologist.2015-0405] [Medline: 26921292]
4. Weingart SN, Li JW, Zhu J, Morway L, Stuver SO, Shulman LN, et al. US cancer centre implementation of ASCO/Oncology Nursing Society chemotherapy administration safety standards. J Oncol Pract 2012 Jan;8(1):7-12 [FREE Full text] [doi: 10.1200/JOP.2011.000379] [Medline: 22548004]
5. Sabate E. World Health Organization. 2013. Adherence to long-term therapies: evidence for action URL: http://www.who.int/chp/knowledge/publications/adherence_report/en/ [accessed 2018-05-29] [WebCite Cache ID 73rcDuJSB]
6. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
7. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
8. Statista. 2016. Number of mobile phone users worldwide from 2015 to 2020 (in billions) URL: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/ [accessed 2018-05-29] [WebCite Cache ID 6zljBqeWV]
9. Zenith Media. 2018. Smartphone penetration to reach 66% in 2018 URL: https://www.zenithmedia.com/smartphone-penetration-reach-66-2018/ [accessed 2018-05-29] [WebCite Cache ID 6zljnuKM1]
10. Deloitte. 2016. Mobile consumer survey, the Australian cut hyper connectivity: clever consumption URL: http://landing.deloitte.com.au/rs/761-IBL-328/images/tmt-mobile-consumer-2016-final-report-101116.pdf [accessed 2018-05-29] [WebCite Cache ID 6zlkG5uFc]
11. Kiselev AR, Gridnev VI, Shvartz VA, Posnenkova OM, Dovgalevsky PY. Active ambulatory care management supported by short message services and mobile phone technology in patients with arterial hypertension. J Am Soc Hypertens 2012;6(5):346-355. [doi: 10.1016/j.jash.2012.08.001] [Medline: 22995803]
12. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
13. Lim S, Kang SM, Shin H, Lee HJ, Won YJ, Yu SH, et al. Improved glycemic control without hypoglycemia in elderly diabetic patients using the ubiquitous healthcare service, a new medical information system. Diabetes Care 2011 Feb;34(2):308-313 [FREE Full text] [doi: 10.2337/dc10-1447] [Medline: 21270188]
14. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
15. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

17. Spoelstra SL, Given CW, Sikorskii A, Coursaris CK, Majumder A, DeKoekkoek T, et al. Proof of concept of a mobile health short message service text message intervention that promotes adherence to oral anticancer agent medications: a randomized controlled trial. Telemed J E Health 2016 Dec;22(6):497-506 [FREE Full text] [doi: 10.1089/tmj.2015.0126] [Medline: 26716365]
18. Greer J, Jacobs JM, Pensak N, Nisotel LE, Fishbein J, MacDonald J, et al. Randomized trial of a smartphone mobile app for adherence to oral chemotherapy. J Clin Oncol 2017 May 20;35(15_suppl):10055-10055. [doi: 10.1200/JCO.2017.35.15_suppl.10055]
19. Fishbein JN, Nisotel LE, MacDonald JJ, Amoyal PN, Jacobs JM, Flanagan C, et al. Mobile application to promote adherence to oral chemotherapy and symptom management: a protocol for design and development. JMIR Res Protoc 2017 Apr 20;6(4):e62 [FREE Full text] [doi: 10.2196/resprot.6198] [Medline: 28428158]
20. Pereira-Salgado A, Westwood JA, Russell L, Ugalde A, Ortlepp B, Seymour JF, et al. Mobile health intervention to increase oral cancer therapy adherence in patients with chronic myeloid leukemia (The REMIND System): clinical feasibility and acceptability assessment. JMIR Mhealth Uhealth 2017 Dec 06;5(12):e184 [FREE Full text] [doi: 10.2196/mhealth.8349] [Medline: 29212628]
21. Brett J, Watson E, Boulton M, Stoner N. Exploring the potential for a smart phone app to improve adherence to adjuvant endocrine therapy (AET) in women after treatment for breast cancer. Psycho-Oncol 2018 Mar 01;27:9-22. [doi: 10.1002/pon.4639]
22. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17(Suppl 1):11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
23. Darlow S, Wen K. Development testing of mobile health interventions for cancer patient self-management: a review. Health Informatics J 2016 Dec;22(3):633-650. [doi: 10.1177/1460458215577994] [Medline: 25916831]
24. Schofield P, Chambers S. Effective, clinically feasible and sustainable: Key design features of psycho-educational and supportive care interventions to promote individualised self-management in cancer care. Acta Oncol 2015 May;54(5):805-812. [doi: 10.3109/0284186X.2015.1010016] [Medline: 25813474]
25. 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]
26. Self-Determination Theory. URL: http://selfdeterminationtheory.org/ [accessed 2018-05-29] [WebCite Cache ID 6zloaDZSr]
27. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037/0033-295X.84.2.191]
28. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease. A meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
29. Smith MK. Facing the heatlh worker crises in developing countries: a call for global solidarity. Bull World Health Organ 2006 Jun 01;84(6):426-426 [FREE Full text] [doi: 10.2471/BLT.06.031757]
30. Poushter J. Pew Research Center. 2016. Emerging, developing countries gain ground in the tech revolution URL: http://www.pewresearch.org/fact-tank/2016/02/22/key-takeaways-global-tech/ [accessed 2018-05-30] [WebCite Cache ID 6znHCYdcn]
31. Partridge AH, Avorn J, Wang PS, Winer EP. Adherence to therapy with oral antineoplastic agents. J Natl Cancer Inst 2002 May 1;94(9):652-661 [FREE Full text] [Medline: 11983753]
32. Roy Morgan. 2016. 9 in 10 Aussie teens now have a mobile (and most are already on to their second or subsequent handset) URL: http://www.roymorgan.com/findings/6929-australian-teenagers-and-their-mobile-phones-june-2016-201608220922 [accessed 2018-05-30] [WebCite Cache ID 6znHsPFxT]
33. Verbrugghe M, Verhaeghe S, Lauwaert K, Beeckman D, Van Hecke A. Determinants and associated factors influencing medication adherence and persistence to oral anticancer drugs: a systematic review. Cancer Treat Rev 2013 Oct;39(6):610-621. [doi: 10.1016/j.ctrv.2012.12.014] [Medline: 23428230]
34. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014;14:89 [FREE Full text] [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
35. Passardi A, Rizzo M, Maines F, Tondini C, Zambelli A, Vespignani R, et al. Optimisation and validation of a remote monitoring system (Onco-TreC) for home-based management of oral anticancer therapies: an Italian multicentre feasibility study. BMJ Open 2017 May 29;7(5):e014617 [FREE Full text] [doi: 10.1136/bmjopen-2016-014617] [Medline: 28554917]

Abbreviations

CORA: ChemOtheRapy Assistant
RCT: randomized controlled trial
SMS: short message service

Edited by G Eysenbach; submitted 29.07.18; peer-reviewed by J Greer, B Given; comments to author 30.08.18; revised version received 03.10.18; accepted 08.10.18; published 21.12.18.

Please cite as:
Skrabal Ross X, Gunn KM, Patterson P, Olver I
Mobile-Based Oral Chemotherapy Adherence–Enhancing Interventions: Scoping Review
JMIR Mhealth Uhealth 2018;6(12):e11724
URL: http://mhealth.jmir.org/2018/12/e11724/
doi:10.2196/11724
PMID:30578182

Original Paper

Developing Typologies of User Engagement With the BRANCH Alcohol-Harm Reduction Smartphone App: Qualitative Study

Joanna Milward[1], MSc; Paolo Deluca[1], PhD; Colin Drummond[1], MD, FRCPsych; Andreas Kimergård[1], PhD

Addictions Department, King's College London, London, United Kingdom

Corresponding Author:
Joanna Milward, MSc
Addictions Department
King's College London
Addictions Sciences Building
4 Windsor Walk
London, SE58BB
United Kingdom
Phone: 44 02087 848 0437
Email: joanna.milward@kcl.ac.uk

Abstract

Background: Understanding how users engage with electronic screening and brief intervention (eSBI) is a critical research objective to improve effectiveness of app-based interventions to reduce harmful alcohol consumption. Although quantitative measures of engagement provide a strong indicator of how the user engages with an app at the group level, they do not elucidate finer-grained details of how apps function from an individual, experiential perspective and why, or how, users engage with an intervention in a particular manner.

Objective: The aim of this study was to (1) understand why and how participants engaged with the BRANCH app, (2) explore facilitators and barriers to engagement with app features, (3) explore how the BRANCH app impacted drinking behavior, (4) use these data to identify typologies of users of the BRANCH app in terms of engagement behaviors, and (5) identify future eSBI app design implications.

Methods: In total, 20 one-to-one semistructured telephone interviews were conducted with participants recruited from a randomized controlled trial, which evaluated the effectiveness of engagement-promoting strategies in the BRANCH app targeting harmful drinking in young adults (aged 18-30 years). The topic guide explored users' current engagement levels with existing health promotion apps, their views toward the effectiveness of such apps, and what they liked and disliked about BRANCH, specifically focusing on how they engaged with the app. Framework analysis was used to develop typologies of user engagement.

Results: The study identified 3 typologies of engagers. *Trackers* were defined by their motivations to use health-tracking apps to monitor and understand quantified self-data. They did not have intentions necessarily to cut down and predominantly used only the drinking diary. *Cut-downers* were motivated to use the app because they wanted to reduce their alcohol consumption Unlike *Trackers*, they did not use a range of different health apps daily, but saw the BRANCH app as an opportunity to test out a different method of trying to cut down their alcohol use. This typology used more features than Trackers, such as the goal setting function. *Noncommitters* were characterized as a group of users who were initially enthusiastic about using the app; however, this enthusiasm quickly waned and they gained no benefit from it.

Conclusions: This was the first study to identify typologies of user engagement with eSBI apps. Although in need of replication, it provides a first step in understanding independent categories of eSBI users, who may benefit from apps tailored to a user's typology or motivation. It also provides new evidence to suggest that apps may be used more effectively as a tool to raise awareness of drinking, instead of reducing alcohol use, and be a step in the care pathway, identifying at-risk individuals and signposting them to more intensive treatment.

Trial Registration: International Standard Randomised Controlled Trial Number ISRCTN70980706; http://www.isrctn.com/ISRCTN70980706 (Archived by WebCite at http://www.webcitation.org/73vfDXYEZ)

(JMIR Mhealth Uhealth 2018;6(12):e11692) doi:10.2196/11692

KEYWORDS

young adults; binge drinking; drinking; smartphone; mobile phone; mHealth

Introduction

Background

User engagement rates with app-based electronic and screening interventions (eSBI) are consistently reported as lower than expected in app-based eSBI trials [1,2]. A recent study of the Drinkaware app demonstrated that less than 50% of those who downloaded the app used it for more than a week, and only 14% were using it after a month [1]. Engagement with eSBI is critical because studies demonstrate an association between the level of a user's engagement and the effectiveness of the Web-based behavior change intervention [1,3-5]. Understanding how users engage with eSBI is a key research objective to improve effectiveness of app-based interventions to reduce harmful alcohol consumption [6-8].

User engagement with eSBI apps is a complex and multifaceted behavior, which requires nuanced exploration of the subtleties of how individuals interact with the electronic health (eHealth) intervention [9-12]. Engagement has been conceptualized in different ways, both in terms of how the user interacts with the technology as well as how the user experiences it [9]. The interaction may be simple such as visiting a particular feature of the eHealth intervention or completing a more complex task such as filling out a diary. Engagement can include how long or how often the participant uses the eHealth intervention [13,14].

Engagement has also been defined qualitatively. From this perspective, it is characterized as a *state* of involvement with the technology, such as becoming fully absorbed in the process. Outside of the alcohol field, research has examined *typologies* or *trajectories* of user engagement that enable distinctions to be made between different types of user behavior, which has the potential to target more tailored interventions to users [15]. Yardley et al [12] make a useful distinction of conceptualizing engagement at the micro and macro level. The micro level reflects the moment-to-moment interactions that occur as a user engages with features of the technology, whereas the macro-level engagement refers to how the user engages with the overall behavior change goal.

Although quantitative measures of log-ins or page visits measured in previous eSBI app trials provide a strong indicator of how the user engaged with the app at a group level, they neither elucidate finer-grained details of aspects of why users engaged with the app in a particular manner nor provide information on features that worked well, and those that did not, from an individual experiential perspective [9-12].

A qualitative exploration allows for examination of engagement, which puts the characteristics of the individual user experiences as the primary focus, and helps to contextualize the findings of the quantitative usage patterns to understand why users engaged as they did. [11,12]. It is argued that measures of engagement that rely purely on usage data do not provide a comprehensive measure of the subtleties of engagement behavior [11,12,15]. For example, quantitative measures of feature usage may identify those that were most frequently used, but not why they were used more than others, or what impact they had. Therefore, qualitative investigations of engagement are required to fill this gap in the existing knowledge of engagement.

There are various models of user engagement with digital interventions. O'Brien and Toms' [16] well-established conceptual model of user engagement (CMUE) is distinguishably qualitative than other models as they conceptualize engagement as a *quality of user experiences* characterized by a series of attributes such as esthetics and novelty. They summarize these attributes comprising *threads of experience* through the different stages of engagement. The *threads of experience* comprise sensual, emotional, and spatiotemporal categories, which are particularly useful in conceptualizing engagement from a qualitative perspective. Sensual experience refers to esthetic or novel elements, which promote attention and interest to the program, and emotional experience refers to the positive or negative affect elicited from the program. Finally, spatiotemporal experiences refer to the perceptions and awareness of time and physical surroundings the user feels when using the program, for example, experiencing a state of *flow* or the fast passing of time.

Short et al [17] propose the first evidence-informed model, specifically dedicated to the issue of engagement with eHealth interventions. The model considers the relationships between the individual (user characteristics), environment (social and physical characteristics), and intervention-level factors (content and features) that contribute to a user's engagement with an eHealth intervention. At the core of the model is the concept of tailoring interventions, with the aim of producing an intervention that is relevant, novel, appealing, and motivating to the user. According to the Elaboration Likelihood Model, people are more motivated to process information elaborately, leading to long-lasting effects, if the message is personally relevant or *tailored* to them [17]. The model brings together theory from O'Brien and Toms' [16] CMUE as well as Oinas-Kukkonen and Harijuanas' persuasive systems design (PSD) [17] (see Ryan et al [10] for a comprehensive scoping review of engagement theory).

Surprisingly, examination of engagement within the eSBI field from a qualitative research perspective is somewhat limited. Quantitative measures of engagement, such as log-ins or time spent on the Web, are more typical across the literature, and qualitative measures have been largely ignored. Studies have examined *usability* with a qualitative research design in the alcohol field [6,18]; however, *usability* typically refers to a single factor that influences engagement at the interaction level with the program and does not encompass the full breadth of elements that comprise engagement [9,10,12,19].

A few studies from the wider health care literature have directly explored engagement qualitatively, for example, in the smoking cessation and weight loss fields [15,20]. The most comprehensive study is by Smith et al [15] who examined

trajectories of engagement and disengagement for a *story-based* smoking cessation app with an interview-based method. The app presented quitters' stories and allowed users to read and post content. They proposed a new *conceptual trajectories'* model of different types of engagement, ranging from productive to counterproductive engagement as well as productive and nonproductive disengagement. Productive engagement is the desired state of use where users are highly invested in the smoking cessation tool and fully identify with the program components. Counterproductive engagement and nonproductive disengagement referred to experiences that had an effect opposite to that which was intended such as antipathy toward the program and even increases in smoking level. Productive disengagement, on the other hand, occurred when users positively engaged with the intervention and then went on to make a quit attempt. The trajectories model has implications in terms of how disengagement is considered and the level of control a user should be given over how they use the technology. Of particular note is that disengagement is not necessarily always an undesired behavior, and users should be free to disengage with the app if they feel it has had a positive impact. The model has, however, only been examined with a smoking cessation population, and although potentially applicable across other substances such as alcohol, the model has not been applied across drug and alcohol substances.

As is evident, few studies have qualitatively explored engagement with health apps, and none has comprehensively examined engagement with a qualitative research design with eSBI apps. In addition to the many quantitative studies exploring engagement using simple server and usage data, rich and detailed explorations of engagement from a qualitative perspective, such as the trajectories study by Smith et al [15] described above, are required to meaningfully understand the complexities and subtleties that engagement encompasses and elucidate insights and contextual factors that are beyond the abilities of a purely quantitative design.

Developing qualitative *typologies* of user engagement enhances existing quantitative typologies developed in the broader health behavior change field and contributes to the literature aiming to tailor apps to specific user types, which is a well-established engagement-promoting technique.

Objectives

The aim was to explore participants' experiences of engaging with the BRANCH app over a 28-day period to (1) understand why and how participants engaged with the BRANCH app, (2) understand facilitators and barriers to engagement with the app features and intervention goals, (3) explore how the BRANCH app may have impacted drinking behavior, (4) identify from these data different typologies of users in terms of engagement behaviors, and (5) identify future eSBI app design implications.

Methods

Study Context

This qualitative study was part of a larger research project, which developed an eSBI smartphone app (called BRANCH), aimed at promoting engagement, targeting harmful drinking in young adults, and evaluating it with a mixed-methods approach including both an randomized controlled trial (RCT) and qualitative interviews with trial participants. The qualitative study is the focus of this paper. For context, the RCT evaluated the effectiveness of a comprehensive version of the BRANCH app, including novel, innovative engagement-promoting strategies (EPSs, intervention arm) to increase app engagement as measured by user log-ins compared with a basic version (control arm), including minimal EPSs, which included only established screening and brief intervention techniques [21]. The RCT (Milward et al, in preparation) also measured change in harmful drinking between arms. See the following publications for further detail on the BRANCH app and a description of the development phases [6-8].

Design

One-on-one semistructured telephone interviews were conducted to explore individual user experiences with the app in the context of their daily lives. Individual interviews provided the opportunity for an in-depth understanding of the individual perspective around which the research objectives were situated [22]. In addition, exploration of how participants used the app to monitor their own drinking was more appropriately conducted in a one-on-one interview design compared with a focus group setting to minimize the potential for participants feeling uncomfortable about discussing sensitive topics about harmful drinking.

Sample

A purposive sample of 20 participants who took part in the BRANCH app RCT was recruited, with approximately equal numbers of high and low engagers from either the comprehensive or basic app group. On the basis of the analysis methods outlined in the protocol of the RCT (ISRCTN registration number 70980706), high engagers were classified as having logged in more than once and low engager as having logged in once. Sampling for interviews occurred across the first 3 months of recruitment in the RCT (January 2017 to March 2017).

In total, 20 participants were recruited to achieve a level of representational generalization [22] to uncover the breadth and nature of the views and experiences of the participants, which reflected those of the wider population from the RCT. In qualitative research, generalization cannot be achieved statistically but instead in terms of reaching saturation in the data [22]. Moreover, 20 participants were also selected as this would include 10 high engagers and 10 low engagers across both trial arms.

Recruitment

Participants were considered eligible 28 days after randomization into the trial so as to not bias the primary outcome of the RCT, which was collected at day 28 post randomization. Participants' details from the RCT were extracted from the Web-based trial management system. Participants were categorized by app version (comprehensive vs basic) and whether they were high or low engagers. Eligible participants were contacted by email in blocks of 20, stratified by month of RCT recruitment and engagement type. Blocks were selected

by date, starting with participants randomized on the first day of the month. Interviews were allocated on a first-come-first-serve basis. If no interviews were completed after the first block, consecutive blocks of 20 participants were emailed until the quota for each month and engagement level were achieved. Of the 211 RCT participants who were emailed, 30 replied requesting for an interview.

Description of the BRANCH App

The BRANCH app was developed in 3 stages: a systematic review of EPSs for Web-based substance use interventions [7]; scoping focus groups to determine user preferences for content, features, and style [8]; and usability testing on the prototype BRANCH app with a sample of the target population [6]. From the systematic review, 3 EPSs were identified that may promote engagement, tailoring, reminders, and delivery strategies. From the scoping focus groups, 2 main themes were identified. The *meaningfulness* theme reflected how young adults thought apps needed to be tailored to the interests and values of their age group, particularly emphasizing on content and feedback about broader health and well-being factors such as exercise, diet, and image. The *community* theme suggested that young adults wanted to be able to engage with other app users both in groups of friends and with Web-based users for motivation and support. From the usability testing, an easy-to-use interface with minimum required user-input was a critical usability issue for young adults. Clear, consistent, and visually appealing design was integral to the level of usability. The option for social connectivity was important, as were the high levels of personalization. Poor functionality was considered a major usability barrier.

The core alcohol harm reduction components of BRANCH were based on the Feedback, Responsibility, Advice, Menu of options, Empathy, Self-efficacy (FRAMES) model of alcohol brief interventions [21], which has been previously adapted for eSBI [23]. The FRAMES model is based on the principles of motivational interviewing, an established and evidence-based method to reduce alcohol-harm [24]. Core functions, which constituted the basic (control) version of BRANCH, included a drinking diary for recording of alcohol consumption and a goal setting function where users could set weekly goals based on alcohol units as well as setting a drink-free day. Users could monitor their drinking over time and receive feedback on it both descriptively and graphically. Information on drinking risks and benefits of cutting down was available to users.

The *comprehensive* (intervention) version of BRANCH included a number of targeted EPSs. Several theoretical models were used, including O'Brien and Toms' CMUE as described above and Oinas-Kukkonen and Harijuanas' PSD [17]. The comprehensive version included a Twitter or Facebook style newsfeed enabling interaction between app users, as well as providing tailored notifications, motivational messaging (including positive reinforcement and praise), and in-app reminders based on goals. The research team could also upload relevant material such as links to news articles, YouTube videos, and photographs. BRANCH was tailored; when signing up for the app, users selected their motivations for cutting down drinking. Personalized feedback and tailored information were delivered to users based on their selection of motivators. Additional goals were included such as reducing sugar, calories, and money spent on alcohol. Participants were encouraged to set goals based on their selected motivator. Extended infographic-style information targeted to motivations for cutting down were included. Users were allocated to a team based on these motivators. A *teams* page with a separate newsfeed channel was available for each motivator. Users could compare their progress against other users in their team as well as between teams and were awarded points for engaging with the app in line with gamification principles. Content was provided in a multimedia format, with a single exposure of content (all at once as opposed to staged).

BRANCH was a Web-based app, which meant it was hosted on the Web and on a server instead of being on a user's phone. For users, this meant that the app was not accessed and downloaded from the app store but was logged into the Web at each point of use.

Ethical Approval

Ethical approval was obtained from the University Research Ethics Committee (reference number RESCMR-16 or 17-2896).

Data Collection

Eligible participants were invited to participate in a semistructured interview lasting between 30 and 60 min and conducted by telephone. They were provided with an information sheet via email. On completion, participants were reimbursed with a £20 voucher for their time.

Consent was audio-recorded at point of interview. Semistructured interviews explored users' current engagement levels with health promotion apps, their views toward the effectiveness of such apps, and what they liked and disliked about the BRANCH app, specifically focusing on how they engaged with the app. The questions were open ended, encouraging participants to share as much information as possible about their experiences. The topic guide was broken into stages of engagement, drawn from O'Brien and Toms' CMUE [16], asking participants to describe their experience at each stage of engagement, for example, their motivations for first using the app, first use, and ongoing and disengagement experiences. The topic guide was tailored to arm allocation and user engagement level (Multimedia Appendix 1).

Reflexivity

The first author conducted all the interviews and had significant experience of moderating both 1:1 and focus group interviews. She had no previous relationship with any of the participants before the interview other than email contact to set up the interview. The only information participants were given about the interviewer was that they were a member of the app research team. In terms of reflexivity, the first author developed the app herself; therefore, special attention was paid to reduce social desirability bias in the interviews, whereby a positive response may have been elicited from the participants to be supportive of the author's work. To overcome this bias, participants were not told that it was the interviewer who developed the app and were encouraged to provide both positive and negative feedback.

Data Analysis

The interviews were recorded, and data were transcribed using a professional transcription service. Transcripts were checked for accuracy against the recordings, and names as well as other personally identifiable data were changed. The data were coded by the first author using QSR International's NVivo version 11.4.1 software. A Framework approach was used to analyze the data [22]. Framework analysis was developed to support the aims of applied research, such as mapping out the range and nature of a phenomenon, finding linkages and patterns within data, and creating typologies of behaviors or attitudes. Its objectives are to meet specific needs with actionable outcomes such as evaluating an intervention. Its key features are that it is grounded deeply in the accounts of the participant and that it is systematic, comprehensive, and allows for cross-case and between-case analysis. A framework matrix was created comprising descriptive themes and subthemes for each individual participant (or case). Specifically, the framework analysis comprised 5 key steps. First, the researcher read through the raw transcripts to gain familiarity with the data and create an initial map of themes derived from the research questions and identified themes from the data. Second, the researcher used this initial thematic framework to code (or index) the raw data using NVivo and indexed the raw data extracts into their relevant thematic category. The thematic framework was then reviewed and refined to see if all important themes and subthemes were covered and were practically distinguishable from one another. Next, the data were summarized into a framework matrix to reduce the data for later thematic abstraction. This was completed in the data processing tool, Microsoft Excel. Each theme was given its own matrix, with each subtheme allocated a column and each participant (or case) a row. For each theme, the researcher summarized the raw data coded at each of the subthemes, keeping as close to the data as possible using key terms, expressions, or phrases from the participants.

The final stage was the abstraction and interpretation of the summarized data. For each theme and subtheme for each case, the researcher read all the summaries and noted key elements, perceptions, and views in an additional column inserted into an Excel spreadsheet. From these underlying dimensions, higher-order categories were developed and abstracted from the data. This process moved the data from the descriptive to the more conceptual level. From these abstracted data, linkages were developed: linkages refer to connections or patterns between different sets of phenomena in the data or can also refer to whether there are links between sets of phenomena such as beliefs, experiences, or behaviors and different subgroups in the study—for example, linking together experiences of app usage and disengagement behaviors or understanding the link between app use and impact on drinking. For clarity, thematic frameworks around app engagement were conceptualized according to Short et al's [19] model of user engagement where engagement was categorized according to individual-level, environmental-level, or app-level engagement factors.

To explore distinct types of engagers, complex typologies were developed involving the interconnection between dimensions such as particular beliefs, experiences, or behaviors (or *positions*) in the data. Multiple-linkage typologies were developed [22], which refer to unique clusters or combinations of *positions* that create distinct typologies. Although the same *position*, such as a specific view, behavior, or belief, can occur across more than 1 typology, it is the unique combinations of positions that create the distinct typologies. Typologies were first developed at the individual case level across the framework matrices and were then abstracted to the phenomena level (as opposed to case level). To check the robustness of the typologies, the researcher went back to the case level to check for fit against each individual case. Finally, to provide recommendations for future app development, higher order findings across all the aims were summarized into key priorities and *meta* findings.

Results

Participant Characteristics

A total of 20 participants completed the interview, 11 from the intervention group and 9 from the control group. There were more females (16/20, 80%) than male participants (Table 1). This reflects the overall sample of the RCT where 70% participants were female. The mean age of the sample was 24 years (SD 3.0). Participants were spread over a wide geographic area of the United Kingdom in a range of professions and education. The majority of participants (12/20, 60%) were students, which again was consistent with the overall sample of the RCT. The mean number of log-ins to the app was 8 (SD 10.3), with a range of 1 to 35.

Table 1. Participant demographics, allocation in randomized controlled trial (RCT), and use characteristics.

Description of sample	Whole sample	High engagers	Low engagers
Intervention arm, n (%)	11 (55)	6 (30)	5 (25)
Control arm, n (%)	9 (45)	5 (25)	4 (20)
Age in years, mean (SD)	24.0 (3.0)	25 (3.1)	22 (2.2)
Female, n (%)	16 (80)	8 (40)	8 (40)
Occupation, n (%)			
Student	12 (60)	5 (25)	7 (35)
Employed	7 (35)	5 (25)	2 (10)
Unemployed	1 (5)	0 (0)	1 (5)
Log-ins, mean (SD)	8.0 (10.3)	15.0 (10.9)	1.0 (0)

Typologies of User Engager

A principal aim of the analyses was to identify distinct user typologies, of which 3 were identified: the Tracker, the Cut-downer, and the Noncommitter. These typologies are outlined in detail below.

The Tracker: Monitoring and Tracking Alcohol Consumption

This was the most common type of app engager. The defining feature of the participants in this group was that their motivations for using the app were not primarily to cut down but mostly to monitor and keep track of their alcohol use. Some were also interested in monitoring their spending and finding out whether or not they were at risk of the harmful effects of drinking. Participants in this group were very conscious about monitoring not only their alcohol usage but also their health and lifestyle in general. Trackers used multiple health apps to track a range of lifestyle factors such as calories, exercise, finances, menstrual cycle, and sleeping patterns. For Trackers, using BRANCH was just an extension of their current health app usage, which fitted easily into their daily habits of entering data into a variety of health apps:

> *I find it quite satisfying to have it logged down accurately. But I enjoy things like this, like I've always kept a diary since I was young, I've always used like apps to track my menstrual cycle and always kept up to date with them, and use them really accurately...* [Female, 24, high engager]

Trackers described themselves as organized individuals who wanted the structure that monitoring can provide to their day-to-day lives. Trackers had a strong positive emotional response to keeping track of their alcohol use, which made them feel in control and empowered.

How Trackers Engaged With BRANCH Features

Trackers consisted of both high and low engagers. The high engagers would use the app consistently and meticulously; whereas some used it daily or weekly, others stepped in and out of using it. Low-engaging Trackers typically just put in a few drinks but still identified their motivations to be to track their health as opposed to cutting down. High-engaging Trackers were strongly motivated to use the app but were focused particularly on entering drinks into the drinking diary. Trackers would typically just log in and log out to enter drinks into the drinking diary, ignoring the other features. Trackers had little to say about the other features and were mostly not interested in setting a goal as reduction was not their primary motivation:

> *I guess what I wanted the app for, for my own exploration of my drinking habits and sort of seeing how much I'm drinking, the drinking diary made the most sense to focus on...I don't really know what the other stuff does on there because I haven't, like the goals that I did look at, I haven't set any goals, I just looked at them...* [Female, 20 years, high engager]

They reported the information on alcohol presented in the app to be unmeaningful and not relevant to them. This is except for the Information section on units, which allowed them to enter drinks more accurately. In terms of the feedback, Trackers appreciated being able to get feedback on their spending on drinking but did not praise the drinking risk feedback as it was not regularly updated and did not change when they entered their drinking data. Trackers did not use the social features of BRANCH viewing tracking as a solitary activity, not one necessarily shared with others. None of the Trackers used the Team section.

Impact on Drinking for Trackers

The majority of Trackers did not have intentions to cut down their alcohol use. Some discussed how they thought they might drink too much, particularly in a social context, and viewed the app as an opportunity to find out whether they were drinking at a harmful level. In terms of impact on drinking, they mostly described being made more aware of their drinking patterns and habits, particularly about drinking more mindfully. Some learned that they drank more alcohol than they thought they did. Having different options to visualize the data in different graphs was considered a helpful way to understand drinking habits. A few did describe cutting down, being more motivated to turn down a drink when offered when out with friends, and making more sensible choices when choosing a drink since using the app. However, overall, the app's impact on drinking was described as increased awareness and not significant change in drinking level:

> *I wouldn't say it has had a massive impact. I would say it has made me a lot more aware of how much I drink when I drink. I don't drink as frequently as other people but I probably drink more than some people do when I do drink. I think that has made me a lot more aware of that, the amount that I drink at one time and so I have been making a conscious effort to kind of restrict that.* [Male, 21 years, high engager]

Facilitators to Use for Trackers

At the app level, engagement facilitators centered around the provision of different features to monitor drinking, costs and calories, and different ways of viewing these data. Accuracy of data was key to Trackers, as was ease of use to enter drinks into the drinking diary such as autofill cues to enter data. For Trackers, any function that made entering drinks faster and more functional was a facilitator to engagement with BRANCH. At the individual level, the provision of monitoring features for this group led to strong positive emotional responses to the app, which fostered motivation to use it. For example, it triggered feelings of control, security, health, empowerment, and autonomy.

Barriers to Use for Trackers

At the app level, usability issues such as Web-based app access problems and nonfamiliarity with Web-based apps in general were frequently highlighted and were the main reasons low-engaging Trackers cited for low engagement. Registration problems and length of time taken to enter data were also barriers to engagement for Trackers. However, Trackers were also likely to try to overcome any registration or usability issues when they occurred because of higher levels of motivation to use the app. Although not familiar with Web-based apps, they

would work out on how to either pin the icon to their home screen or save the Web page link. A lack of short message service (SMS) text messaging or push notification reminders meant that Trackers sometimes forgot to enter data. Trackers were wary of the privacy issues about the social component: cutting down drinking was considered somewhat a nonacceptable social activity, unlike, for example, fitness trends, or weight loss, which often share data via apps. At the environmental level, barriers to engagement included life constraints such as tiredness, time taken to add in data, being busy at work, or being away on a holiday.

The Cut-Downer: Intention to Reduce Alcohol Consumption

The Cut-downers' primary reason for using the BRANCH app was to reduce their alcohol intake. Participants from this group were worried about the health risks associated with alcohol use and would typically believe that they already drank too much:

> *Interviewer: And did you want to cut down your drinking at all, was that in your mind?*
>
> *Interviewee: Yes, definitely, because I know when I was like 17, 18, 19, 20, 21, I used to binge drink, like a lot, so it'd be like nearly every weekend, if not like twice a week, and like as I've got older, I've cut back on it but this has helped me see, like how much I have cut back and where it is I still need to carry on trying to cut back.* [Female, 25 years, high engager]

Alternatively, they might have wanted to cut down for other reasons such as those of a financial nature. A few of them had previous quit attempts. Unlike Trackers, Cut-downers were not interested in monitoring all aspects of their health but were predominantly focused on alcohol use. Cut-downers did not use many different health apps but may have tried 1 or 2 in the past, with no regular use.

How Cut-Downers Engaged With BRANCH Features

Similar to Trackers, Cut-downers were also very focused on the monitoring features where they could record their alcohol intake and set goals for themselves (see matrix, Multimedia Appendix 2, for a comparison of typology characteristics):

> *What I mainly used from the app was the drinking diary and the goals. So, the goals of, say, if I was having five or six pints a week, maybe cutting that down to four, something like that...I think the drinking diary helped as well. I filled that in as I was going along.* [Female, 26 years, high engager]

Cut-downers were all high engagers and typically used the app consistently either daily or at least two to three times per week. Cut-downers were also concerned about the importance of being able to accurately input drinks in the app, particularly about having enough options for brands of drinks and how to record drinks when not pouring or buying one's own. Cut-downers also had a broader interest in other features of BRANCH, which helped them achieve their goal of cutting down their drinking. Cut-downers were more likely to set a goal than other typologies, with the majority setting at least one goal or a drink-free day. One Cut-downer set a goal for himself without the support of the app, as he did not feel he needed the app to help him with it, suggesting that using the app might encourage behavior change beyond the support offered in the app's features.

Cut-downers also engaged with the feedback section; they found the feedback on drinking and risks not only eye opening and often shocking but also frustrating as this did not change week to week. One Cut-downer commented on the Newsfeed, but no other participants used the social features stating that it was either not relevant to their goals or, similar to the Trackers, that cutting down on alcohol use is a solitary activity. Concerns were shared over privacy issues of the social component. No Cut-downers used the Team section.

Impact on Drinking for Cut-Downers

One participant reported having stopped drinking for a week, and another reported having cut back during the time they were using the app. Cut-downers did not express regret when they did not achieve their goals; instead, similar to the Trackers, they reported an increase in awareness of the amount of alcohol they were consuming as opposed to reduced consumption. This was expressed as a positive outcome and an achievement in itself. Cut-downers also reported being surprised by how much they were actually drinking, about how many units there are in alcohol, and the associated risks:

> *I think its helped in the fact that I now know what I'm drinking and how much I'm drinking. However, it's not really helped me cut down as such, because I'm surrounded by it all the time. I'm still trying to cut down.* [Female, 22 years, low engager]

Facilitators to Use for Cut-Downers

Ease of use was cited as a facilitator to using the app. Some Cut-downers favored the design of BRANCH as a Web-based app as it did not take up data space on their phone. At the individual level, Cut-downers experienced strong emotional responses to the app when they achieved a goal of cutting down in comparison with Trackers who felt this way when they successfully entered data for a period. Motivation level also played a large part in their usage of the app as they had a specific goal they wanted to achieve, which encouraged them to use the app more frequently.

Barriers to Use for Cut-Downers

At the app level, Cut-downers were not expecting a Web-based app and some experienced Web-based app accessibility issues. However, similar to the Trackers, having high levels of motivation to use the app meant that they were willing to put in effort to learn how to use Web-based apps. Registration issues, such as passwords not working, were cited as barriers. Cut-downers who did not reduce their alcohol-use suggested that lack of reminders contributed to this. Similarly, some struggled to make plans ahead to cut down drinking and highlighted that it was not practical to set goals when they had social occasions planned. This demonstrates the conflict between the functioning of the app in a potentially risky environment of high exposure to alcohol consumption at social events. One participant commented that the social newsfeed was not used frequently enough for it to be engaging. They also commented that the Information and Feedback sections were not regularly

updated with new information to make it interesting enough to return. Regarding the social feature, similar to Trackers, Cut-downers were concerned with the privacy element of this component. One participant reported that he disengaged with the app because it helped him achieve his aim of reducing his alcohol consumption, which is a term known as *effective (dis)engagement* [12].

The Noncommitter: Lack of Motivation to Use Health Apps

The Noncommitters did not engage with the BRANCH app at all and were all low engagers. They cited a variety of reasons for being attracted to the app in the first place, such as spending, curiosity, health, or understanding their drinking habits, but their initial interest and motivation quickly faded. Although they had good intentions about using the app at first, such as becoming healthier, Noncommitters reported lacking the motivation to even log in to the app, let alone use the app features. They did not regularly use other health apps and would say that this was because they were too *disorganized* or *lazy* to use them:

> *I didn't use it for very long because I thought it was too tedious and I wasn't willing to input every kind of drink I had…it was just kind of like a fad thing that I would start doing for a bit and then forget about.* [Male, 23 years, low engager]

How Noncommitters Engaged With BRANCH Features

Noncommitters typically had a quick look about the newsfeed and the drinking diary and perhaps entered a drink or 2 into the diary. Like the other 2 typologies, Noncommitters commented on concerns over the social element. These were primarily about concerns of privacy and about how discussing harmful drinking is not socially acceptable and is still a taboo. Noncommitters did not want other people to find out that they were using the app, with 1 commenting on how using an app such as BRANCH is different from participating in an event such as Dry January, which is a common and an acceptable activity to partake in because it is not assuming that one has a *problem* with alcohol. Noncommitters did not use the other features, such as goals, information, or feedback, lacking the motivation to use them or not even knowing they existed. A few scanned over the features but did not engage with them meaningfully:

> *I think I added in one drink and I remembered something else I drank, so I went back in and added it in, I haven't been back since.* [Male, 25 years, low engager]

Impact on Drinking for Noncommitters

Noncommitters did not report any significant impact on their drinking. One participant stated that it may have had a temporary effect, which was quickly lost as he disengaged. In contrast to the other 2 typologies, they did not report the app increasing any awareness in their drinking:

> *I don't think it has really affected it much but I just think that's because I haven't used it a lot. But I think if I did, it would've, because I'd have seen like the amounts that I have been drinking.* [Female, 20 years, low engager]

Facilitators to Use for Noncommitters

As the group that engaged least with the BRANCH app, there were few facilitators to use for the Noncommitters. One participant mentioned that having a Web-based app meant that it did not take up storage on the phone. A few also commented that they appreciated the functionality of the drinking diary and that the graphs were helpful. However, this was not sufficient to encourage them to return to the app.

Barriers to Use for Noncommitters

At the app level, participants were expecting a native app and were confused about how to use a Web-based app. There were also barriers in terms of needing internet connection, there being no reminders, and too many steps to input data. This meant they were put-off using the app again. As the individual-level motivation was a major barrier, although initially citing good intentions to want to use the app, these seemed to quickly wane. Noncommitters typically expressed not being *bothered* to go through the various required steps, feeling that it was *too much effort* to input data and having to access the app on the Web. This contrasts to the Trackers and Cut-downers who while experiencing similar issues had the motivation to overcome these barriers and learn how to use the app effectively. This suggests that the level of motivation the Noncommitters had to track and cut down their drinking was significantly lower than the other 2 typologies.

Discussion

Principal Findings

Understanding how users interact with and use Web-based and digital health technologies is an important step to improving engagement rates and user experience. The examination of usage patterns can identify typologies, exploring whether there are important differences in how users interact with the app. The major finding of this study was the identification of 3 discreet typologies of engagers to explore the individual experiences of using the BRANCH app: the Tracker, the Cut-downer, and the Noncommitter. To the authors' knowledge, this is the first study to qualitatively identify a typology of engagers with a smartphone-based alcohol intervention. These data can be used to improve content and functionality of the intervention and to tailor the intervention to the user or certain groups of users, thereby increasing engagement and potentially the effectiveness of the intervention. On the basis of these findings, the key feature to improve BRANCH would be to determine the user typology on app registration, such as via a short questionnaire, and target the app content to their typology.

The Tracker had the largest group of users and was defined by their motivations to use health-tracking apps to monitor and understand quantified self-data. Users did not have intentions necessarily to cut down but purely to track and predominantly used only the drinking diary in the app to enter drinking data. This gave them a sense of control over their lives. They were characterized by frequent usage of a range of health monitoring apps such as fitness, calories, sleep, and spending. Cut-downers

were motivated to use the app because they wanted to reduce their alcohol consumption. This group's users, unlike Trackers, did not use many different health apps but saw the BRANCH app as an opportunity to try to cut down their alcohol use. Similar to Trackers, they not only used the monitoring features of the app but also the goal setting and feedback functions. Noncommitters were characterized as a group of users who were initially enthusiastic about using the app; however, this enthusiasm quickly waned and they only used the app once or twice, gaining no benefit from it. Noncommitters were particularly defined as lacking in motivation to make any behavior changes and were easily put-off in terms of using the app by any usability issues or perceptions that the app required too much time and effort to use.

Although the 3 typologies did occupy discreet categories of engagers, there were similarities between the groups in terms of how they perceived barriers and facilitators to engagement. This is important for future alcohol brief intervention app development as across typologies, barriers and facilitators imply *core* or *fundamental* usability and component issues that need to be considered when designing an app. For example, almost none of the participants were familiar with Web-based apps and cited this as a barrier to use as it required additional effort and time to log in. This is consistent with previous usability testing research [6], which suggests that users want features *at their fingertips*, with minimal required input and effort. Indeed, app-level barriers were the most frequently cited barrier to engagement across the typologies. Issues such as lack of reminders and content updates were highlighted by all typologies. Although improvements were made to the prototype version of the app during its development, clearly, usability is a persistent and enduring issue, which is a priority for optimal eSBI app development.

Comparisons With Previous Work

No previous qualitative study has explored typologies of user engagement in the alcohol field. Smith et al [15] explored a similar concept of *trajectories of use* with a smoking cessation app and identified 4 different engagement trajectories: (1) productive engagement, (2) counterproductive engagement, (3) productive disengagement, and (4) counterproductive disengagement. Considering this study in light of this conceptual model, Trackers and Cut-downers displayed characteristics of productive engagement and disengagement, being invested in the program and (some of) its components, using adaptive strategies to overcome any barriers of usage and describing disengagement with the app because they achieved their goals of use. Similarly, Noncommitters displayed characteristics of counterproductive engagement and disengagement, having a negative response to the program, not relating to intervention content and disengaging because of usability or motivation factors. Although smoking and alcohol intervention apps often share similar behavior change techniques such as setting goals and monitoring [25], there are differences in the objectives between them. For example, in the smoking domain, cessation is the goal of apps, whereas in the alcohol field, apps typically target hazardous and harmful drinkers, as opposed to dependent drinkers, where reduction not abstinence is the goal. As such, the goals and level of motivation of the app user may differ, which may potentially differentially affect user engagement.

Smith et al [15] focused on identifying trajectories of usage through the program life cycle. This study extends their work by building individual-level factors into the engagement framework, such as personal user characteristics to define engagement not only through usage but also through personality, other app usage, and detailed exploration of motivations and intentions for usage. Subsequently, this study contributes to creating a comprehensive picture of how people engage with eSBI at the individual, environmental, and technological level.

Another similarity across the Tracker and Cut-downer typologies was that most of these users did not report a reduction in alcohol use but instead an increased awareness of level of alcohol consumption and promotion of more *mindful* drinking. Indeed, even for the Cut-downers, when prompted about how they felt when did not achieve their alcohol reduction goal, they constructed this failure with a positive narrative in which they gained awareness, which was achievement enough. This is an important finding because it raises the question of what is the primary goal of eSBI apps and what are the subsequent clinical implications these findings may have. In terms of risk zones for harmful drinking, the app may slow down or stop transitions between zones by identifying at-risk individuals, or it may sit on the pathway between as an identification tool to support patients into treatment. Such steps are critical, as it is important to treat alcohol use in the early stages before it has developed into dependence.

Although a recent Cochrane systematic review examining personalized digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations reported that participants using a digital intervention drank approximately about 3 UK units less than those who received no or minimal interventions; this included all types of digital interventions and was not specific to apps. The evidence for eSBI apps to reduce alcohol consumption is currently inconclusive [2,26-29], and it has been argued that there are differences in the way that eSBI apps may be used by target users compared with eSBIs, which are computer based [6,8]. For example, app-based interventions can be used quickly while on the go, collecting data in the moment; however, computer-based interventions require users to sit down and dedicate time and effort to the program. This may influence the level of the effectiveness of the program or may highlight differences between users. Those who are dedicated to using a computer-based eSBI may have higher levels of motivation to reduce their alcohol use, which would explain why the evidence for eSBI apps is inconclusive. Looking at *effectiveness* from a different perspective, perhaps, the aim of eSBI apps may not be a reduction in units but increased awareness leading to other interventions. For example, Cut-downers may attend their general practitioner (GP) after using the app, and the subsequent GP intervention may reduce their level of drinking.

The lack of use of the social and gamification component was a surprising finding considering these features had been designed in collaboration with the target user group and because previous research has supported the use of such features in the broader

health behavior change literature [30,31]. The main reasons cited for nonusage were that (1) they were concerned about privacy issues and (2) cutting down on alcohol use is considered an individual instead of a group activity. Unlike other types of health behavior change such as exercise or weight loss, there exists a persistent social stigma about cutting down on alcohol use; it is, therefore, consistent that participants were wary of using the social component in BRANCH app. The effectiveness of social components in digital alcohol harm reduction interventions is unclear. Although preliminary developmental work for the BRANCH supported the use of social features [8], research on existing smoking and alcohol interventions has suggested that users do not want to share their progress on social networks [25]. A recent systematic review [7] of RCTs that examined the effectiveness of EPSs in Web-based substance misuse interventions reported ambiguous outcomes for the use of social features. It may be that social features need to be adapted further for substance-use interventions focusing on establishing trust about privacy and targeting issues of stigma about cutting down on alcohol use.

In terms of implications for future design of eSBI apps, based on the current research, it is recommended that push notifications or SMS reminders are used. Overall, users were unfamiliar with Web-based apps, and native apps may be a more appropriate user platform. Frequently updated information and improved usability should also be considered in future eSBI app design.

Tailoring is also an important feature for improvement of eSBI apps. A previous study [7] highlighted tailoring to self-efficacy as potentially effective, as well as some support for tailoring to motivations for quitting, abstinence status, and a personalized source. This qualitative study suggests that tailoring to engagement typology might also be associated with improved outcomes. Future iterations may benefit from creating an app where certain features could be turned on and off, tailoring features to different typologies. For example, Trackers might benefit most from an eSBI app that prioritized optimization of the monitoring features such as having the drinking diary as the home page and more sophisticated data visualizations. Cut-downers may need an app that is targeted to supporting users reach their goals, for example, more emphasis on positive reinforcement of goal completion and tangible milestones for cutting down integrated into the app. Nonengagers might need an app with features to enhance motivation to engage with behavior change as a precursor to further app usage.

This study questions how we measure the *effectiveness* of eSBI apps. The objective of eSBI apps may not be to reduce alcohol consumption but to serve as a tool to support people to seek other types of interventions and treatment. Perhaps eSBI is not a stand-alone treatment for reducing alcohol use (as opposed to computer-based eSBI) but is in fact a low-intensity tool to facilitate harmful drinkers to become aware of their level of drinking and be a step in the treatment pathway toward behavior change. From a clinical perspective, such apps could be prescribed by clinicians to patients to make them aware of the level of their drinking, and then more intensive treatments such as computer-based interventions could be recommended. However, as the first study to report such findings, this needs further investigation before absolute recommendations can be made.

Another question that this study raises is whether eSBI apps are a suitable intervention for everybody or whether they are an appropriate intervention for certain typologies. Trackers and Cut-downers both reported benefits from the program although from an awareness rather than a reduction perspective. However, Noncommitters did not report any tangible benefit. On the one hand, it may be that certain individuals enjoy using health monitoring apps and are more likely to see a positive effect. Others, such as the Noncommitters, do not enjoy entering data into apps and find it difficult to engage with such programs and, therefore, may benefit from more traditional face-to-face interventions or apps that aim to enhance motivation to engage in further interventions or behavior change. From a clinical perspective, it may be that there is not a *one-size-fits-all* solution to alcohol brief interventions (BIs), and patients should be offered a range of tools, including digital and face-to-face interventions, to select what works best for them and their lifestyle.

On the other hand, it is also plausible that there is scope within eSBI to target motivation levels within individuals who do not initially engage, such as the Noncommitters. The findings of the study suggest that because different typologies have different levels of motivation to cut down, eSBI may be compatible with the application of the Transtheoretical Model of Change [32] to increase motivation level to cut down. For example, the app could be tailored to stage of change, such as through a questionnaire at registration stage, followed by specific intervention components tailored to the user's motivation level. However, further research is needed to explore whether or not this would be an effective feature. It is suggested that future research examines further engagement from the perspective of the Transtheoretical Model of Change. With future enhancements in technology and the potential linkage of health to digital health watches such as Fitbits, it is plausible that such devices could measure blood alcohol concentration (BAC) transdermally, eradicating the need for self-report input and many of challenges faced by current eSBI. In the future, apps in conjunction with automated BAC calculations may be able to automatically measure usage and provide tailored support and advice without the need for human input.

Limitations

The majority of participants were female (16/20, 80%). Although potentially introducing bias, this is consistent with previous research published on the BRANCH app where 90% of the qualitative sample was female [6,18]. This is also consistent with the characteristics of participants of the main RCT, with a 70% female sample. Research suggests that women are more likely to use Web-based resources to access health information [33,34]. As such, this sample may reflect the type of individual most likely to engage with eSBI interventions. Future studies may wish to explore in more detail why females are more likely to engage and examine how eHealth interventions can be tailored to gender.

Recruitment may have discouraged potential participants who had negative experiences or limited engagement with the app.

Some potential participants contacted declined to participate as they felt they did not have enough feedback to offer because of low engagement. However, efforts were made to outline that participants who did not use the app frequently were still invited to participate; subsequently, an equal split of high and low engagers were recruited. Qualitative interviews can be subject to response bias where participants provide views that they believe the researcher wants to hear. A couple of *nonengagers* (logged in once), identified from app-usage data extracted from the server, did report using the app more than once. This may imply a response bias, or poor memory recall, and may have resulted in some participants over-reporting app usage, potentially biasing the analysis. However, a strength of the study overall was that it recruited both high and low engagers, so a range of views was provided. Characteristics of trial participants who declined to participate in the qualitative interviews were not collected. Therefore, there may exist a bias in the characteristics of participants recruited, such as the high-engaging participants being more motivated to provide positive experiences of using the app than low engagers who declined to participate. All participants were selected as harmful drinkers (Alcohol Use Disorder Identification Test (AUDIT) score 16+). As alcohol is still considered a stigmatizing subject, participants may have not shared all their experiences of drinking at a harmful level. Efforts were made by the researcher to provide a nonjudgmental space in which to discuss their experiences of drinking. The data were analyzed by only a single researcher, which may have biased the results. As this is the first study in the alcohol field to qualitatively define typologies of user engagement, the generalizability of the findings requires further research. However, as the findings are consistent with quantitative work [35,36] and qualitative work from the smoking cessation field [15], this suggests the findings are generalizable to the broader population.

Conclusions

This study has identified 3 typologies of eSBI app users. Trackers use only monitoring features and are not interested in cutting down, only measurement; Cut-downers use the app to reduce their alcohol use and will use more features such as goal setting functions; and Noncommitters have good intentions but quickly disengage from using the app. Although in need of replication, it provides a first step in understanding how eSBI apps can be tailored to different user types to improve engagement and ultimately effectiveness. It also questions what the purpose or utility of eSBI apps is. As opposed to reducing consumption, eSBI apps may serve as a tool to provide a stepping stone in the pathway of treatment to prevent individuals developing more serious alcohol-related conditions. With the consistent findings from eSBI app trials that apps may not be as effective as computer-based methods as previously thought, perhaps it is time to rethink how we conceptualize the purpose and function of eSBI apps in the future.

Acknowledgments

The research was funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South London at King's College Hospital National Health Service (NHS) Foundation Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

CD is partly funded by the NIHR Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust and King's College London and partly funded by the NIHR Collaborations for Leadership in Applied Health Research and Care South London at King's College Hospital NHS Foundation Trust.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Topic guide.

[PDF File (Adobe PDF File), 23KB - mhealth_v6i12e11692_app1.pdf]

Multimedia Appendix 2

Matrix of engagement typologies.

[PDF File (Adobe PDF File), 26KB - mhealth_v6i12e11692_app2.pdf]

References

1. Attwood S, Parke H, Larsen J, Morton KL. BMC Public Health. 2017. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application URL: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-017-4358-9 [accessed 2018-10-29] [WebCite Cache ID 73Wgyrugw]
2. Deluca P, Coulton S, Boniface S, Donoghue K, Gilvarry E, Kaner E, et al. Effectiveness and cost effectiveness of a smartphone based electronic alcohol intervention for adolescents: findings from the SIPS jr trials. In: Proceedings of the

14th annual conference of INEBRIA. Addiction Science & Clinical Practice. 2017 Presented at: INEBRIA 2017; Sept 14-15 2017; New York, NY, USA p. A55.
3. Witkiewitz K, Desai SA, Bowen S, Leigh BC, Kirouac M, Larimer ME. Development and evaluation of a mobile intervention for heavy drinking and smoking among college students. Psychol Addict Behav 2014 Sep;28(3):639-650 [FREE Full text] [doi: 10.1037/a0034747] [Medline: 25000269]
4. Hutton HE, Wilson LM, Apelberg BJ, Tang EA, Odelola O, Bass EB, et al. A systematic review of randomized controlled trials: Web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011 Apr;13(4):227-238. [doi: 10.1093/ntr/ntq252] [Medline: 21350042]
5. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010 Nov 18;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]
6. Milward J, Deluca P, Drummond C, Watson R, Dunne J, Kimergård A. Usability testing of the BRANCH smartphone app designed to reduce harmful drinking in young adults. JMIR Mhealth Uhealth 2017 Aug 8;5(8):e109 [FREE Full text] [doi: 10.2196/mhealth.7836] [Medline: 28790022]
7. Milward J, Drummond C, Fincham-Campbell S, Deluca P. What makes online substance-use interventions engaging? A systematic review and narrative synthesis. Digit Health 2018;4:2055207617743354 [FREE Full text] [doi: 10.1177/2055207617743354] [Medline: 29942622]
8. Milward J, Khadjesari Z, Fincham-Campbell S, Deluca P, Watson R, Drummond C. User preferences for content, features, and style for an app to reduce harmful drinking in young adults: analysis of user feedback in app stores and focus group interviews. JMIR Mhealth Uhealth 2016 May 24;4(2):e47 [FREE Full text] [doi: 10.2196/mhealth.5242] [Medline: 27220371]
9. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
10. Ryan C, Bergin M, Wells JS. Theoretical perspectives of adherence to web-based interventions: a scoping review. Int J Behav Med 2018 Dec;25(1):17-29. [doi: 10.1007/s12529-017-9678-8] [Medline: 28730402]
11. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
12. 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]
13. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
14. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2 Suppl):TC9-T20. [doi: 10.1097/NCN.0b013e3182155274] [Medline: 21372645]
15. Smith W, Ploderer B, Wadley G, Webber S, Borland R. Trajectories of engagement and disengagement with a story-based smoking cessation app. In: CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. New York: ACM; 2017 Presented at: 2017 CHI Conference on Human Factors in Computing Systems; May 6th -11 2017; Denver USA p. 3045-3056. [doi: 10.1145/3025453.3026054]
16. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
17. Oinas-Kukkonen H, Harjumaa M. Communications of the Association for Information Systems. 2009. Persuasive systems design: Key issues, process model, and system features.Communications of the Association for Information Systems Internet URL: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3424&context=cais [accessed 2018-02-01]
18. Crane D, Garnett C, Brown J, West R, Michie S. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: think aloud and interview studies. Front Public Health 2017 [FREE Full text] [doi: 10.3389/fpubh.2017.00039]
19. Short CE, Rebar AL, Plotnikoff RC, Vandelanotte C. Designing engaging online behaviour change interventions: a proposed model of user engagement. Eur Health Psychol 2015;17(1):32-38 [FREE Full text]
20. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
21. Miller WR, Sanchez VC. Motivating young adults for treatment and lifestyle change. In: Howard GS, Nathan PE, editors. Alcohol use and misuse by young adults. Notre Dame, IN, USA: University of Notre Dame Press; 1994:55-81.
22. Ritchie J, Lewis J, Nicholls CM, Ormston R. Qualitative research practice: A Guide for Social Science Students and Researchers. London, UK: Sage Publications Ltd; 2014.
23. Deluca P, Coulton S, Alam MF, Cohen D, Donoghue K, Gilvarry E, et al. Linked randomised controlled trials of face-to-face and electronic brief intervention methods to prevent alcohol related harm in young people aged 14-17 years presenting to

Emergency Departments (SIPS junior). BMC public health 2015 Apr 15;15 [FREE Full text] [doi: 10.1186/s12889-015-1679-4]
24. Miller RW, Rollnick S. Motivational interviewing: Helping people change. New York: Guilford press; 2012:1-482.
25. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Dec 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
26. Davies EL, Lonsdale AJ, Hennelly SE, Winstock AR, Foxcroft DR. Personalized digital interventions showed no impact on risky drinking in young adults: a pilot randomized controlled trial. Alcohol Alcohol 2017 Nov 1;52(6):671-676. [doi: 10.1093/alcalc/agx051] [Medline: 29016711]
27. Gajecki M, Berman AH, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. Addict Sci Clin Pract 2014 Jul 2;9:11 [FREE Full text] [doi: 10.1186/1940-0640-9-11] [Medline: 24985342]
28. Berman AH, Rosendahl I, Andersson C, Gajecki M, Sinadinovic K, Blankers M. Smartphone apps targeting risky excessive drinking patterns among university students show differing subgroup effects over 20 weeks. In: Proceedings of the 14th annual conference of INEBRIA. Addiction Science & Clinical Practice. 2017 Presented at: INEBRIA 2017; Sept 14-15 2017; New York, NY, USA p. A54.
29. Garnett C, Crane D, Brown R, West R, Michie S. Estimating effectiveness of components of a smartphone app- Drink Less to reduce excessive alcohol consumption: a factorial randomised control trial. In: Proceedings of the 14th annual conference of INEBRIA. Addiction Science & Clinical Practice. 2017 Presented at: INEBRIA 2017; Sept 14th 2017; New York, NY, USA p. A53. [doi: 10.3389/conf.FPUBH.2017.03.00004]
30. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
31. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. PLoS One 2017;12(3):e0173403 [FREE Full text] [doi: 10.1371/journal.pone.0173403] [Medline: 28362821]
32. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
33. Fox S. Pew Internet. Washington DC: Pew Internet & American Life Project; 2011 May 12. The Social Life of Health Information URL: http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/ [accessed 2017-05-15] [WebCite Cache ID 6qTe92sFy]
34. Bidmon S, Terlutter R. Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. J Med Internet Res 2015 Jun 22;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]
35. Goh G, Tan NC, Malhotra R, Padmanabhan U, Barbier S, Allen JJ, et al. Short-term trajectories of use of a caloric-monitoring mobile phone app among patients with type 2 diabetes mellitus in a primary care setting. J Med Internet Res 2015 Feb 3;17(2):e33 [FREE Full text] [doi: 10.2196/jmir.3938] [Medline: 25648130]
36. Morrison C, Doherty G. Analyzing engagement in a web-based intervention platform through visualizing log-data. J Med Internet Res 2014 Nov 13;16(11):e252 [FREE Full text] [doi: 10.2196/jmir.3575] [Medline: 25406097]

Abbreviations

BAC: blood alcohol concentration
CMUE: conceptual model of user engagement
eHealth: electronic health
EPS: engagement-promoting strategy
eSBI: electronic screening and brief intervention
FRAMES: Feedback, Responsibility, Advice, Menu of options, Empathy, Self-efficacy
GP: general practitioner
NIHR: National Institute for Health Research
NHS: National Health Service
PSD: Persuasive Systems Design
RCT: randomized controlled trial
SMS: short message service

Edited by G Eysenbach; submitted 27.07.18; peer-reviewed by W Smith, C Ryan, O Perski; comments to author 06.09.18; revised version received 19.09.18; accepted 04.10.18; published 13.12.18.

Please cite as:
Milward J, Deluca P, Drummond C, Kimergård A
Developing Typologies of User Engagement With the BRANCH Alcohol-Harm Reduction Smartphone App: Qualitative Study
JMIR Mhealth Uhealth 2018;6(12):e11692
URL: https://mhealth.jmir.org/2018/12/e11692/
doi:10.2196/11692
PMID:30545806

Original Paper

Unveiling the Black Box of Diagnostic and Clinical Decision Support Systems for Antenatal Care: Realist Evaluation

Ibukun-Oluwa Omolade Abejirinde[1,2,3], MSc, MD; Marjolein Zweekhorst[1], PhD; Azucena Bardají[3], MD, PhD; Rudolf Abugnaba-Abanga[4], MPhil; Norbert Apentibadek[5], BA; Vincent De Brouwere[2], MD, PhD; Jos van Roosmalen[1,6], MD, PhD; Bruno Marchal[2], MD, PhD

[1]Athena Institute, Faculty of Science, Vrije Universiteit, Amsterdam, Netherlands

[2]Department of Public Health, Institute of Tropical Medicine, Antwerp, Belgium

[3]ISGlobal, Barcelona Institute for Global Health, Hospital Clínic - Universitat de Barcelona, Barcelona, Spain

[4]Presbyterian Health Services-North, Bolgatanga, Upper East Region, Ghana

[5]Association of Church Development Projects (ACDEP), Northern Region, Ghana

[6]Department of Obstetrics, Leiden University Medical Centre, Leiden, Netherlands

Corresponding Author:
Ibukun-Oluwa Omolade Abejirinde, MSc, MD
Athena Institute
Faculty of Science
Vrije Universiteit
WN Building, S-544
De Boelelaan 1105
Amsterdam, 1081HV
Netherlands
Phone: 31 205983143
Fax: 31 205987031
Email: i.o.abejirinde@vu.nl

Abstract

Background: Digital innovations have shown promise for improving maternal health service delivery. However, low- and middle-income countries are still at the adoption-utilization stage. Evidence on mobile health has been described as a black box, with gaps in theoretical explanations that account for the ecosystem of health care and their effect on adoption mechanisms. Bliss4Midwives, a modular integrated diagnostic kit to support antenatal care service delivery, was piloted for 1 year in Northern Ghana. Although both users and beneficiaries valued Bliss4Midwives, results from the pilot showed wide variations in usage behavior and duration of use across project sites.

Objective: To strengthen the design and implementation of an improved prototype, the study objectives were two-fold: to identify causal factors underlying the variation in Bliss4Midwives usage behavior and understand how to overcome or leverage these in subsequent implementation cycles.

Methods: Using a multiple case study design, a realist evaluation of Bliss4Midwives was conducted. A total of 3 candidate program theories were developed and empirically tested in 6 health facilities grouped into low and moderate usage clusters. Quantitative and qualitative data were collected and analyzed using realist thinking to build configurations that link intervention, context, actors, and mechanisms to program outcomes, by employing inductive and deductive reasoning. Nonparametric *t* test was used to compare the perceived usefulness and perceived ease of use of Bliss4Midwives between usage clusters.

Results: We found no statistically significant differences between the 2 usage clusters. Low to moderate adoption of Bliss4Midwives was better explained by fear, enthusiasm, and high expectations for service delivery, especially in the absence of alternatives. Recognition from pregnant women, peers, supervisors, and the program itself was a crucial mechanism for device utilization. Other supportive mechanisms included ownership, empowerment, motivation, and adaptive responses to the device, such as realignment and negotiation. *Champion* users displayed high adoption-utilization behavior in contexts of participative or authoritative supervision, yet used the device inconsistently. Intervention-related (technical challenges, device rotation, lack of performance feedback, and refresher training), context-related (staff turnover, competing priorities, and workload), and individual factors (low technological self-efficacy, baseline knowledge, and internal motivation) suppressed utilization mechanisms.

Conclusions: This study shed light on optimal conditions necessary for Bliss4Midwives to thrive in a complex social and organizational setting. Beyond usability and viability studies, advocates of innovative technologies for maternal care need to consider how implementation strategies and contextual factors, such as existing collaborations and supervision styles, trigger mechanisms that influence program outcomes. In addition to informing scale-up of the Bliss4Midwives prototype, our results highlight the need for interventions that are guided by research methods that account for complexity.

(JMIR Mhealth Uhealth 2018;6(12):e11468) doi:10.2196/11468

KEYWORDS

systems analysis; Ghana; clinical decision support; antenatal care; mHealth; program evaluation

Introduction

Background

Digital health innovations have gained support as a means to improve health service delivery while strengthening health systems [1,2]. Mobile technologies (mobile health, mHealth) for maternal health in low-resource settings can play a role in addressing information, skills, and resource needs at various points in the continuum from prenatal to postnatal care [1,3,4].

The majority of digital health innovations for maternal health involve use of short messaging services, voice calls, point-of-care diagnostics, and health information management systems [1,4]. Other less explored areas recently gaining attention include its use for clinical decision support and remote monitoring. This is particularly important in the context of poor road networks, remote geographical locations, weak referral chains, and alarming workforce shortages. Diagnostic and decision-support systems are a group of digital health innovations that aim to address challenges of timely and effective health care, using evidence-based principles [2]. Despite evidence of their importance for task shifting and promoting adherence to clinical practice guidelines, attempts to embed them into large-scale service structures are yet to be attained [4].

Evidence on mHealth has been critiqued for being a black box with little knowledge from pilot projects to inform prototype development and scale-up [5]. The dominant discourse is that low technological skills alongside infrastructural barriers are at the root of poor mHealth uptake in low- and middle-income countries (LMICs). An alternative and less explored explanation is that factors unique to the ecosystem of health service delivery need to be accounted for, motivating calls for knowledge on mHealth that is grounded in theoretical understanding [6-8]. A recent theory-based analysis on what works or not for mHealth in maternal health service delivery has shown that LMICs are still at the adoption-utilization stage [9]. In their review, Chib et al also highlight a knowledge gap on mechanisms for mHealth adoption and the role of theoretical explanations in addressing these gaps [8].

This study aims to identify causal factors underlying the variation in mHealth usage during the adoption and utilization phases of an intervention and understand how to overcome or leverage these in subsequent implementation cycles. Findings will contribute to the body of evidence on contextual and domain-specific applications of similar innovations in other low-resource settings.

Description of the Intervention

In 2016, a consortium of 7 organizations representing a south-north public-private partnership embarked on a project to prove the viability of a modular integrated diagnostic kit tagged the Bliss4Midwives (B4M) device (unpublished data [10]). The B4M device supports instant informed diagnosis during antenatal care (ANC) by enabling noninvasive point-of-care screening for preeclampsia, gestational diabetes, and anemia—3 main screening components of ANC. The components of the device include a noninvasive hemoglobin reader with infrared sensors mounted on a finger clip, a self-inflating blood pressure cuff, and an automated urinary dipstick reader for measuring urinary protein and glucose. In the absence of B4M, target beneficiaries in remote areas would otherwise have to travel to other health facilities to conduct these tests, delaying timely detection and management of high-risk complications [11]. B4M was introduced in 7 health facilities in the upper east region and northern region of Ghana. Additional details on the device, project setting, viability, and beneficiary experiences have been reported elsewhere (unpublished data [10];[11]).

Although both users and beneficiaries valued B4M, results from the pilot showed wide variations in usage behavior and duration of use across project sites (unpublished data [10]). Beyond establishing viability of the intervention, application of a theory-based approach requires assessing why and how exactly it works [12]. In line with the long-term goals of the consortium, evaluation findings will inform the design and implementation of an adapted B4M prototype.

Methods

Study Setting

A total of 6 prototype devices were deployed in 7 predominantly rural locations—4 facilities in the upper east region and 3 in the northern region. A total of 25 maternal health workers were trained to operate B4M. As the device was withdrawn from 1 facility in the second month of the intervention, the evaluation focused on 6 of the 7 health facilities: facilities A to D in the upper east region and facilities E and F in the northern region. Facility A is the ANC unit of a district hospital and the first-level referral point for facilities B, C, and D, which are health centers. Facility E is an independent public health unit of a district hospital, whereas F is a health center. With the exception of facilities B and C, which shared a single B4M device on a rotating schedule, the other facilities had stable access to 1 device each.

Table 1. Adoption and utilization per health facility.

Adoption level/Utilization level	Low adoption	Moderate adoption	High adoption
Low utilization	Facilities B and E	Facility C	Facility A
Moderate utilization	N/A[a]	Facility F[b]	Facility D
High utilization	N/A	N/A	N/A

[a]N/A: not applicable.

[b]Due to data loss and inability to track the usage trend in facility F, we relied on cumulative usage data and reports from monitoring visits.

Table 2. Clustering of cases.

Usage combinations	Clusters[a]		
	Low usage	Moderate usage	High usage
Low adoption—low utilization	Facilities B and E	N/A[b]	N/A
Moderate adoption—low utilization	Facility C	N/A	N/A
Moderate adoption—moderate utilization	N/A	Facility F	N/A
High adoption—low utilization	N/A	Facility A	N/A
High adoption—moderate utilization	N/A	Facility D	N/A
High adoption—high utilization	N/A	N/A	None

[a]As utilization covered a longer period than adoption and total duration of use varied between facilities, when defining clusters, cases were stepped down to account for this.

[b]N/A: not applicable.

Study Design

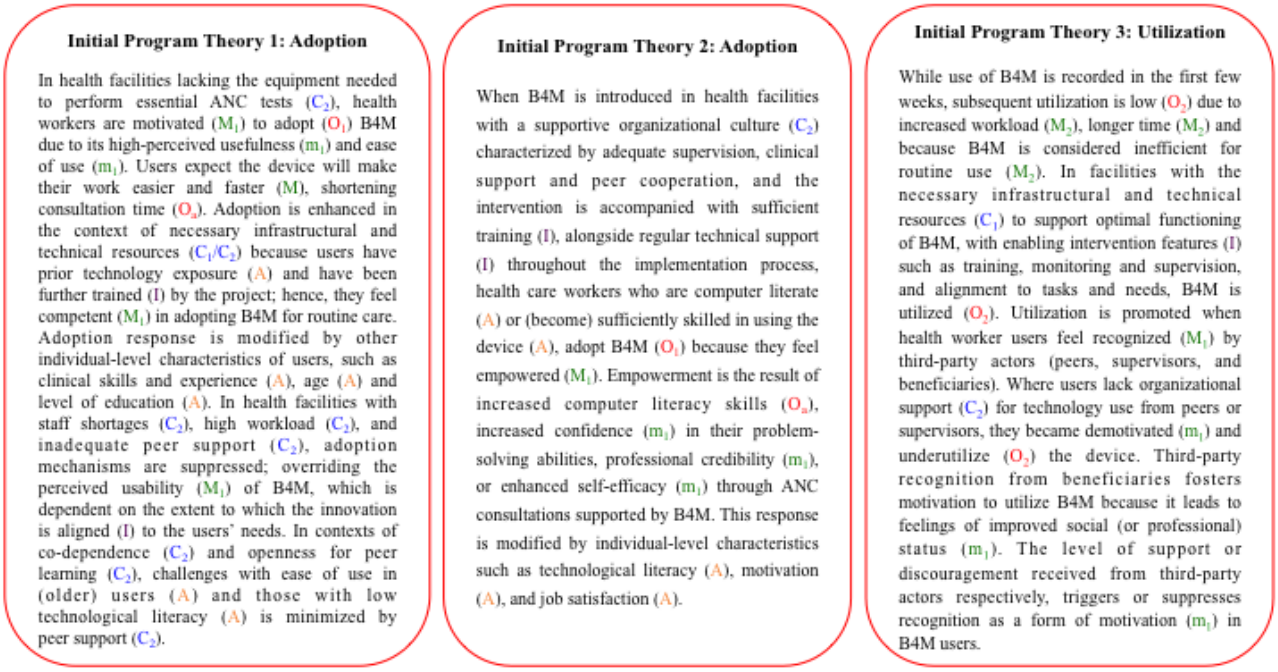
We employed a multiple case study design, defining a case as 1 B4M health facility [13]. Informed by knowledge of the project, ANC volume per facility and trend analysis on adoption (first 2 months) and utilization (continued or prolonged use over time) of the device over a 10-month period (unpublished data [10]), health facilities were classified as low (average number of screenings <15 per month), moderate (average number of screenings ≥16 and ≤40 per month), or high (average number of screenings ≥41 and ≤75 per month) adoption and utilization (Table 1). Cases were subsequently grouped into 3 usage clusters: low, moderate, and high, whereby the term usage is a composite term describing adoption and utilization (Table 2). No health facility fell under the high usage cluster, which was recognized as the ideal state. The evaluation sought to understand usage variation between low and moderate usage clusters and reflect on how a high usage state may be attained in implementing an improved prototype.

Evaluation Methodology

Realist evaluation is a theory-based approach for opening the black box on complex interventions [14,15]. It has shown promise in unraveling explanations for complex interventions in health, international development, and technological innovation [16-18]. It involves an iterative process beginning and ending with program theories, systematically moving from the specific to the abstract, described as “climbing the ladder of abstraction” [19,16]. Realist methodology is suited for evaluating B4M because it is method neutral and can aid an in-depth understanding of the explanatory processes for program outcomes as well as in the identification of implicit and explicit mechanisms underlying them.

Due to its theoretical underpinning and applicability in real-life settings, realist methodology was applied to assess differences between low and moderate B4M usage clusters. This involved developing and subsequently testing initial program theories using qualitative and quantitative data. Identified causal explanations underlying variation in mHealth usage between clusters were framed in configurations that showed the interrelationship between the *Intervention*, implementation *Context*, participating *Actors*, explanatory *Mechanisms*, and *Outcomes*. Simply put, *ICAMO* configurations. Using this analytical heuristic, 2 main layers of context may be differentiated: the broad external environment in which interventions are situated (C_1) and the health system or health facility setting in which mobile technology is introduced (C_2). Where mechanisms broadly refer to the reasoning and responses to the B4M intervention underlying observed outcomes, main mechanisms (M) were differentiated from subexplanatory mechanisms (m).

Figure 1. Initial program theories. Features and characteristics of the intervention- (I); Contextual factors are denoted (C_1) and (C_2) for environmental and health system context respectively; Outcomes are denoted (O_1) or (O_2) representing adoption and utilization respectively; Mechanisms are identified (M_1) or (M_2) following the outcomes they are linked to, with related explanatory mechanisms further depicted (m_1) or (m_2); Actor or user characteristics are denoted (A); (O_a) represents additional outcomes. ANC: antenatal care; B4M: Bliss4Midwives.



Initial Program Theories

The initial program theories of the B4M intervention, which describe how the intervention was expected to work, were developed using a 2-pronged approach:

- A realist review of how mHealth influences performance of maternal health workers in LMICs was conducted. A total of 4 factors necessary for the successful adoption and utilization of mHealth were identified: general environmental context, organization of the health system, intervention factors, and individual factors [9].
- To ensure that the initial program theories were aligned to the unique prescription of the B4M intervention, we refined literature-based theories by analyzing the research protocol and interim progress reports. We also conducted a focus group discussion and follow-up interviews with members of the program consortium, resulting in 3 initial program theories (Figure 1). These processes informed data collection tools and guided analysis.

Data Collection

Using quantitative and qualitative methods, the 3 candidate initial program theories were empirically tested. Data collection activities are presented in Multimedia Appendix 1 and summarily involved:

- A total of 24 semistructured interviews with device users, health facility managers, local program managers, and district health information officers trained to provide technical support. Interviews were conducted in English and lasted between 22 and 122 min (mean=60 min).
- A total of 14 usability questionnaires measuring perceived usefulness and ease of use of B4M using 12 items each, developed from standardized tools [20,21] and administered to device users (Multimedia Appendix 2). Respondents selected options from strongly disagree (1) to strongly agree (5) on a 5-point Likert scale, totaling 12 to 60 points per construct.
- Health facility checklists at 6 facilities, to assess their capacity to provide ANC services, referral, or management of emergencies (Multimedia Appendix 3). Observation of ANC service provision was conducted in 5 facilities.
- A focus group discussion with project implementers.
- A theory-validation meeting with 16 B4M users.

All interviews and meetings were conducted in English, audio recorded, and transcribed verbatim.

Data Analysis

For the data on usability, negative statements were reverse coded, and raw scores were exported to SPSS. Nonparametric *t* test was used to compare perceived usefulness and perceived ease of use of B4M between clusters. Interview transcripts as well as observation and field notes were analyzed using realist thinking, applying an interpretive lens to build a casual web of explanations from multiple strands of evidence [22]. Using abductive inference, we started from the main outcomes of interest (adoption and utilization) and *worked backward* to trace plausible underlying explanations. We queried the data for mechanisms of perceived usefulness, perceived ease of use and empowerment (self-efficacy and confidence) for adoption, and the mechanism of recognition for utilization, while being open to new configurations.

A cumulative stepwise approach applying inductive and deductive reasoning was employed. First, aided by an Excel spreadsheet, we entered information on each health facility that *spoke* to elements of the ICAMO configuration into rows and columns, including supporting quotes. Furthermore, previous analysis has shown that over time, the intervention itself can become a new contextual layer within the study setting [9]. Nevertheless, we chose to differentiate the intervention (I) from the existing contextual factors (C_1 or C_2) to clarify the resources and support that are specifically introduced by B4M. As our data were closer to the project itself than to the broader environmental context (C_1), we did not have sufficient strands of evidence on this level. Next, the realist thinking of "if C, then O, because M, for A" was applied to develop ICAMO configurations for each cluster. This involved grouping similar patterns and corroborating or voiding strands of preliminary evidence. Although most evidence strands manifested to varying degrees in each facility, when these were not sufficient to explain usage behavior, they were discarded from the configuration. Theory testing and refining were incremental; data from the low usage cluster were first assessed and then compared with data from the moderate usage cluster. Finally, a cross-case comparison between clusters was used to develop refined program theories.

Ethical Considerations

Study approval was granted by the Navrongo Health Research Centre Institutional Review Board (approval ID: NHRCIRB18) and the EMGO+ Scientific Committee of the Amsterdam Public Health Institute (reference number: WC2017-026). Before all interviews, written consent was secured using informed consent forms.

Results

Usability Statistics

Respondents' characteristics and usability scores are presented in Multimedia Appendix 4. Acknowledging individual variations, the perceived usefulness and perceived ease of B4M use were relatively high in all facilities (range 39.0-58.0). The *t* test showed no statistically significant differences between the 2 usage clusters (Table 3).

Next, we present the refined program theories under each outcome of interest in narratives of ICAMO configurations. Intervention features are marked "(I)" factors related to the health system context as "(C_2)" evaluation outcome "(O_1)" represents adoption and "(O_2)" utilization, whereas "(O_a)" represents additional outcomes. Mechanisms are identified "(M_1)" or "(M_2)" following the outcomes they are linked to, with related explanatory mechanisms further marked "(m_1)" or "(m_2)." Actor or user characteristics are marked "(A)." Explanations are included in the narrative, noting differences between cases and usage clusters. An overview of the realist analysis is depicted in Figure 2.

Adoption (O_1)

Adoption of B4M was characterized by an initial upward climb in both clusters. Differences, however, stemmed from experienced technical failures (I), complete or partial presence of an alternative point-of-care device or onsite laboratory (C_2), and dispositions of individual users (A). In health facilities with limited capacity to perform basic ANC screening tests (C_2), trained midwives and community health workers (I) were enthusiastic (m_1) to adopt B4M (O_1). This was due to its novelty (M_1) as a noninvasive automated device (I) and in anticipation of service delivery benefits, which they considered important (m_1) for providing focused ANC:

> *After training, we were just eager [...] If we don't support whatever the project's intention is, it will not be realised. Then it means the support we could have also gotten from it will not come.* [Facility C]

In facilities A, E, and F, long-standing relationships with local project partners (C_2) played a role in their selection as project sites (I). Their adoption response was transactional (M_1), triggered by a sense of obligation (m_1) to the project partner and by pride from being selected (m_1). Where alternative screening options were not functionally reliable (C_2), were not trusted (C_2); as was the case in facility F, required a longer turnaround time (C_2); as in facility A, or when screening was a paid service (C_2), users were motivated (M_1) to adopt B4M (O_1). This is because they considered it to be a necessary alternative (m_1), a trustworthy expert (m_1), a time-efficient resource (m_1), and a cost-effective substitute (m_1):

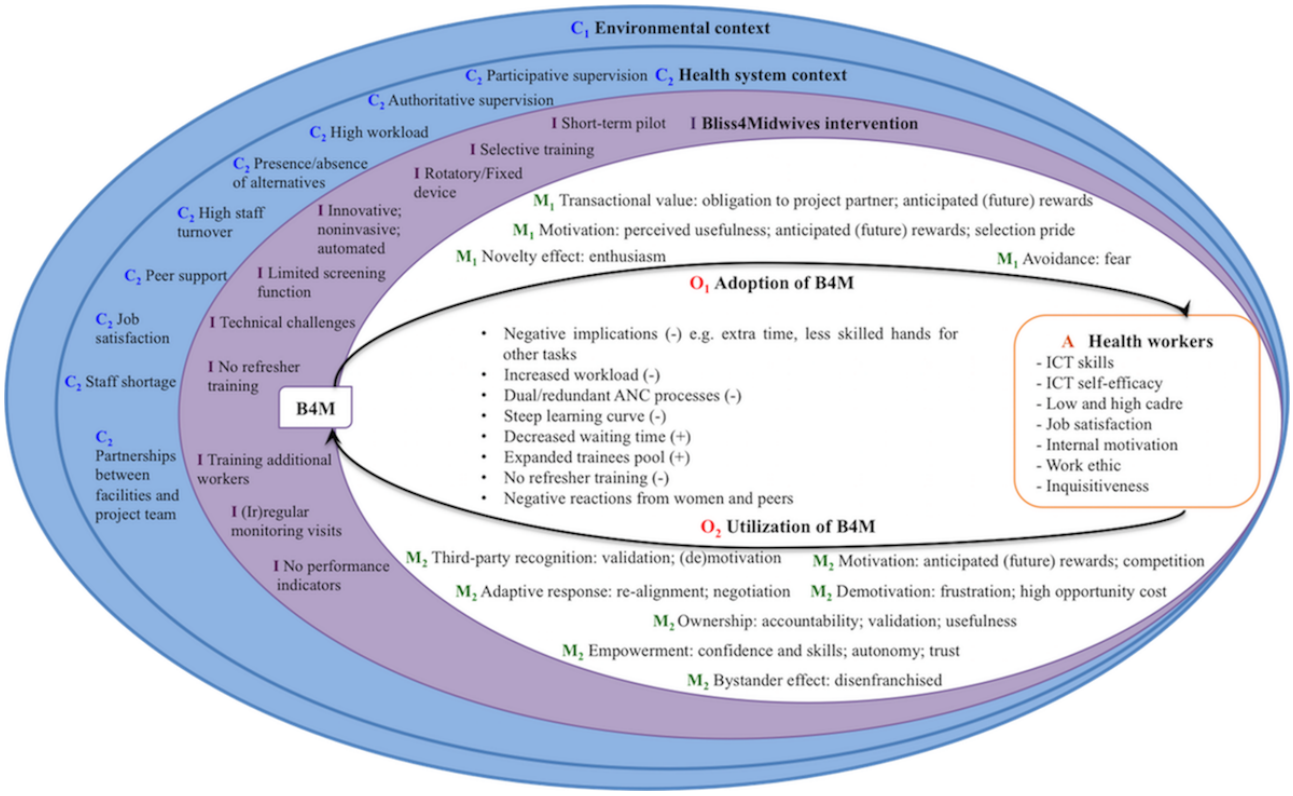
> *Supposing I am here alone, I can't talk to anybody…the machine can tell me what to do [...] So, it is more accurate to use it.* [Facility F]

Table 3. *t* test for equality of means on usability assessment.

Construct	Low usage cluster[a], mean (SD)	Moderate usage cluster[a], mean (SD)	P value
Perceived usefulness	52.0 (4.6)	47.0 (6.2)	.11
Perceived ease of use	50.4 (7.1)	51.4 (8.1)	.81

[a]P<.05.

Figure 2. Summary of findings. Ecosystem of ICAMO factors underlying the adoption (O_1) and utilization (O_2) of B4M within a complex context (concentric circles C_1 and C_2) and features of the B4M intervention (I). M_1 and M_2 are mechanisms related to outcomes O_1 and O_2 , mediated by user characteristics (A). Bullet points highlight other facilitating (+) or inhibitory (–) factors influencing usage behavior. ANC: antenatal care; B4M: Bliss4Midwives; ICT: information and communication technology.



Irrespective of age, baseline technological skills, years of clinical experience, and the additional time use of B4M required, users with high technological self-efficacy (A) or willing to acquire technological skills (A) adopted the device despite an initial (1-2 weeks) difficult learning curve. Workers with low technological self-efficacy (A) avoided B4M (M_1) out of fear (m_1) of damaging it and low effort expectancy, although they reported sufficient training (I) and encouragement from peers and the project (C_2). To fill training attrition gaps due to user disinterest (A) and staff turnover (C_2), additional training (I) was organized for other health workers, thereby expanding the pool of trainees (O_a) who were mostly younger, lower cadre personnel (ie, community health nurses and nursing assistants who are trained professionals):

> *So, when I went for the training on the kit, it was a hell with me and the nurses [...] I don't know the thing and I don't want to make mistakes [...]We have people who know the thing, so why should I be forcing my head to be doing all these things, when these young girls are sitting there?* [Facility E]

Utilization (O_2)

In response to contextual and program factors, postadoption utilization of B4M was explained by the dominance of either suppressive or supportive mechanisms triggered at facility and individual levels. Supportive mechanisms identified in the moderate usage cluster were less prominent or lacking in the low usage cluster.

Implementation Strategy: Fixed Versus Rotatory

The rotating strategy (I) between facilities B and C (from the low usage cluster) required ownership of the rotation process (M_2) and necessary resources (C_2) including fuel and motorbike. Although facility B had more resources, users in facility C demonstrated higher ownership—using the device more consistently when present. As ANC schedules between rotatory sites sometimes overlapped (C_2), the device was often absent from points of need or present without use. More importantly, the rotation strategy not only affected the number of screenings in facilities B and C but also whether a woman repeatedly benefitted from its use throughout pregnancy:

> *It wasn't that (convenient), [...] (sigh) because some women, you have used it on them and then when they come back, it is not there. The next time, it becomes a problem.* [Facility C]

All facilities in the moderate usage cluster had a fixed device (I), and health workers quickly gained dexterity (O_a), when they felt supported by supervisors and peers (C_2), alongside other enabling contextual factors.

Empowerment

Health workers felt empowered (M_2) by B4M in 2 ways. First, users increasingly gained confidence and skills (m_2) in device use and ANC referral, shortening the time needed per screening (O_a). In lower cadres who did not know what to do (C_2) and higher cadres (ie, trained professional midwives) who overlooked warning signs due to work pressure (C_2), B4M was

used to validate hunches and keep users alert (I) because they trusted its accuracy (m_2). Second, both facilities with and without alternatives experienced more autonomy (m_2) and a decreased need for diagnostic referrals, which previously delayed the care cycle:

> *Even though you are experienced and you know what to do, you may be tired or distracted, so the device will not allow you miss a critical case.* [Facility D]

Users with low technological skills (A), without refresher training (I), in facilities with a rotatory implementation (I), or inconsistent usage (O_a) were demotivated (M_2) and frustrated (m_2) because they frequently forgot how to navigate the system (O_a), contributing to nonpartial or partial use (O_2). By limiting access (I), the rotating strategy effectively suppressed the empowering effect of B4M:

> *When you don't use it for some time, you forget.* [Data validation meeting]

Realignment and Negotiation

Misalignment of the device (I) to existing work processes or limited workspaces made usage frustrating (m_2). Evidence from the moderate cluster showed that if trainees felt compelled or were otherwise motivated by current or anticipated benefits to use B4M, their adaptive response (M_2) was realignment (m_2) or negotiation (m_2). In the low usage cluster, the response was rejection and abandonment (m_2):

> *It was cumbersome because of our setting. Where to put (the device) was a problem [...] It was really hindering us, because who would be standing and doing all this?* [Facility E]

> *Oh, it is an interruption but since we've been able to manage it, it is no more an interruption again.* [Facility D]

Realigning workflow as a coping mechanism to B4M involved peer-training other (lower cadre) health workers (O_a) who showed keen interest and were inquisitive (A). Peer-trained users (A), however, had low confidence (m_2) in the thoroughness of training, manifesting low ownership (M_2). Realignment allowed for redistribution of roles (O_a) with at least two workers conducting ANC when the device was in use. This meant that midwives could focus on core maternal health tasks (palpation, deliveries, and counseling), whereas lower cadre staff operated the device. This strategy was not feasible in contexts where support staff had other fixed duties such as outreach visits (C_2), where only 1 midwife was available per time (C_2), or in contexts of high staff turnover or B4M-training attrition due to administrative leave or transfer (C_2). In facility D, users did not only manage their own expectations and avoid dual use of screening options but actively negotiated (m_2) B4M usage with beneficiaries (O_a):

> *You know, when human beings tune their mind to something, they expect only that. I told them that the machine will have to check everything for them and it will tell us what to do [...] In fact, now, we don't talk about it. When they come, everybody is relaxed.* [Facility D]

In contexts of professional isolation (C_2), low (supervisory) recognition (C_2), low job satisfaction (C_2), and high workload (C_2), it was not sustainable for users to persevere against all odds, which manifested in low utilization (O_2).

Opportunity Cost and Competitive Edge

Facilities with high volumes of ANC attendees (C_2), multiple service delivery demands (C_2), or high staff turnover or shortages (C_2) manifested suboptimal utilization of B4M (O_2) despite high perceived usefulness (m_2) and high perceived ease of use (m_2). This was linked to the demotivating (M_2) high opportunity costs (m_2) of usage, including the following: (1) ANC consultations took longer; (2) B4M did not completely remove the need for diagnostic referral for other tests; and (3) B4M was used in addition to the usual ANC routine because it was regarded as a pilot intervention. Where B4M represented a partial solution (I) to a larger diagnostic need and was not fully integrated (I) into ANC workflow, duplication of processes made utilization burdensome (m_2), causing dissatisfaction (M_2) and decreased perception of its usefulness (m_2):

> *It's easy to do either the standard or B4M. It's the combination that is not easy [...] It helps you to waste a lot of your time. It's like the thing became not useful to us again.* [Facility F]

Health workers in moderate usage facilities took ownership (M_2) of the device and utilized it because of their strong work ethic (A), motivation (M_2) to meet service delivery needs, and expectation of appreciation (m_2) at project end. To defend their professional image and as a favor to their local program managers, these users had an internal drive to compete (m_2) and perform better than other facilities:

> *What I can say about the midwives here is that we take our work serious [...] Sometimes there are certain things you don't want to do, but when it comes to our work anything we have to do we do it.* [Facility F]

If users believed that project success and subsequent reward were based on the number of screening records per facility, utilization was higher, with less regard to follow-up screening of beneficiaries at each visit (O_a). Absence of project feedback on performance indicators (I) and lack of direct incentives (I) suppressed (in the low cluster) and dampened (in moderate cluster) the competitive edge:

> *So we needed that they should tell us that what we were doing is actually an important thing [...], then we would put our efforts in getting the whole thing done properly.* [Data Validation Meeting]

Third-Party Recognition

We found that recognition from third-party actors (M_2) as a form of external motivation was an important mechanism underlying utilization, and this derived from multiple sources: (1) peers who supported and encouraged device use, (2) pregnant

women who projected the value of the device to their trust in the health worker, (3) program staff who provided technical support and conducted monitoring visits, and (4) supervisors at facility and district levels. In facilities A and D, peers regarded B4M users as distinguished, belonging to an expert niche. This sometimes increased utilization motivation (m_2), but in many other cases, it caused tension (m_2) when peers felt that trainees had enjoyed preferential selection and benefits from the intervention. Peers, therefore, tagged B4M users as lazy or unserious:

> *I feel proud (when my colleagues call me computer woman).* [Facility D]
>
> *The perception is worse about you who went and learnt because you can now (do these things). But the thing is that you went and signed and took money (ie, participation and per diem during B4M training).* [Data Validation Meeting]

B4M users felt respected by pregnant women who showed increased confidence in health workers' professional credibility (m_2), especially in lower cadre workers (A). However, the comparatively longer time (I) it took compared with the standard ANC routine elicited negative reactions (m_2) manifested in body language or grumbling from pregnant women. This demotivated (m_2) users and led to decreased utilization (O_2):

> *Sometimes, the women think that you are doing it for them and so that kind of trust comes in [...] They are happy that it is madam midwife who is doing it for me, but not necessarily the bliss for midwife that is doing it. So, it sort of gives you that zeal to continue using it.* [Facility C]

Irregularity of monitoring visits (I) and technical problems (I) led to prolonged periods of nonuse (O_2) because users forgot (m_2) about the intervention and no longer considered it a priority (m_2). Due to easy geographical access (C_2) and strong preintervention collaboration (C_2), facilities A and D from the moderate cluster frequently received monitoring visits (I) from the project manager, which kept users *on their toes* (m_2) and stimulated ownership (M_2). It also made users feel validated (m_2) and not exploited by the project to extract usage data:

> *We didn't expect to see money. Money could be one of the things, but regular visits, calls and all those things; we were not getting it at all. So we just said "Aha, so the person just comes to take the (data) and goes away."* [Data Validation Meeting]

In facilities where workers feel unsupported by superiors (C_2) and where aspirations for career progression and professional development are not fostered (C_2), users were demotivated (m_2) and did not take ownership (M_2). The project, therefore, became a platform to silently protest job satisfaction through nonuse:

> *But sometimes it's really heart-breaking. Like why should I really waste my time doing this and at the end of the day nobody appreciates you? Even what we are supposed to work with is not there.* [Data Validation Meeting]

Ownership and Supervision Styles

Ownership (M_2) of device usage trickled down to users from higher-level actors at program, district, and facility levels, based on supervision styles (C_2). If authority figures did not demonstrate the importance of B4M, health workers were less inclined to use the device because they did not feel accountable (m_2) and felt discouraged and unappreciated (m_2). Authority figures in the moderate usage cluster showed more engagement with the program.

In facilities with firm hierarchical structures such as facility A, where users were accustomed to authoritative supervision (C_2), involvement of a high-ranking supervisor (C_2) imposed accountability and responsibility (m_2), reinforcing device use. In facility D, on the other hand, ownership was fostered by supportive participative supervision (C_2) in motivated health workers with high self-efficacy (A) in using technology:

> *Because it came and our matron called and said "I'm putting this thing in your hands, take care of it." So, because it was from her, we were doing it [...] And often the matron would come and ask "Are you people with the box? Are you ok?" Then the next day, again. So if you are not there and she comes and the box is lying there, there would be problem. So, we are always doing it.* [Facility A]

Bystanders and Champion Users

By training only a select number of staff in each facility (I), the project could not leverage collective ownership at facility level and some users felt unsupported by disenfranchised peers (m_2). Even when multiple persons were trained (I), in contexts of low-shared responsibility (C_2) and weak interpersonal relationships (C_2), a bystander effect (M_2) was observed. As seen in moderate usage facilities, responsibility for B4M was indirectly delegated to a *champion* user (A) who had strong internal motivation (A) and in whose absence (C_2) the device was not used (O_2). Nevertheless, given other competing priorities (C_2) and to balance the inconvenience of using the device, usage was restricted to 1 day a week or to a few hours in a day:

> *Yea, at first, the excitement was just too much. But when I trained this lady and she picked it very fast, then I stopped using it.* [Facility D]
>
> *If she is there, nobody will even tell her what to do. [...] I used the word passionate—if you have the zeal to work. For her, I think no matter the pressure she can be called at any time and she won't have any problem, unlike some others.* [Facility F]

Discussion

Overview

The theoretical bases of knowledge on adoption and postadoption have been largely developed in the field of management information systems with a focus on higher-income countries [23,24]. As digital innovation systems continue to

expand in LMICs, the implications of these theories in low-resource settings such as Ghana are making their way into the research agenda [8,25,26]. To our knowledge, this is the first study that applies a realist lens to elicit theory-based explanations on mHealth for maternal health services in LMICs. Our analysis confirmed some components of the initial program theories, voided others, and unveiled additional elements previously unaccounted for. Below, we reflect on key findings and their relevance to the science of mHealth implementation.

Principal Findings

In facilities with limited diagnostic capacity, motivated workers adopted B4M for its novelty and benefits, in contexts of existing collaborations and authoritative or participative supervision styles. Although technology novelty triggered supportive adoption mechanisms, we found that the actual utilization of the device was the most important phase of the usage cycle [23]. Above-average usability scores from most health facilities did not fully explain variation between usage clusters, confirming the disconnect between usability and actual use [27]. Fear, enthusiasm, and high expectations for service delivery, especially in the absence of alternatives, better explained low to moderate adoption of B4M. With increased experience of use, we found that the initial emotive adoption response was replaced by rational behavior in the utilization phase: perceived usefulness being overshadowed by experienced contextual difficulties. Saccol and Reinhard describe this contrast between the perceived magic of technology and the disappointment of its limitations in the real world, which dampens users' initial enthusiasm [28]. Although the program designers' expectation was that all or most facilities would operate under the high usage cluster, that is, high adoption followed by high utilization, the identified supportive or suppressive mechanisms within and between cases shed light on why no health facility fell under this ideal state.

Realignment of mHealth to workflow and beneficiary expectations of ANC was identified as a crucial adaptive mechanism for its utilization. In addition to intrinsic motivation and a sense of accountability in users, utilization was influenced by mechanisms triggered in third-party actors. Negative reactions from pregnant women, bystander effect in peers, and low support or ownership from supervisors and program managers caused low utilization. mHealth adoption has been described as a social process [29], which may explain the strong third-party effect, although its influence has been specifically linked to contexts of mandatory technology use [30]. Despite perceived usefulness and user motivation, utilization mechanisms were suppressed by intervention-related (technical challenges, device rotation, lack of performance feedback, and refresher training), contextual (staff turnover, high workload, competing priorities, and low job satisfaction), and individual (low technological self-efficacy and knowledge) factors. Champion users displayed moderate but inconsistent adoption-utilization behavior, by taking ownership of the device, defying usage barriers. This adaptive behavior of users as a distinguishing factor in usage behavior is in line with other studies [27].

Contrary to the expectation that usage behavior was related to age, we found that internal and external motivations and technological self-efficacy were stronger explanatory factors. However, these are linked to age as a predictor of technology usage. Previous research confirms that older users have lower technological self-efficacy and are intimidated by the steep learning curve, especially when they have low baseline technological skills and inadequate learning support [31-33]. Although we found that empowerment was triggered in the utilization phase, adoption behavior has been shown to predict utilization response [34]. This mechanism might, therefore, manifest in both phases.

Implications for Bliss4Midwives Prototype II and Other Mobile Health Interventions

Beyond initial training, introducing technology requires careful planning and adaptation in low-resource settings where not many users experience job satisfaction or have adequate technological training as part of their professional competencies [35]. Admittedly, most factors related to the intervention context and actors such as service delivery demands, workforce shortages, and staff turnover are beyond the control of implementation teams. Nevertheless, these will have to be constantly negotiated especially in the utilization phase, with a responsive implementation strategy that supports workflow alignment and integration, which are crucial to the success of mHealth [35,36]. A preintervention situation analysis that takes our findings into account would go a long way in ensuring that future interventions are holistic and context-specific. A practical starting point to this could involve incorporating ICAMO elements into applicable implementation research frameworks such as the Consolidated Framework for Implementation Research, which incorporates multilevel factors and is adaptable through program cycles [37].

The temporal nature of pilot projects imposes a false sense of reality. Although users may briefly accommodate the innovation, they will be less invested in making long-term commitments requiring individual and organizational realignment, for short-term gains. In addition to being user-centered and accounting for the context, it is imperative that multiple stakeholder perspectives are leveraged during innovation design [6,7,35]. High-ranking supervisors might seem distant from the usage process but could compel or foster accountability and usage. They can also support adaptive strategies to integrate technologies into routine practice, especially in contexts of hierarchical supervision [30,38].

Selective training of a few workers unintentionally limits collective ownership and accountability for usage behavior. All health workers involved in maternal health service delivery at each site should be trained on device use, with regular monitoring and supervision, and periodical refresher training to help sustain or improve technological self-efficacy and dexterity, consequently preventing frustration and utilization decline [39]. Closer supervision and attention will be necessary in users with lower baseline technological skills and self-efficacy. Although it may cause tension and resistance from higher cadre users or peers, workers who fit the typology of champion users should be identified and encouraged to serve

as opinion leaders within their health facilities. This would improve collective ownership, minimizing the bystander effect and optimizing social pressure [39,40].

The value of preexisting collaborations between the local partner organizations and health facilities and other administrative bodies remains crucial to gain access and influence, motivate, and encourage users. However, sustainability of transactional responses as a favor to program managers is doubtful. Financial incentives as a mechanism for behavior change have elicited mixed reports [41]. Indirect incentives such as encouragement, recognition, and support, which were highly desired and valued by B4M users, can, however, be promoted. To leverage the competitive mechanism and give users regular performance feedback, respectively, the design features of B4M prototype II could include gamification and dashboard analytics.

Limitations

At the time of data collection, the device was in limited use in 3 of the 6 sites, with some respondents reporting not using the device for up to 5 months. This introduced recollection bias, in addition to socially desirable answers. Furthermore, not all user experiences were captured because a small number of trained users were unavailable for interviews. By triangulating data from multiple sources and interviewing at least two users per site, we attempted to compensate for these. The data validation workshop and dissemination meeting also informed group consensus on our findings. A realist approach is best applied throughout the life cycle of a project, from design to evaluation and reporting [22]. The nature of B4M as a short-term pilot, in addition to other constraints, restricted this possibility. Nevertheless, by developing and testing 3 initial program theories, the refined theories as a result of our analysis are sufficient for the next phase of prototype development.

Conclusions

This study shed light on optimal conditions necessary for B4M to thrive in a complex social and organizational setting. Evidence on the growth and potential of mHealth in improving service delivery, especially in a critical domain such as maternal health, may have overshadowed important individual, health system, and implementation factors that preclude its alignment in specific contexts and by certain user types. Beyond usability and viability studies, advocates of innovative technologies for maternal care need to consider how contextual factors, such as existing collaborations and supervision styles, trigger supportive mechanisms that influence program outcomes. This knowledge can be used to design and implement mHealth in similar settings. In addition to informing scale-up of the B4M prototype, our results and approach highlight the need for interventions that are guided by research methods that account for complexity.

Acknowledgments

The authors thank the research assistants, translators, transcribers, study participants, the district administrators, and health facility personnel who were involved in the study. The authors acknowledge the support of the B4M implementing organizations in Ghana and the Netherlands (Association of Church Development Projects, Ghana; Presbyterian Health Services-North, Ghana; Cordaid; Relitech B.V; The Netherlands Organisation for Applied Scientific Research; Simavi; and Enviu Netherlands). The first author thanks Professor Vincent De Brouwere who financially supported the stakeholders' dissemination meeting. The Erasmus Mundus Joint Doctorate Fellowship Specific Grant Agreement 2015-1595, which IOA is a beneficiary of, financed this study. Additional financial support was received from the B4M implementing partners through the Life Sciences and Health for Development (LS&H4D) grant: Number- LSH14GH16, which funded the B4M intervention. Funders had no role in study design, analysis, or preparation of the paper.

Conflicts of Interest

NA and RAA work for organizations involved in the implementation of the B4M proof-of-concept and were actively involved in the project. Other authors declare that they have no competing interests.

Multimedia Appendix 1
Data collection activities.
[PDF File (Adobe PDF File), 28 KB - mhealth_v6i12e11468_app1.pdf]

Multimedia Appendix 2
Usability survey questionnaire.
[PDF File (Adobe PDF File), 64 KB - mhealth_v6i12e11468_app2.pdf]

Multimedia Appendix 3
Health facility checklist for antenatal care (ANC) service provision.
[PDF File (Adobe PDF File), 44 KB - mhealth_v6i12e11468_app3.pdf]

Multimedia Appendix 4
Characteristics of Bliss4Midwives (B4M) users and usability scores.

[PDF File (Adobe PDF File), 40 KB - mhealth_v6i12e11468_app4.pdf]

References

1. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
2. World Health Organization. Geneva: World Health Organization; 2018. Classification of digital health interventions URL: http://apps.who.int/iris/bitstream/handle/10665/260480/WHO-RHR-18.06-eng.pdf?sequence=1 [accessed 2018-07-02] [WebCite Cache ID 70cDnlfDg]
3. Modi D, Gopalan R, Shah S, Venkatraman S, Desai G, Desai S, et al. Development and formative evaluation of an innovative mHealth intervention for improving coverage of community-based maternal, newborn and child health services in rural areas of India. Glob Health Action 2015;8:26769 [FREE Full text] [Medline: 25697233]
4. 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]
5. 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]
6. Cho Y, Lee S, Islam SMS, Kim S. Theories applied to m-Health interventions for behavior change in low- and middle-income countries: a systematic review. Telemed J E Health 2018 Feb 13. [doi: 10.1089/tmj.2017.0249] [Medline: 29437546]
7. Morgan B, Hunt X, Tomlinson M. Thinking about the environment and theorising change: how could Life History Strategy Theory inform mHealth interventions in low- and middle-income countries? Glob Health Action 2017;10(1):1320118 [FREE Full text] [doi: 10.1080/16549716.2017.1320118] [Medline: 28617198]
8. 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]
9. Abejirinde IO, Ilozumba O, Marchal B, Zweekhorst M, Dieleman M. Mobile health and the performance of maternal health care workers in low- and middle-income countries: a realist review. Int J Care Coord 2018 Jun 19:205343451877949. [doi: 10.1177/2053434518779491]
10. Abejirinde IO, De Brouwere V, van Roosmalen J, van der Heiden M, Apentibadek N, Bardaji A, et al. Viability of diagnostic decision support for antenatal care in rural settings: findings from the Bliss4Midwives Intervention in Northern Ghana. J Glob Health 2018 (Accepted September 2018, forthcoming).
11. Abejirinde IO, Douwes R, Bardají A, Abugnaba-Abanga R, Zweekhorst M, van Roosmalen J, et al. Pregnant women's experiences with an integrated diagnostic and decision support device for antenatal care in Ghana. BMC Pregnancy Childbirth 2018 Jun 05;18(1):209 [FREE Full text] [doi: 10.1186/s12884-018-1853-7] [Medline: 29871596]
12. Mechael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. The Earth Institute, Columbia University. 2010 May. Barriers and gaps affecting mHealth in low and middle income countries: Policy white paper URL: http://crcrth670-f11-sullivan.wikispaces.umb.edu/file/view/mHealth_EARTH+Institute_Low+Mid+Income+Co's_Barriers_WhitePaper_Columbia+U+'10.pdf/280818744/mHealth_EARTH+Institute_Low+Mid+Income+Co's_Barriers_WhitePaper_Columbia+U+'10.pdf [accessed 2018-07-02] [WebCite Cache ID 70cFxoQZk]
13. Mills AJ, Durepos G, Wiebe E. Encyclopedia of Case Study Research. United States of America: Sage Publications Inc; 2010.
14. Pawson R, Tilley N. Realist Evaluation. UK: British Cabinet Office; 2004.
15. Astbury B, Leeuw FL. Unpacking black boxes: mechanisms and theory building in evaluation. Am J Eval 2010 Sep 13;31(3):363-381. [doi: 10.1177/1098214010371972]
16. Punton M, Vogel I, Lloyd R. Centre for Development Impact. Brighton: Centre for Development Impact; 2016 Apr. Reflections from a realist evaluation in progress: scaling ladders and stitching theory URL: https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/11254/CDIPracticePaper_18.pdf [accessed 2018-07-02] [WebCite Cache ID 70cHsG2vD]
17. Marchal B, Van Belle S, De Brouwere V, Witter S. Studying complex interventions: reflections from the FEMHealth project on evaluating fee exemption policies in West Africa and Morocco. BMC Health Serv Res 2013 Nov 08;13:469 [FREE Full text] [doi: 10.1186/1472-6963-13-469] [Medline: 24209295]
18. 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]
19. Pawson R, Tilley N. Realistic Evaluation. London: Sage Publications Inc; 1997.
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
21. Yen PY, Wantland D, Bakken S. Development of a customizable health IT usability evaluation scale. AMIA Annu Symp Proc 2010;2010:917-921 [FREE Full text] [Medline: 21347112]

22. Pawson R. The Science of Evaluation: A Realist Manifesto. London: Sage Publications; 2013.
23. Jasperson J, Carter PE, Zmud RW. A comprehensive conceptualization of post-adoptive behaviors associated with information technology enabled work systems. ACS Chem Biol 2014 Nov 21;9(11):2447-2458. [doi: 10.2307/25148694]
24. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
25. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
26. Kaphle S, Chaturvedi S, Chaudhuri I, Krishnan R, Lesh N. Adoption and usage of mhealth technology on quality and experience of care provided by frontline workers: observations from rural india. JMIR Mhealth Uhealth 2015 May 28;3(2):e61 [FREE Full text] [doi: 10.2196/mhealth.4047] [Medline: 26023001]
27. Zheng K, Padman R, Johnson MP, Diamond HS. Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system. Int J Med Inform 2005 Aug;74(7-8):535-543. [doi: 10.1016/j.ijmedinf.2005.03.007] [Medline: 16043083]
28. Saccol AZ, Reinhard N. The Hospitality Metaphor as a theoretical lens for understanding the ICT adoption process. J Inf Technol 2006;21(3):154-164. [doi: 10.1057/9780230250611_15]
29. Straub ET. Understanding technology adoption: theory and future directions for informal learning. Rev Educ Res 2017 Jan 10;79(2):625-649. [doi: 10.3102/0034654308325896]
30. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
31. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. Comput Human Behav 2013 Jul;29(4):1715-1724. [doi: 10.1016/j.chb.2013.02.006]
32. Kuerbis A, Mulliken A, Muench F, A. Moore A, Gardner D. Older adults and mobile technology: factors that enhance and inhibit utilization in the context of behavioral health. Ment Health Addict Res 2017;2(2). [doi: 10.15761/MHAR.1000136]
33. Magsamen-Conrad K, Upadhyaya S, Joa CY, Dowd J. Bridging the divide: using UTAUT to predict multigenerational tablet adoption practices. Comput Human Behav 2015 Sep 01;50:186-196 [FREE Full text] [doi: 10.1016/j.chb.2015.03.032] [Medline: 25937699]
34. Karahanna E, Straub DW, Chervany NL. Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Q 1999 Jun;23(2):183. [doi: 10.2307/249751]
35. Peiris D, Miranda JJ, Mohr DC. Going beyond killer apps: building a better mHealth evidence base. BMJ Glob Health 2018;3(1):e000676 [FREE Full text] [doi: 10.1136/bmjgh-2017-000676] [Medline: 29527353]
36. Granja C, Janssen W, Johansen A. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 1;20(5):e10235. [doi: 10.2196/10235]
37. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
38. Staples DS, Seddon P. Testing the technology-to-performance chain model. J Organ End User Comput 2004;16(4):17-36. [doi: 10.4018/joeuc.2004100102]
39. Sezgin E, Özkan-Yildirim S, Yildirim S. Understanding the perception towards using mHealth applications in practice. Inf Dev 2016 Dec 19;34(2):182-200. [doi: 10.1177/0266666916684180]
40. Hao H, Padman R, Telang R. Physician's Usage Of Mobile Clinical Applications In A Community Hospital: A Longitudinal Analysis Of Adoption Behavior. In: Association for Information Systems Electronic Library. 2013 Mar 19 Presented at: UK Academy for Information Systems Conference; 2013; United Kingdom p. 1-22 URL: http://aisel.aisnet.org/ukais2013/11
41. Lynagh MC, Sanson-Fisher RW, Bonevski B. What's good for the goose is good for the gander. Guiding principles for the use of financial incentives in health behaviour change. Int J Behav Med 2013 Mar;20(1):114-120. [doi: 10.1007/s12529-011-9202-5] [Medline: 22094998]

Abbreviations

ANC: antenatal care
B4M: Bliss4Midwives
ICAMO: intervention-context-actor-mechanism-outcome
LMICs: low- and middle-income countries
mHealth: mobile health

Edited by G Eysenbach; submitted 03.07.18; peer-reviewed by J Hensel, J Lynch; comments to author 06.09.18; revised version received 29.09.18; accepted 04.10.18; published 21.12.18.

Please cite as:
Abejirinde IOO, Zweekhorst M, Bardají A, Abugnaba-Abanga R, Apentibadek N, De Brouwere V, van Roosmalen J, Marchal B
Unveiling the Black Box of Diagnostic and Clinical Decision Support Systems for Antenatal Care: Realist Evaluation
JMIR Mhealth Uhealth 2018;6(12):e11468
URL: http://mhealth.jmir.org/2018/12/e11468/
doi:10.2196/11468
PMID:30578177

Original Paper

Social Networking App Use Among Primary Health Care Professionals: Web-Based Cross-Sectional Survey

Francesc X Marin-Gomez[1,2,3,4], MD, BSc (Comp Eng); Francesc Garcia Cuyas[4,5], PhD, MD; Ramon Reig-Bolano[4,6], Prof Dr; Jacobo Mendioroz[2,3], MPH, MD; Pere Roura-Poch[2], MPH, MD; Margalida Pico-Nicolau[2,7], MD; Josep Vidal-Alaball[2,3], MPH, MD

[1]Servei d'Atenció Primària d'Osona, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Vic, Spain

[2]Unitat de Suport a la Recerca Catalunya Central, Institut Universitari d'Investigació en Atenció Primària Jordi Gol, Sant Fruitós de Bages, Spain

[3]Health Promotion in Rural Areas Research Group, Institut Català de la Salut, Sant Fruitós de Bages, Spain

[4]Digital Care Research Group, Centre for Health and Social Care Research, Universitat de Vic–Universitat Central de Catalunya, Vic, Spain

[5]Department of Information and Communications Technology in Health, Universitat de Vic–Universitat Central de Catalunya, Vic, Spain

[6]Department of Engineering, Universitat de Vic–Universitat Central de Catalunya, Vic, Spain

[7]Centre d'Atenció Primària Sant Quirze de Besora, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Quirze de Besora, Spain

Corresponding Author:
Francesc X Marin-Gomez, MD, BSc (Comp Eng)
Servei d'Atenció Primària d'Osona
Gerència Territorial de la Catalunya Central
Institut Català de la Salut
Pl Divina Pastora, 7
Vic, 08500
Spain
Phone: 34 938890222
Email: xavy.marin@gmail.com

Abstract

Background: Several studies have been conducted to analyze the role social networks play in communication between patients and health professionals. However, there is a shortage of studies in relation to communication among primary health professionals, in a professional context, using the various mobile phone apps available.

Objective: The objective of our study was to explore mobile phone social networking app use among primary health care professionals for work-related purposes, by comparing the most widely used apps in the market.

Methods: We undertook a cross-sectional study using an anonymous Web survey among a convenience sample of 1635 primary health care professionals during August and September 2017.

Results: Of 483 participants in the survey, 474 (98.1%, 95% CI 97.1%-99.4%) were health professionals who commonly accessed social networking sites and 362 (74.9%, 95% CI 71.1%-78.8%) accessed the sites in a work-related context. Of those 362 respondents, 219 (96.7%, 95% CI 94.8%-98.5%) preferred WhatsApp for both personal and professional uses. Of the 362 respondents who used social networking sites in a work-related context, 276 (76.2%, 95% CI 71.9%-80.6%) rated social networking sites as useful or very useful to solve clinical problems, 261 (72.1%, 95% CI 67.5%-76.7%) to improve their professional knowledge, and 254 (70.2%, 95% CI 65.5%-74.9%) to speed up the transmission of clinical information. Most of them (338/362, 94.8%, 95% CI 92.5%-97.0%) used social networking sites for interprofessional communications, and 204 of 362 (56.4%, 95% CI 51.2%-61.5%) used them for pharmacological-related consultations.

Conclusions: Health professionals frequently accessed social networking sites using their mobile phones and often for work-related issues. This trend suggests that social networking sites may be useful tools in primary care settings, but we need to ensure the security of the data transfer process to make sure that social networking sites are used appropriately. Health institutions need to increase information and training activities to ensure the correct use of these tools.

(JMIR Mhealth Uhealth 2018;6(12):e11147) doi:10.2196/11147

KEYWORDS

mHealth; social networking; smartphone; attitude; primary health care; telemedicine; cell phone use

Introduction

Background

The introduction and development of mobile technology and the expansion of social networking are changing social relationships and modifying behaviors and attitudes, especially among the younger generations. Mobile phones are not only used as working tools, but often continue to be used for similar purposes at home, thus extending the normal working hours. In 2016, a survey showed that 90.4% of general practitioners owned a mobile phone with 1 to 3 medical-related apps [1]. There are not many studies on the current use of social networking among primary care professionals and even fewer comparing the use of the various apps available in the market.

Social networking sites (SNSs) are Web-based services that allow individuals to construct a public or semipublic profile within a bounded system to share information, ideas, personal messages, and other content in online communities [2]. Their forms of online communication vary greatly depending on their features (eg, photo-sharing or video-sharing capabilities, built-in blogging, and instant messaging technology). Most Web-based SNSs also support mobile interactions. There are considerably more mobile users than personal computer users, but how individuals decide to access SNSs (ie, through personal computers, iPads, tablets, or mobile phones) still depends on their own choice. Mobile phone apps allow for creating, sharing, and exchanging information, images, or videos with other users through a mobile portable format, and probably this is the main reason why the use of apps has grown rapidly among SNSs. In 2016, the number of apps downloaded to connected devices worldwide was 149.3 billion [3], and there were 3196 million active SNSs users, or about 42% of the global population. Of those, 2958 million, about 39% of the world's population, accessed SNSs through their mobile devices [4]. Apps such as WhatsApp, Facebook, YouTube, Instagram, Twitter, Spotify, Telegram, LinkedIn, or Snapchat have contributed to increase the number of mobile phone users and, in turn, the number of SNS users around the world. A recent report estimated that in Spain more than 15 million people aged between 16 and 55 years were active users of SNS technologies. WhatsApp and Facebook were the favorite SNSs of mobile phone users (76%) and people spent most of their time in WhatsApp. This increasing use of SNSs has been attributed to their being a way to socialize with peers (to chat or send messages) [5].

From a health care perspective, the use of mobile phones by clinicians could improve clinical communication, increase the practice of evidence-based medicine, enable access to information tools at the point of care, and improve education and research [6-9]. Apps designed for health professionals can be used to diagnose diseases, consult data on medications, perform clinical calculations, search scientific evidence, exchange clinical experiences, improve the management of chronic diseases, and conduct health care research [10]. The benefits of mobile technology for health professionals include the ability to make decisions more quickly and more reliably, thus improving the quality of health care and data management [11,12]. Apps to access SNSs stand out in the improvement of accessibility to health information, both as a support tool and for public health surveillance [13-16]. A greater connection with other professionals has been highlighted as one of the main benefits associated with the use of SNS media in the field of health care. Health professionals, from all categories, are using apps as social media for their professional development, to connect with colleagues, and to be up-to-date with the latest medical literature. Health care organizations around the world are taking initiatives to expand mobile health use and to demonstrate its efficiency [17]. The focus areas for future development of these technologies probably will be mobile telehealth and disease surveillance with SNS media and clinical decision support systems using machine learning. During the recent outbreak of Ebola virus in Africa, mobile phones and their apps were used for research, surveillance, and health education and to follow its dissemination [18,19]. It is likely that the use of apps in cases like these will increase in the future due to their potential to improve the health outcomes of patients in various health care settings.

Few studies have been undertaken on primary care professionals' use of various apps to access SNSs in a professional context. In a survey [20] on the use of mobile phones at work, in which about half the sample of 416 respondents were registered nurses, 58% of these nurses used their mobile phones at work; this use increased to 81% among physicians. The importance of this phenomenon and its foreseeable future impact require additional research on the use made by health care professionals in all types of social networks and devices. Primary care professionals (physicians, nurses, midwives, medical social workers, etc) are usually establishing the first contact with patients, and this type of SNS app, in a portable format, allows for remote support that seems useful and effective. For this reason, it is necessary to evaluate SNSs' impact and benefits perceived by members of primary care teams [21].

However, there is a growing fear and some controversies in relation to extending the use of social networks in health data communication contexts, which have their origin in the threat to privacy and confidentiality and the risk of misinformation, fake news, and the impersonation of professionals as recently reported in some media stories [22]. The increase in reports of these situations shows that these are risks to be taken into serious consideration [23]. If we add to this the risks associated with storing and transporting images, multimedia files, or text files on these mobile devices that go wherever the user goes and that often connect through low-reliability Wi-Fi networks, security risks rise exponentially. Lack of clarity on the boundaries between personal and professional life, increased risk of liability arising from the use of SNSs for professional purposes, low methodological rigor in studies on the use of social media, and poor accuracy, quality, and reliability of information are creating

serious doubts about extending SNS use among health care professions [6-9,11-15].

Objectives

Considering the need for more studies on SNS use and the growing trend toward the use of social networks to disseminate and discuss knowledge, we chose Bloom's taxonomy as an evaluative tool [24-26]. Our aim in this study was integrate this taxonomy into our exploration of primary health care professionals' use of SNSs and their main reasons for using them.

Methods

Design

This was a descriptive cross-sectional study to explore, through a Web-based survey, primary health care professionals' use of various social networking apps. The survey was conducted anonymously from August to September 2017.

Sample and Settings

The target population for the survey was a convenience sample of 1635 practicing primary health care professionals registered in SISAP (the Catalan acronym for Information Systems for Primary Care Services) [27] who worked in the central region of Catalonia, Spain. Those invited to take part in this study had an account to access electronic health records, had a valid email address, and had previously given consent to be contacted.

We distributed a link to the questionnaire by email. The email invited potential participants to take voluntary part in the questionnaire and explained the aim of the study.

Web-Based Survey

We used a voluntarily accessed survey developed using the Google Forms tool (Google LLC, Mountain View, CA, USA). Participants had access to the survey through a link sent in a personalized email. It was a closed survey, and no personal identification data were collected, thus protecting the confidentiality of participants. There was no financial incentive for participating in the study.

We carried out a pilot test with a group of 47 health care professionals (similar to the target group) to ensure the clarity of the questions and the validity of the rating scale. We introduced no major changes.

The first page of the survey informed participants about the total number of questions, the approximate response time, and the aim of the study. Participants were encouraged to contact the main investigator if they had any questions requiring clarification (contact details were also on the same page). The questionnaire, consisting of 10 multiple-choice questions, had only 1 conditional question referring to the use of social networks in a professional context. Depending on the response, it allowed access to a second section. The questionnaire was distributed in 3 distinct sections with all questions, except the last one, being mandatory. Some questions allowed free-text content (eg, apps used) and others allowed combined answers (options were "none" and "all"). As the questions were mandatory, incomplete answers were not registered. Only 1 response was allowed for each email sent. We kept no records of the respondents who quit the survey and analyzed only the completed questionnaires. We did not apply any statistical weighting.

Study Variables

The survey was divided into 3 sections: (1) type of apps used by the health care professional, (2) type of apps used in a professional context, and (3) professional perception of the benefits and impact of the apps on their clinical practice and professional development. For this last part, professionals were asked about the usefulness of using apps, classifying the answers as "not useful," "of little use," "useful," and "very useful" in terms of their benefits and impact. We used 8 distinct categories based on the 2 dimensions of Bloom's taxonomy (knowledge and cognitive processes), previously used in similar studies [28]: knowledge, clinical reasoning, critical thinking, clinical skills, problem solving, creativity, decision making, and outcome on the patient. An additional closed-ended question asked respondents to indicate whether they used SNSs for work-related purposes and, if they did, they were asked about the main reasons for this use.

The apps we chose to evaluate in the survey were those reported as being the most used in Spain (Facebook, WhatsApp, Twitter, Instagram, and other) [5]. We collected sociodemographic data (age, sex, education level, and work experience) using a demographic form. We determined professional category (physician, nurse, midwife, odontologist, social worker, or other) using a jobs checklist; we also recorded type of work (classified as "academic only," "clinical only," "academic and clinical, " or "other") and years of work experience. We did not evaluate the qualitative data collected for this study.

Ethical Considerations

We obtained ethical approval of the study from the University Institute for Primary Care Research Jordi Gol Clinical Research Ethics Committee (P17/174), Barcelona, Spain. The invitational email described the study's aims and procedures, and security and confidentiality of data. It also informed invitees about their right to decline to participate. The study observed data protection laws in effect at the time it was conducted.

Statistical Analysis

We made a bivariate comparison using the Pearson chi-square test between the professionals who used the apps in a professional context and those who did not, considering sociodemographic, professional knowledge, and attitude variables.

We performed a multivariate analysis using logistic regression, including the use of SNSs in a professional context as the dependent variable and taking $P<.05$ in the bivariate analysis. We also determined the adjusted odds ratio (adjusted OR). We conducted the analysis using IBM SPSS version 18 (IBM Corporation) and we reported the summary statistics as frequencies and percentages.

Results

Participant Characteristics

Of the 503 respondents, we included 483 as study participants and excluded 21 who had no clinical activity (ie, academic or research professionals; Figure 1).

The median age of the 483 participants was 45 years (SD 10.44, range 24-65). Most of them (393/483, 81.4%) were women and had a median work experience of 19 years (SD 10.88, range 1-58). In the professional category, 211 of the 483 participants (43.7%) were physicians and 215 (44.5%) were nurses. Of the 483 participants, 385 (75.6%) had a bachelor's, a graduate, or a diploma degree, and 118 (24.4%) had a master's or a doctoral degree.

App Use Analyses

To evaluate the frequency of use of the apps by the professionals surveyed, we considered the responses in which they had selected the option "often" or "constantly" as an indication of usual use. Among the 483 respondents, 474 (98.1%) were regular users of social networks and 362 (74.9%) also used them in work-related situations. WhatsApp was the most used app, in both personal and professional contexts. Respondents indicated using WhatsApp in 467 of 483 (92.6%) cases and Facebook in 209 (41.5%; Figure 2).

Of the 483 participants, 362 used their mobile phone to access SNSs in a work-related context (74.9%, 95% CI 71.1%-78.8%). This proportion was significantly higher in 3 situations: in the age span between 20 and 30 years (37/44, 84.1%, 95% CI 73.3%-94.9%); among professionals who used their mobile phone more than 3 hours daily (100/118, 84.7%, 95% CI 78.3%-91.2%); and among those with less than 15 years of work experience (142/175, 81.1%, 95% CI 75.3%-86.9%; Table 1).

Figure 1. Flow of participants through the study. EHR: electronic health record.

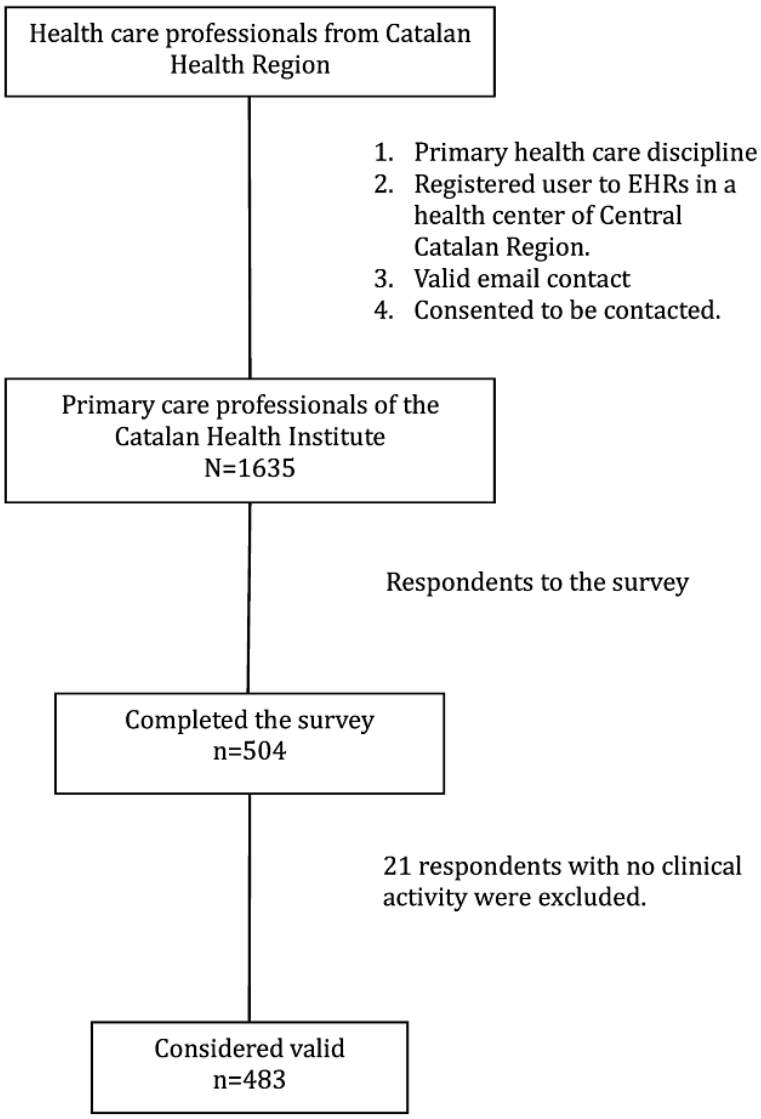


Figure 2. Percentage of respondents using the various apps for personal and professional use.

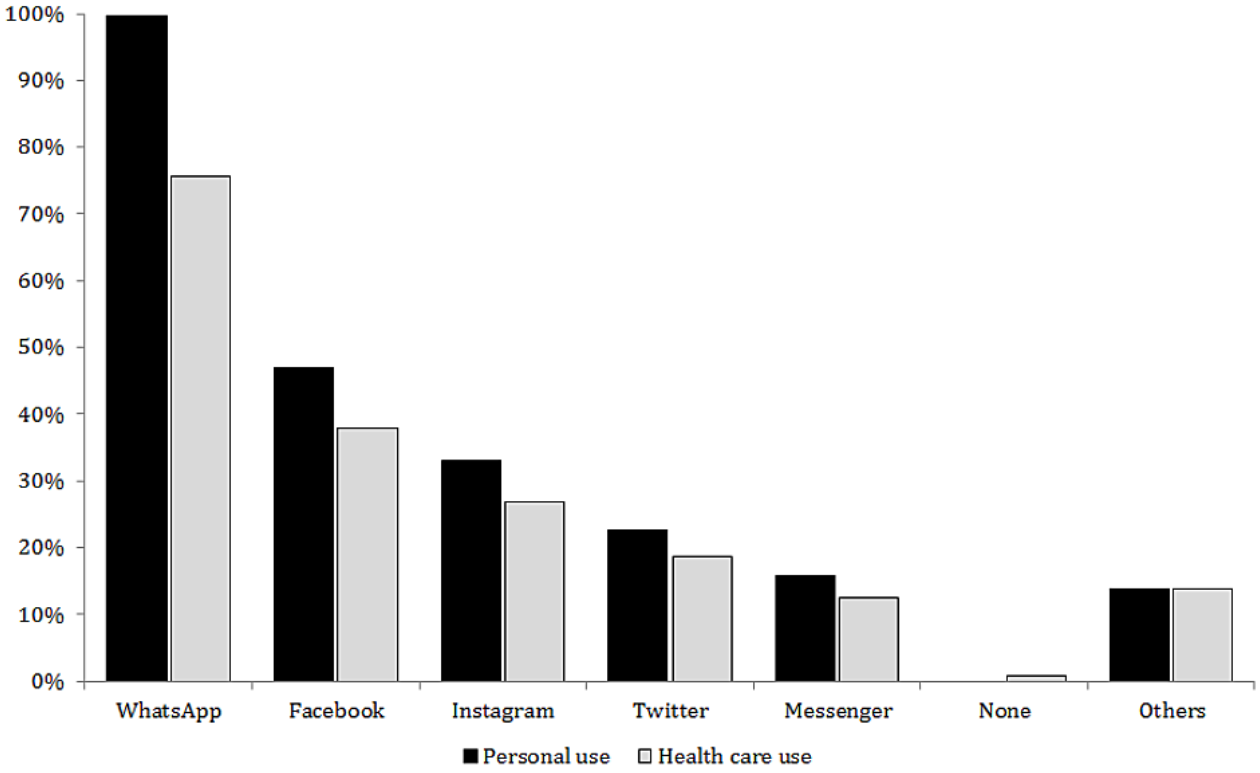


Table 1. Demographic data according to the use of social networking site apps.

Characteristics	Total (N=483)	Users (n=362)	Nonusers (n=121)	Crude odds ratio (95% CI)	*P* value
Age (years)[a], n (%)					
>50	161	110 (68.3)	51 (31.7)	1	
41-50	158	120 (75.9)	38 (24.1)	1.46 (0.9-2.4)	.13
31-40	120	95 (79.2)	25 (20.8)	1.76 (1-3.1)	.04
20-30	44	37 (84.1)	7 (15.9)	2.45 (1-5.9)	.04
Median (SD)		44 (10.4)	48 (10.1)		.005[b]
Sex, n (%)					
Male	90	64 (71.1)	26 (28.9)	1	
Female	393	298 (75.8)	95 (24.2)	1.3 (0.8-2.1)	.35
Health profession, n (%)					
Nurse	211	159 (75.4)	52 (24.6)	1	
Physician	215	162 (75.3)	53 (24.7)	1 (0.6-1.6)	.99
Midwife	31	24 (77.4)	7 (22.6)	1.1 (0.5-2.8)	.80
Social worker	21	14 (66.7)	7 (33.3)	0.5 (0.8-3)	.44
Dentist	5	3 (60)	2 (40)	0.7 (0.3-1.7)	.38
Role, n (%)					
Clinical	406	301 (74.1)	105 (25.9)	1	
Clinical and academic	77	61 (79.2)	16 (20.8)	1.3 (0.7-2.4)	.34
Mobile phone daily use (hours), n (%)					
<1	190	126 (66.3)	64 (33.7)	1	
1-3	175	136 (77.7)	39 (22.3)	1.8 (1.1-2.8)	.01
>3	118	100 (84.7)	18 (15.3)	2.8 (1.6-5)	.001
Work experience (years), n (%)[b]					
>35	46	29 (63)	17 (37)	1	
26-35	103	72 (69.9)	31 (30.1)	1.4 (0.7-2.8)	.40
16-25	159	119 (74.8)	40 (25.2)	1.7 (0.9-3.5)	.11
≤15 years	175	142 (81.1)	33 (18.9)	2.5 (1.2-5.1)	.01
Median (SD)	N/A[c]	18 (10.6)	22 (11.3)	N/A	.002[b]

[a]Pearson correlation coefficient (age and work experience) =.9; *P*<.001.

[b]*P* value for the linear trend test (analysis of variance).

[c]N/A: not applicable.

The factors independently associated with the use of apps to access SNSs in a professional-related context were having less than 15 years of work experience (adjusted OR 2.11, 95% CI 1.02-4.36) and a frequency of mobile phone use greater than 3 hours a day (adjusted OR 1.90, 95% CI 1.07-3.38; Table 2).

Most of the 362 respondents (mean 67.5%, SD 6.1%) considered using mobile phones to access SNSs in a professional context as useful or very useful in all 8 domains studied. Considering valuations rated as “useful” or “very useful” as indicators of usefulness, the best-rated domain was problem-solving skills (276/362, 76.2%, 95% CI 71.8%-80.6%), followed by knowledge about the profession (261/362, 72.1%, 95% CI 67.5%-76.7%) and speed and clinical safety (254/362, 70.2%, 95% CI 65.4%-74.9%; Table 3).

When we compared the apps according to the same domains, we observed that WhatsApp, Facebook, and Twitter were well valued for the acquisition of professional knowledge, creativity and innovation, and critical thinking skills. WhatsApp and Facebook were valued positively for their speed in helping to reach clinical decision, whereas WhatsApp was the only app positively valued for problem solving as well (177/219, 80.8%, 95% CI 75.6%-86.0%; Table 4).

We also asked the respondents to select their main reasons for using the apps. The reasons most frequently cited were communication between professionals and drug or clinical

consultations (Table 5). Among the reasons added by professionals, 8 of the 362 (2.2%) respondents reported using SNS apps to send photographs to other professionals and 5 of 362 (1.4%) reported using them to register clinical information.

These preferences varied according to the apps preferred by the health care professionals. However, it is notable that communication with other professionals was reported by 213 of 219 (97.3%, 95% CI 95.1%-99.4%) WhatsApp users (Table 4).

Table 2. Multivariate analysis of factors associated with work-related use of social networking site apps by primary care professionals.

Associated factors	Adjusted odds ratio (95% CI)	*P* value
Work experience (years)		
>35	1	N/A[a]
16-35	1.54 (0.79-2.99)	.19
≤15	2.11 (1.02-4.36)	.04
Daily use of mobile phone (hours)		
<3	1	N/A
≥3	1.90 (1.07-3.38)	.02

[a]N/A: not applicable.

Table 3. Assessment of the usefulness of social networking sites in the 8 domains analyzed (n=362).

Domains	Rating, n (%)			
	Not useful	Of little use	Useful	Very useful
Problem solving	18 (5.0)	68 (18.8)	202 (55.8)	74 (20.4)
Knowledge about profession	25 (6.9)	76 (21.0)	206 (56.9)	55 (15.2)
Speed and clinical safety	23 (6.4)	85 (23.5)	190 (52.5)	64 (17.7)
Patient care	25 (6.9)	85 (23.5)	199 (55.0)	53 (14.6)
Clinical decisions	20 (5.5)	91 (25.1)	203 (56.1)	48 (13.3)
Clinical skills	33 (9.1)	97 (26.8)	185 (51.1)	47 (13.0)
Creativity and innovation	30 (8.3)	112 (30.9)	171 (47.2)	49 (13.5)
Critical thinking	36 (9.9)	116 (32.0)	170 (47.0)	40 (11.0)

Table 4. Assessment of the impact of 4 apps compared according to the 8 domains analyzed.

Main uses of the apps	WhatsApp (n=219)	Facebook (n=22)	Twitter (n=20)	Instagram (n=8)
Domains, n (%)				
Problem solving	177 (80.8)[a]	19 (97.6)	15 (75.0)	6 (75.0)
Knowledge about profession	167 (76.3)[b]	21 (95.5)[a]	17 (85.0)[c]	8 (100)
Speed and clinical safety	164 (74.9)[c]	21 (95.5)[d]	19 (95.0)[a]	6 (75.0)
Patient care	155 (70.8)	19 (86.4)	18 (90.0)[e]	6 (75.0)
Clinical decisions	166 (75.8)[f]	21 (95.5)[g]	17 (85.0)	8 (100)
Clinical skills	146 (66.7)	20 (90.9)[d]	17 (85.0)	8 (100)
Creativity and innovation	147 (67.1)[c]	20 (90.9)[h]	18 (90.0)[g]	6 (75.0)
Critical thinking	140 (63.9)[i]	21 (95.5)[j]	17 (85.0)	6 (75.0)
Utility, n (%)				
Communication with other professionals	213 (97.3)[j]	21 (95.5)	18 (90.0)	8 (100)
Pharmacological or clinical consultations	124 (56.6)	14 (63.6)	16 (80.0)[k]	6 (75.0)
Professional development	72 (32.9)	9 (40.9)	11 (55.0)[l]	4 (50.0)
Health promotion	60 (27.4)	13 (59.1)[j]	14 (70.0)[j]	2 (25.0)
Communication with patients	50 (22.8)	5 (22.7)	3 (15.0)	1 (12.5)
Social networks	50 (22.8)	13 (59.1)[j]	12 (47.5)[j]	4 (50.0)
Work or research opportunities	44 (20.1)[i]	7 (31.8)	5 (25.0)	3 (37.5)
Other	9 (4.1)	0 (0.0)	0 (0.0)	0 (0.0)

[a]*P*=.01.
[b]*P*=.03.
[c]*P*=.02.
[d]*P*=.04.
[e]*P*=.006.
[f]*P*=.003.
[g]*P*=.007.
[h]*P*=.001.
[i]*P*=.002.
[j]*P*=.005.
[k]*P*<.001.
[l]*P*=.009.

Table 5. Reasons given by the professionals (n=362) for using social networking site apps.

Reasons for using the apps[a]	n (%)	95% CI
Communication with other professionals	338 (93.4)	90.8-95.9
Pharmacological or clinical consultations	204 (56.4)	51.2-61.5
Professional development	106 (29.3)	24.6-34.0
Health promotion	86 (23.8)	19.4-28.2
Communication with patients	72 (19.9)	15.8-24.0
Social networks	71 (19.6)	15.5-23.7
Work or research opportunities	57 (15.7)	12.0-19.5
Other	19 (5.2)	2.9-7.6
Sending images or clinical photos	8 (2.2)	0.7-3.7
Clinical information record	5 (1.4)	0.2-2.6
Assistance support tools	3 (0.8)	0-1.8
Professional email	3 (0.8)	0-1.8

[a]Respondents could choose more than 1 reason.

Discussion

Principal Findings

The results of this study indicate that most of the primary health care professionals surveyed were using apps to access SNSs in a professional context and that WhatsApp, Twitter, and Facebook, in this order, were the most used, in both personal and professional contexts [20,29]. In terms of its benefits, WhatsApp was generally perceived as more useful for improving professional knowledge and clinical problem solving [13]. These findings suggest that these apps can be powerful tools to involve health professionals in their professional activities and that they can be used as a model to develop new and more secure apps in the future [21].

The study showed a higher proportion of SNS users among professionals with shorter work experience and, although the univariate analysis didn't achieve statistical significance, a multivariate analysis demonstrated that age and work experience were significantly correlated variables (linear correlation) and, together with hours of mobile phone use, generated a good response model. New generations of professionals, as expected, made greater use of mobile phones and everything that use entails (eg, participating in social networks or conducting internet searches). The health system should be adapted to this, both ethically (for the sharing of photos and patient data) and in relation to documentary and assisted support. If we were to repeat our study in 15 years' time, it would show a completely different picture.

Professionals perceived that using these apps had an impact in several domains, the most prominent of these being the apps' role in improving knowledge and problem solving, as well as their speed and clinical security. When we inquired about applied uses, respondents emphasized the use of apps as a communication tool and, although the amount of data we obtained did not allow for deep analysis, a significant number of professionals claimed to have sent patient images or photographs to other colleagues and a small percentage had sent clinical information. Some studies carried out with mobile phones mentioned that telemedicine offers an opportunity to send photos and video clips, representing a source of clinical support for obtaining a second opinion from other colleagues and experts [30,31]. In an environment of scarce resources, the use of mobile phones for medical communication could be of great value. However, we should not forget that sending health information through apps, such as WhatsApp, can imply a serious risk to patient data safety. Professionals are using SNS tools such as WhatsApp and Facebook commonly to communicate and share clinical information, and this use of social media as a health tool raises ethical issues in part because of the possible inappropriate use of individuals' personal and sensitive information and the possible breach of data security regulations (such as the European Union's General Data Protection Regulation). Health institutions must give special attention to advising health professionals about these risks.

The use of SNSs as a means of communicating with patients has been reported as being of little use, probably, according to other studies, due to the lack of legal protection, because their use could be a source of errors or distractions [32], or because of the preference for face-to-face contact with their physicians by a large part of the population [33]. This trend could change in the near future, as pointed out by some studies carried out in places where mobile phones are mostly used, since it can improve patient care and make the use of resources more efficient [29-31,34].

Limitations

The study had several limitations. A selection bias was caused by the type of convenience sample used (closed cohort). This problem could be solved in future research by expanding the recruitment to self-selected professionals on the internet. Another limitation originated in the low response rate and the bias inherent in using a Web-based survey that those with better technology would be likelier to respond and, therefore, more likely to use apps for professional purposes. There was another

important bias in relation to the high percentage of physicians and nurses who responded to the questionnaire, caused in part by the higher number of those professional categories registered as electronic health record users and in part by the low participation rate of other clinical categories included in the study.

Because this was a descriptive study, we were not able to establish a cause-effect relationship.

Comparison With Prior Work

Our findings are in accordance with those observed in other studies [6,11]. The most popular social and messaging platforms used by health professionals were the same, and they had similar usage patterns in their professional context. The limited use of other more specialized groups of health apps in our study differed from the findings of other studies conducted in populations that used these apps constantly, especially those that are used for direct patient management (eg, medical guidelines and medical calculators) [29,34], which could be explained by poor knowledge of the current market or by technological barriers, especially among certain age segments of users. Although there are more specialized health apps that offer similar communication features and tend to have better safety profiles and certification in the handling of data [21], the lack of information and poor knowledge about them could be preventing their use. This leaves open the possibility that promotion and dissemination of such tools in professionals' environments could improve their use.

Social networking is a form of social media, and SNS users typically download services that offer social media functionality to their mobile devices (eg, mobile phones and tablet computers), but they can also access SNSs on desktop computers or laptops [35]. Studies to determine which devices health professionals use to access social media are lacking. Some sources suggest that the rate of using mobile phones or mobile devices to using computers and laptops for accessing social media is 2 to 1 [36]. Our study specifically focused on accessing SNSs through mobile device apps, assuming that these devices are used most frequently, but more studies need to be done on this particular subject.

Some studies found that physicians use predetermined browsers in their mobile phones to access SNSs, search clinical practice guidelines or patient information, or access medical information through the Web [37,38]. Our survey may not have caught this functionality, carried out with mobile phones. Other studies reported the use of mobile technology in primary care as a good tool to provide medical care in hard-to-reach areas, making it easier to guarantee health services and resources [11,18,19]. The combination of SNSs and mobile health offers a great opportunity to strengthen information systems transforming health systems. However, the implementation of this combination should carefully consider aspects such as the security, privacy, and confidentiality of user information, but it also needs to take into consideration health professionals' preferences [20,22,23]. The results of this study provide new insights into the use and perceived benefits of apps among primary care professionals and, specifically, about the uses and needs relating to social networks. The demonstration of health professionals' use of SNSs should warn us about the need to improve and enhance their benefits, but also to facilitate the proper and secure use of these new tools. Further analytical-experimental research using more exhaustive methods to recruit participants will be essential to confirm and extend the results of this study.

Conclusions

The vast majority of primary health care professionals surveyed, 362 of 483 (74.9%) respondents, accessed SNSs with their mobile phones in a work-related context. WhatsApp was the most used, in both personal and professional contexts. Mobile phone apps with access to SNSs in health care are frequently used for communication between professionals, but they are also used for the exchange of files and images or recorded clinical data. The use of these apps, according to the professionals surveyed, affects problem solving, but their use for communicating with patients is not yet widespread. We recommend that health institutions assess the need to improve the general and specific knowledge about the available apps and, thereby, improve and facilitate their use among health professionals as a way to prevent the risks of inappropriate use.

Acknowledgments

We appreciate the collaboration provided in the development of the study by the components of the Research Unit of Central Catalonia and the primary care professionals who responded to the survey.

Conflicts of Interest

None declared.

References

1. Liu Y, Ren W, Qiu Y, Liu J, Yin P, Ren J. The use of mobile phone and medical apps among general practitioners in Hangzhou City, Eastern China. JMIR Mhealth Uhealth 2016 May 24;4(2):e64 [FREE Full text] [doi: 10.2196/mhealth.4508] [Medline: 27220417]
2. Barreto JE, Whitehair CL. Social media and web presence for patients and professionals: evolving trends and implications for practice. PM R 2017 May;9(5S):S98-S105. [doi: 10.1016/j.pmrj.2017.02.012] [Medline: 28527508]
3. Statista. Annual number of global mobile app downloads 2016-2021. URL: https://www.statista.com/statistics/271644/worldwide-free-and-paid-mobile-app-store-downloads/ [accessed 2018-04-19] [WebCite Cache ID 6ynHSFx8g]

4. Statista. Number of social network users worldwide from 2010 to 2020. URL: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ [accessed 2018-04-19] [WebCite Cache ID 6ynHrE6MR]
5. Estudio Redes Sociales. Madrid, Spain: IAB Spain; 2016 Apr 20. URL: http://www.iabspain.net/wp-content/uploads/downloads/2016/04/IAB_EstudioRedesSociales_2016_VCorta.pdf [accessed 2018-04-19] [WebCite Cache ID 6ynILrbHy]
6. McCallum C, Rooksby J, Gray CM. Evaluating the impact of physical activity apps and wearables: interdisciplinary review. JMIR Mhealth Uhealth 2018 Mar 23;6(3):e58 [FREE Full text] [doi: 10.2196/mhealth.9054] [Medline: 29572200]
7. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69 [FREE Full text] [doi: 10.2196/mhealth.7141] [Medline: 28536089]
8. Lalloo C, Shah U, Birnie KA, Davies-Chalmers C, Rivera J, Stinson J, et al. Commercially available smartphone apps to support postoperative pain self-management: scoping review. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e162 [FREE Full text] [doi: 10.2196/mhealth.8230] [Medline: 29061558]
9. Zhang M, Ying J, Song G, Fung DS, Smith H. Attention and cognitive bias modification apps: review of the literature and of commercially available apps. JMIR Mhealth Uhealth 2018 May 24;6(5):e10034 [FREE Full text] [doi: 10.2196/10034] [Medline: 29793899]
10. Dorsey ER, Yvonne CY, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
11. Chan WS, Leung AY. Use of social network sites for communication among health professionals: systematic review. J Med Internet Res 2018 Mar 28;20(3):e117. [doi: 10.2196/jmir.8382] [Medline: 29592845]
12. Roehrs A, da Costa CA, Righi RDR, de Oliveira KSF. Personal health records: a systematic literature review. J Med Internet Res 2017 Jan 06;19(1):e13 [FREE Full text] [doi: 10.2196/jmir.5876] [Medline: 28062391]
13. Ganasegeran K, Renganathan P, Rashid A, Al-Dubai SAR. The m-Health revolution: exploring perceived benefits of WhatsApp use in clinical practice. Int J Med Inform 2017 Dec;97:145-151. [doi: 10.1016/j.ijmedinf.2016.10.013] [Medline: 27919374]
14. Brown J, Ryan C, Harris A. How doctors view and use social media: a national survey. J Med Internet Res 2014;16(12):e267 [FREE Full text] [doi: 10.2196/jmir.3589] [Medline: 25470407]
15. Anguiano B, Brown-Johnson C, Rosas LG, Pechmann C, Prochaska JJ. Latino adults' perspectives on treating tobacco use via social media. JMIR Mhealth Uhealth 2017 Feb 08;5(2):e12 [FREE Full text] [doi: 10.2196/mhealth.6684] [Medline: 28179217]
16. Eckert S, Sopory P, Day A, Wilkins L, Padgett D, Novak J, et al. Health-related disaster communication and social media: mixed-method systematic review. Health Commun 2017 Aug 21:1-12. [doi: 10.1080/10410236.2017.1351278] [Medline: 28825501]
17. Holt-Lunstad J, Robles TF, Sbarra DA. Advancing social connection as a public health priority in the United States. Am Psychol 2017 Sep;72(6):517-530. [doi: 10.1037/amp0000103] [Medline: 28880099]
18. Brangel P, Sobarzo A, Parolo C, Miller BS, Howes PD, Gelkop S, et al. A serological point-of-care test for the detection of IgG antibodies against Ebola virus in human survivors. ACS Nano 2018 Jan 23;12(1):63-73 [FREE Full text] [doi: 10.1021/acsnano.7b07021] [Medline: 29303554]
19. Nic Lochlainn LM, Gayton I, Theocharopoulos G, Edwards R, Danis K, Kremer R, et al. Improving mapping for Ebola response through mobilising a local community with self-owned smartphones: Tonkolili District, Sierra Leone, January 2015. PLoS One 2018;13(1):e0189959 [FREE Full text] [doi: 10.1371/journal.pone.0189959] [Medline: 29298314]
20. Moore S, Jayewardene D. The use of smartphones in clinical practice. Nurs Manag (Harrow) 2014 Jul;21(4):18-22. [doi: 10.7748/nm.21.4.18.e1225] [Medline: 24967805]
21. Marin-Gomez FX, Vidal-Alaball J, Garcia Cuyàs F, Reig-Bolano R. Attending home care patients in primary care using a smartphone application (WhatsICS): a feasibility study. Int Arch Med 2017 Oct 19;10(261):1-8. [doi: 10.3823/2531]
22. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. Health Commun 2018 Sep;33(9):1131-1140. [doi: 10.1080/10410236.2017.1331312] [Medline: 28622038]
23. Al Khaja KAJ, AlKhaja AK, Sequeira RP. Drug information, misinformation, and disinformation on social media: a content analysis study. J Public Health Policy 2018 Aug;39(3):343-357. [doi: 10.1057/s41271-018-0131-2] [Medline: 29795521]
24. Verenna AA, Noble KA, Pearson HE, Miller SM. Role of comprehension on performance at higher levels of Bloom's taxonomy: findings from assessments of healthcare professional students. Anat Sci Educ 2018 Jan 18. [doi: 10.1002/ase.1768] [Medline: 29346708]
25. Neal AE, Lehto E, Miller KH, Davis E, Ziegler C. A qualitative assessment of pediatric cardiology core content: comments from Kentucky trainees, pediatricians, and pediatric cardiologists. Congenit Heart Dis 2018 Jul 22. [doi: 10.1111/chd.12626] [Medline: 30033580]
26. Mashamba-Thompson TP, Sartorius B, Stevens FCJ, Drain PK. Experiential Bloom's Taxonomy learning framework for point-of-care diagnostics training of primary healthcare workers. Afr J Lab Med 2016;5(1):449 [FREE Full text] [doi: 10.4102/ajlm.v5i1.449] [Medline: 28879117]
27. Coma Redon E, Méndez Boo L. Experiencias para compartir SISAP: 4 años buceando en mares de datos. Actual Med Fam 2010;6(8):473-476 [FREE Full text]

28. Alsobayel H. Use of Social Media for Professional Development by Health Care Professionals: A Cross-Sectional Web-Based Survey. JMIR Med Educ 2016 Sep 12;2(2):e15 [FREE Full text] [doi: 10.2196/mededu.6232] [Medline: 27731855]
29. Jebraeily M, Fazlollahi ZZ, Rahimi B. The most common smartphone applications used by medical students and barriers of using them. Acta Inform Med 2017 Dec;25(4):232-235 [FREE Full text] [doi: 10.5455/aim.2017.25.232-235] [Medline: 29284911]
30. Chao C, Chen Y, Shih C, Hou S, Seethala RR, Aisiku IP, et al. Smartphone transmission of electrocardiography images to reduce time of cardiac catheterization laboratory activation. J Chin Med Assoc 2018 Jun;81(6):505-510 [FREE Full text] [doi: 10.1016/j.jcma.2017.11.009] [Medline: 29287703]
31. Fernandes-Taylor S, Gunter RL, Bennett KM, Awoyinka L, Rahman S, Greenberg CC, et al. Feasibility of implementing a patient-centered postoperative wound monitoring program using smartphone images: a pilot protocol. JMIR Res Protoc 2017 Feb 22;6(2):e26 [FREE Full text] [doi: 10.2196/resprot.6819] [Medline: 28228369]
32. Nerminathan A, Harrison A, Phelps M, Alexander S, Scott KM. Doctors' use of mobile devices in the clinical setting: a mixed methods study. Intern Med J 2017 Dec;47(3):291-298. [doi: 10.1111/imj.13349] [Medline: 27925381]
33. 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]
34. Perlman EM, Greenberg PB, Browning D, Friday RP, Miller JW. Solving the hydroxychloroquine dosing dilemma with a smartphone app. JAMA Ophthalmol 2018 Feb 01;136(2):218-219. [doi: 10.1001/jamaophthalmol.2017.6076] [Medline: 29302689]
35. Heinrichs JJ, Lim JS, Lim KS. Influence of social networking site and user access method on social media evaluation. J Consum Behav 2011 Dec 23;10(6):347-355. [doi: 10.1002/cb.377]
36. Bemis M. How do people access social media and why? Crimson Hexagon analyzes social user device preferences. Boston, MA: Crimson Hexagon; 2014 Sep 15. URL: https://www.crimsonhexagon.com/blog/people-access-social-media/ [accessed 2018-08-09] [WebCite Cache ID 71YAQ8sDL]
37. Manganello J, Gerstner G, Pergolino K, Graham Y, Falisi A, Strogatz D. The relationship of health literacy with use of digital technology for health information: implications for public health practice. J Public Health Manag Pract 2017;23(4):380-387. [doi: 10.1097/PHH.0000000000000366] [Medline: 26672402]
38. Kim R, Kim H, Jeon B. The good, the bad, and the ugly of medical information on the internet. Mov Disord 2018 May;33(5):754-757. [doi: 10.1002/mds.27324] [Medline: 29436748]

Abbreviations

OR: odds ratio
SNS: social networking site

Edited by G Eysenbach; submitted 27.05.18; peer-reviewed by M Merolli, J Brown; comments to author 19.07.18; revised version received 13.08.18; accepted 25.08.18; published 21.12.18.

Please cite as:
Marin-Gomez FX, Garcia Cuyas F, Reig-Bolano R, Mendioroz J, Roura-Poch P, Pico-Nicolau M, Vidal-Alaball J
Social Networking App Use Among Primary Health Care Professionals: Web-Based Cross-Sectional Survey
JMIR Mhealth Uhealth 2018;6(12):e11147
URL: https://mhealth.jmir.org/2018/12/e11147/
doi:10.2196/11147
PMID:30578175

Original Paper

Uptake and Utilization of the Management of Anticoagulation in the Periprocedural Period App: Longitudinal Analysis

Alex C Spyropoulos[1*], MD; Anne Myrka[2*], MAT, RPH; Darren M Triller[3*], PharmD; Stephen Ragan[2*], MD; Collin York[2*], BS; Jaz-Michael King[2*], MS; Ti-Kuang Lee[2*], ScM

[1]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, New York, NY, United States

[2]IPRO, Albany, NY, United States

[3]WellScriptEd Consulting Inc, Delmar, NY, United States

[*]all authors contributed equally

Corresponding Author:
Anne Myrka, MAT, RPH
IPRO
20 Corporate Woods Boulevard
Albany, NY, 12211
United States
Phone: 1 518 320 3591
Email: Anne.Myrka@area-I.hcqis.org

Abstract

Background: Anticoagulants are major contributors to preventable adverse drug events, and their optimal management in the periprocedural period is particularly challenging. Traditional methods of disseminating clinical guidelines and tools cannot keep pace with the rapid expansion of available therapeutic agents, approved indications for use, and published medical evidence, so a mobile app, Management of Anticoagulation in the Periprocedural Period (MAPPP), was developed and disseminated to provide clinicians with guidance that reflects the most current medical evidence.

Objective: The objective of this study was to assess the global, national, and state-level acquisition of a mobile app since its initial release and characterize individual episodes of use based on drug selection, procedural bleeding risk, and patient thromboembolic risk.

Methods: Data were extracted from a mobile app usage tracker (Google Analytics) to characterize new users and completed episodes temporally (by calendar quarter) and geographically (globally, nationally, and in the targeted US state of New York) for the period between April 1, 2016 and September 30, 2017.

Results: The app was acquired by 2866 new users in the measurement period, and the users completed nearly 10,000 individual episodes of use. Acquisition and utilization spanned 51 countries globally, predominantly in the United States and particularly in New York State. Warfarin and rivaroxaban were the most frequently selected drugs, and completed episodes most frequently included the selection of high bleeding risk (4888/9963, 49.06%) and high thromboembolic risk categories (4500/9963, 45.17%).

Conclusions: The MAPPP app is a successful means of disseminating current guidance on periprocedural anticoagulant use, as indicated by broad global uptake and upward trends in utilization. Limitations in access to provider and patient-specific data preclude objective evaluation of the clinical impact of the app. An ongoing study incorporating app logic into electronic health record systems at participant health systems will provide a more definitive evaluation of the clinical impact of the app logic.

(JMIR Mhealth Uhealth 2018;6(12):e11090) doi:10.2196/11090

KEYWORDS

adverse drug event; anticoagulant; app; mobile phone; periprocedural; warfarin

Introduction

Anticoagulants are highly prescribed medications that have been identified as major contributors to adverse drug events (ADEs), of which bleeding is the most common and dangerous. Anticoagulants have repeatedly been identified among the drugs most frequently associated with emergency department (ED) visits, and national estimates indicate that the rates of such events are increasing [1-4]. While warfarin had been the only available oral anticoagulant for decades, 5 new direct oral

anticoagulants have come to market in the United States since 2010, paralleled by an increase in outpatient anticoagulant use [5]. The proportion of ED visits due to anticoagulants has increased by 57% in that period, and anticoagulants now account for nearly 18% of all ADE-associated ED visits in the United States [1-4,6]. Because many serious ADEs are thought to be preventable through improvements in care delivery, anticoagulants are identified as a specific target for health system improvement in the US Department of Health and Human Services' National Action Plan for Adverse Drug Event Prevention [7-16]. Quality Innovation Network-Quality Improvement Organizations (QIN-QIOs) have also been contracted by the Centers for Medicare & Medicaid Services (CMS) to work with providers, Medicare beneficiaries, and other stakeholders to improve the quality of care for patients who are prescribed anticoagulants [17].

For patients utilizing chronic oral anticoagulation, medical management in the time period leading up to and following an invasive medical procedure or surgery (ie, the periprocedural period) is particularly challenging. Such patients are exposed to an increased risk of serious or life-threatening bleeding when clinically relevant anticoagulation is present in the absence of complete hemostasis, and conversely, they are at increased risk of thrombosis when the degree of anticoagulation present is insufficient to address underlying thromboembolic risk (ie, when an anticoagulant is held for surgery). Under such circumstances, clinicians must balance the risk of bleeding associated with the procedure with that of the underlying thromboembolic risk of condition(s) that has prompted the need for anticoagulation. With approximately 10%-15% of chronically anticoagulated patients requiring anticoagulant interruption for an invasive procedure each year, experts have endeavored to create guidance for optimal periprocedural management, addressing surgical risk, thromboembolic risk, and the pharmacokinetic profiles of the individual anticoagulants [18-23]. However, the rapid expansion of available agents, approved indications for use, and related medical evidence serve to minimize the utility of traditional published guidelines, which are revised infrequently and are not reflective of the most current practices [24]. Similarly, it can be difficult to recall or update clinical tools disseminated in hard copy or pdf formats, and neither supports passive data collection on utilization or impact on patient care.

To increase clinician access to the most current expert guidance on periprocedural anticoagulation management, the New York State Anticoagulation Coalition (NYSACC) led the development and dissemination of the Management of Anticoagulation in the Periprocedural Period mobile app (MAPPP; see Multimedia Appendix 1) [25]. This paper describes the creation of the app and its adoption and utilization globally, in the United States, and in the state of New York. It also characterizes episodes of use by drug, procedural bleeding risk, and patient thromboembolic risk.

Methods

Clinical Tool Development

The NYSACC was created in 2012 by IPRO, the CMS-designated QIO for New York State, to advance the drug safety priorities identified in the CMS QIN-QIO 10th Statement of Work [26]. The NYSACC's multidisciplinary membership consists of more than 150 representatives from clinical practice, academia, industry, and advocacy organizations with interest in anticoagulation management quality. The NYSACC identified inadequate availability of tools for periprocedural management as a barrier to quality anticoagulant management that warranted action.

In the spring of 2013, the group developed and disseminated a novel clinical tool, the MAPPP Tool. This tool, disseminated in hard copy and pdf formats, utilized a 3×3 matrix to help clinicians simultaneously categorize underlying thromboembolic risk (high, moderate, or low) and procedural bleeding risk (high, low, or minimal) and provided evidence-based guidance on (1) whether the anticoagulant should be interrupted; (2) timing for preprocedural discontinuation of anticoagulant (if necessary); (3) whether "bridging" with heparin products is warranted, with details of timing and laboratory monitoring; and (4) timing and dosing of re-initiation of anticoagulant in the postprocedural period. The tool included guidance for all oral anticoagulants available for use in the United States at the time of release, including apixaban, dabigatran, edoxaban, rivaroxaban, and warfarin.

App Development

Recognizing the barriers to disseminating and updating the MAPPP in its original form, and appreciating the limited number of mutually exclusive steps involved in navigating the tool, it was determined that the structure of the MAPPP lent itself to the creation of a mobile app that would potentially allow for broader dissemination, remote updates to reflect most current knowledge, and the collection of data on downloads and utilization. The MAPPP app design team was convened, which included clinical content experts and IPRO app developers, with the goal of creating an app that would provide a high-value proposition for the clinical user (evidence-based, accurate, and quick results) with simplicity of design (few pages). Construction of a wireframe and subsequent user interface sample screenshots using MAPPP app logic were designed to iteratively develop the app through use cases, preliminary usability testing, and clinical content expert feedback. The resulting prototype was refined into the final product via iterative cycles of clinical user testing.

The MAPPP app was developed using the open source framework Cordova (Apache Software Foundation, Wakefield, MA), which utilizes HTML, CSS, and JS to develop apps across multiple platforms from a single code base. The MAPPP app was made available through both the iOS App Store and Google Play; it functions with Web browsers that support modern Web technologies. The app was released in April 2016 and, based on user feedback, underwent minor modifications in May 2016 to enhance the disclaimers, improve the visibility of the button linking to more information on selection options, and expand

flexibility of use through the addition of a "Back" button. While no major changes have been made to the clinical content of the app to date, a process is in place to support app updates as necessary based on the most current evidence-based recommendations in the field of perioperative antithrombotic therapy.

Analysis of App Uptake and Utilization

Data were extracted from Google Analytics (including iOS, Android, and Web browser data) to characterize new users and completed episodes temporally (by calendar quarter) and geographically (globally, United States, and the targeted US state of New York) for the period between April 1, 2016 and September 30, 2017. A completed episode was defined by the project team as a user interaction that resulted in the app presenting a recommendation page based on sequential selection of an anticoagulant drug, assignment of bleeding risk category for a procedure, and indication of underlying thromboembolic risk of the patient [21]. Additional analysis of completed episodes was performed to characterize utilization by individual drug, procedural bleeding risk category, and underlying thromboembolic category.

Results

There were 353 new users of the app globally in its first full calendar quarter of use, primarily in the United States (276/353, 78.18%; Figure 1) and in the project's home state of New York (168/353, 47.59%). App acquisition increased through Q3 2017, accruing 2866 total global new users, dominated by use in the United States (2013/2866, 70.24%), particularly in New York (1067/2866, 37.23%). The app was downloaded and used for at least one episode in 51 different countries during the measurement period. Among US states, California (85) and New Mexico (69) had the second and third greatest numbers of new users. The quarterly total of all users (ie, the sum of new and returning users) trended upwards in like manner (Figure 2).

Users completed nearly 10,000 episodes during the study period, exceeding 2000 episodes in each of the most recent calendar quarters (Table 1). Overall, utilization was highest in the United States (6748/9963, 67.73%), particularly in New York (3618/9963, 36.31%). Globally, episodes were spread approximately evenly between new users (4571/9963, 45.88%) and returning users (5392/9963, 54.12%).

Among the completed episodes, the most commonly selected medications were warfarin (4074/9963, 40.89%), followed by rivaroxaban (2347/9963, 23.56%; Table 2). Each completed episode required the selection of a drug, a procedural bleeding risk category, and an underlying thromboembolic risk category, thus, allowing for 45 possible drug-risk-risk combinations. Of these possible combinations, those indicating high bleeding risk (4888/9963, 49.06%, episodes) and high thromboembolic risk (4500/9963, 45.17%, episodes) were the most commonly selected. Warfarin at 13.29% (1324/9963) and rivaroxaban at 6.52% (650/9963) were the agents most frequently associated with high bleeding risk–high thromboembolic risk episodes.

More in-depth analysis of trends among drug selections and risk categorizations of completed episodes would require evaluation of drug utilization trends among individual nations, facility-level data (eg, procedures performed, patients characteristics), user profiles (eg, medical residents vs experienced clinicians), and other variables that were not available for the current analysis, so no such analysis was attempted.

Figure 1. New users of the Management of Anticoagulation in the Periprocedural Period App by calendar quarter (Q) as defined by Google Analytics.

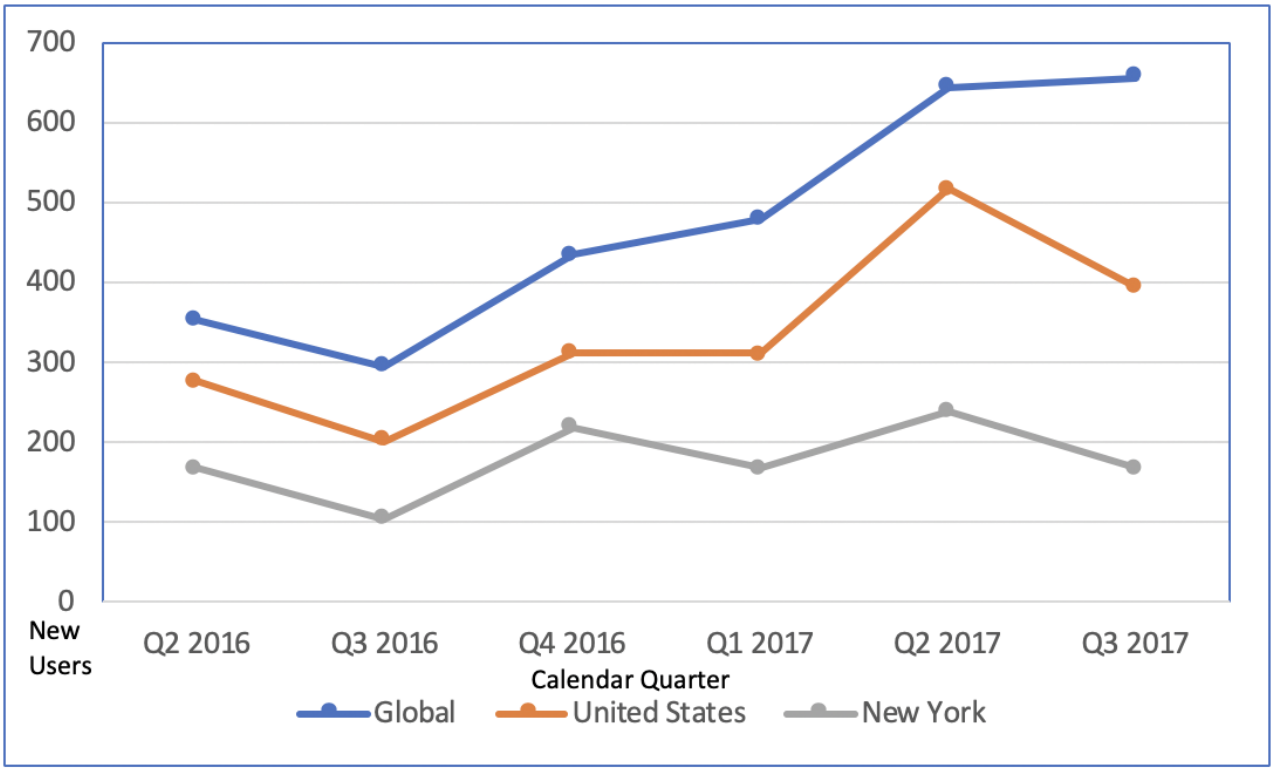


Figure 2. All users of the Management of Anticoagulation in the Periprocedural Period App by calendar quarter (Q) as defined by Google Analytics.

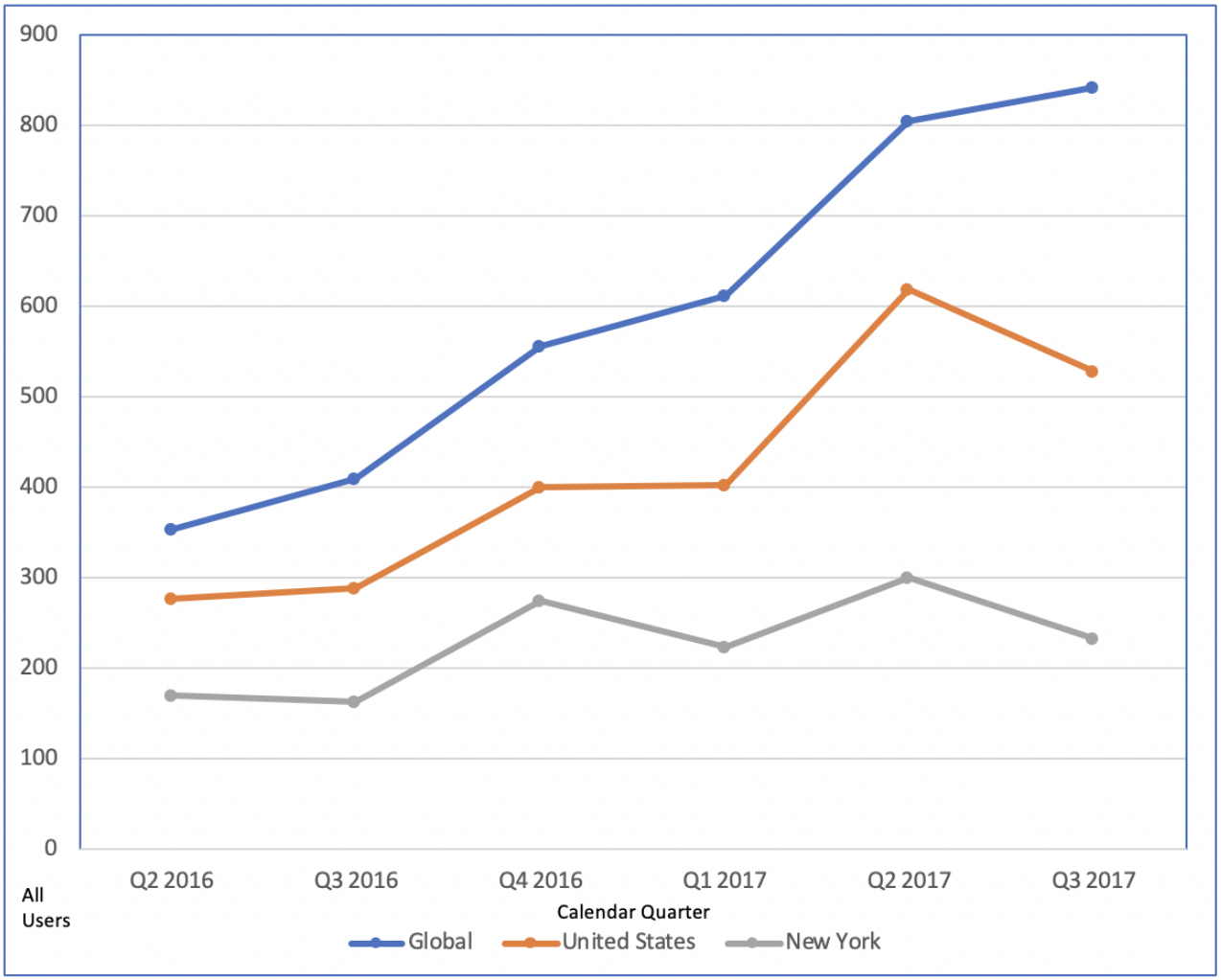


Table 1. Completed episodes by calendar quarter.

Geography	Quarter 2 2016 (n=1083), n (%)	Quarter 3 2016 (n=1064), n (%)	Quarter 4 2016 (n=1718), n (%)	Quarter 1 2017 (n=1766), n (%)	Quarter 2 2017 (n=2071), n (%)	Quarter 3 2017 (n=2261), n (%)	Total to date (N=9963), n (%)
Global							
New users	567 (52.35)	426 (40.04)	743 (43.25)	813 (46.04)	937 (45.24)	1085 (47.99)	4571 (45.88)
Returning users	516 (47.65)	638 (59.96)	975 (56.75)	953 (53.96)	1134 (54.76)	1176 (52.01)	5392 (54.12)
United States							
Total users	779 (71.93)	760 (71.43)	1169 (68.04)	1018 (57.64)	1588 (76.68)	1434 (63.42)	6748 (67.73)
New users	432 (39.89)	279 (26.22)	466 (27.12)	476 (26.95)	721 (34.81)	612 (27.07)	2986 (29.97)
Returning users	347 (32.04)	481 (45.21)	703 (40.92)	542 (30.69)	867 (41.86)	822 (36.36)	3762 (37.76)
New York							
Total users	525 (48.48)	383 (36)	759 (44.18)	543 (30.75)	712 (34.38)	696 (30.78)	3618 (36.31)
New users	255 (23.55)	140 (13.16)	297 (17.29)	261 (14.78)	302 (14.58)	262 (11.59)	1517 (15.23)
Returning users	270 (24.93)	243 (22.84)	462 (26.89)	282 (15.97)	410 (19.8)	434 (19.2)	2101 (21.09)

Table 2. Details of completed episodes of use (N=9963).

Bleeding risk and thromboembolic risk	Apixaban, n (%)	Dabigatran, n (%)	Edoxaban, n (%)	Rivaroxaban, n (%)	Warfarin, n (%)	Row values, n (%)
High						
High	526 (5.28)	368 (3.69)	132 (1.32)	650 (6.52)	1324 (13.29)	3000 (30.11)
Moderate	215 (2.16)	169 (1.7)	35 (0.35)	319 (3.2)	560 (5.62)	1298 (13.03)
Low	118 (1.18)	72 (0.72)	25 (0.25)	137 (1.38)	238 (2.39)	590 (5.92)
Low						
High	189 (1.9)	99 (0.99)	29 (0.29)	256 (2.57)	430 (4.32)	1003 (10.07)
Moderate	327 (3.28)	234 (2.35)	52 (0.52)	435 (4.37)	615 (6.17)	1663 (16.69)
Low	200 (2.01)	84 (0.84)	21 (0.21)	151 (1.52)	343 (3.44)	799 (8.02)
Minimal						
High	100 (1)	55 (0.55)	30 (0.3)	124 (1.24)	188 (1.89)	497 (4.99)
Moderate	80 (0.8)	69 (0.69)	21 (0.21)	115 (1.15)	174 (1.75)	459 (4.61)
Low	148 (1.49)	91 (0.91)	53 (0.53)	160 (1.61)	202 (2.03)	654 (6.56)
Total	1903 (19.10)	1241 (12.46)	398 (3.99)	2347 (23.56)	4074 (40.89)	9963 (100)

Discussion

Principal Findings

The MAPPP app was developed and disseminated to address a pressing public health need (ie, the rate of ADEs associated with anticoagulant use). The technology was perceived to have specific advantages over traditional distribution of clinical tools as hard copies or pdf files, including ease of access when providing direct patient care and the ability to reach an expanded recipient audience, update content, and passively collect data. This analysis of the initial uptake and utilization of the MAPPP app supports the premise that such apps can not only be a useful means of disseminating clinical tools but also help identify relevant weaknesses and opportunities for improvement.

The app demonstrated the ability to reach a broad audience of recipients within not only the state of New York (the target geography of the CMS-funded quality improvement project) but also across the United States and in other countries. Despite the absence of any formal dissemination plan beyond an initial launch press release on April 29, 2016, direct provider interactions in New York, and sharing by members of the NYACC, the app has continued to generate new users and a growing number of completed user episodes. Evaluation of utilization data also suggests that the app is frequently used to access guidance for patients undergoing procedures with the highest bleeding and thromboembolic risk. While no changes were made to the clinical content of the app during the period of this evaluation, several formatting changes were made successfully, demonstrating the ability to update information to providers without having to remove outdated hard copies or pdf files. Finally, the analysis was performed using data passively collected by Google Analytics, which is impossible with hard copy tools or pdf files without advanced formatting functionality.

However, the app and the current analysis are not without limitations. Because the app does not identify the end user or establish a connection with actual patient medical records, there is no way of knowing whether the observed episodes were utilized for actual patient cases. However, data from New York's comprehensive ADE reduction quality improvement initiative do suggest that the app is a promising component of a successful intervention campaign to improve anticoagulation safety. In its routine quality improvement role for CMS, IPRO assessed the quarterly rates of bleeding and thromboembolic events resulting in hospitalization within 30 days of an elective surgical procedure among Medicare beneficiaries who had Part D claims for anticoagulants and who were residing within the metro areas of New York State where MAPPP app sessions were identified. In the first calendar quarter of 2016 (ie, prior to app release), IPRO identified 154 ADEs among all eligible cases (154/10,855, 1.42%). In the last quarter of that year (ie, post-MAPPP app launch), among 13,948 cases, 1.13% (157/13,948) ADEs were identified, indicating an approximately 20% relative reduction in the rate of ADEs.

While such improvements cannot be attributed directly to use of the app, efforts are currently underway under a CMS Special Innovation Project awarded to IPRO to facilitate the integration of the app's logic into the electronic health record systems of 3 health systems as active clinical decision support using Substitutable Medical Applications reusable technologies on Fast Healthcare Interoperability Resources when possible, which will allow a direct assessment of the patient-level impact. Incorporation of the app into existing clinical workflows with executive-level adoption buy-in as facility policy has the potential to affect all patients undergoing relevant procedures and will provide access to the data needed to objectively evaluate the app's clinical and financial impact. Training on MAPPP app use via recorded webinar and associated patient education materials can be found on the MAPPP app website [25]. Electronic health record integration is anticipated to scale MAPPP use globally. Results of the project are anticipated in 2019.

Similar apps have been developed by reputable and authoritative organizations such as the University of Michigan's MAQI2 Anticoagulation Toolkit and the ManageAnticoag app developed by the American College of Cardiology [27,28].

Conclusions

The MAPPP app is a successful means of disseminating current guidance on periprocedural anticoagulant use, as indicated by broad global uptake and upward trends in utilization. Lack of access to provider- and patient-specific data precludes objective evaluation of the clinical impact of the app. An ongoing study incorporating app logic into the electronic health records of 3 health systems will provide a more definitive evaluation of the clinical impact of the app logic.

Acknowledgments

The authors disclose receipt of the following financial support for the research, authorship, and publication of this article: The consensus statement was produced by the Care Transitions Task Force of the New York State Anticoagulation Coalition, which was assembled and supported by IPRO, the Centers for Medicare & Medicaid Services (CMS)- designated Quality Innovation Network-Quality Improvement Organization for New York State and lead for the Atlantic Quality Innovation Network under the 10th and 11th Statements of Work. The analyses upon which this publication is based were performed under Contract Number HHSM-500-2014-QIN013I, funded by the CMS, an agency of the US Department of Health and Human Services. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services; furthermore, the mention of trade names, commercial products, or organizations does not imply endorsement by the US government. The authors assume full responsibility for the accuracy and completeness of the ideas presented.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Management of Anticoagulation in the Peri-Procedural Period Mobile App (MAPPP App). Brief video demonstration of utilizing of the app to guide clinical decisions regarding anticoagulants prior to elective invasive medical procedures and surgeries.

[MP4 File (MP4 Video), 14MB - mhealth_v6i12e11090_app1.mp4]

References

1. Budnitz DS, Shehab N, Kegler SR, Richards CL. Medication use leading to emergency department visits for adverse drug events in older adults. Ann Intern Med 2007 Dec 04;147(11):755-765. [Medline: 18056659]
2. Shehab N, Sperling LS, Kegler SR, Budnitz DS. National estimates of emergency department visits for hemorrhage-related adverse events from clopidogrel plus aspirin and from warfarin. Arch Intern Med 2010 Nov 22;170(21):1926-1933. [doi: 10.1001/archinternmed.2010.407] [Medline: 21098354]
3. Budnitz DS, Lovegrove MC, Shehab N, Richards CL. Emergency hospitalizations for adverse drug events in older Americans. N Engl J Med 2011 Nov 24;365(21):2002-2012. [doi: 10.1056/NEJMsa1103053] [Medline: 22111719]
4. Shehab N, Lovegrove MC, Geller AI, Rose KO, Weidle NJ, Budnitz DS. US Emergency Department Visits for Outpatient Adverse Drug Events, 2013-2014. JAMA 2016 Nov 22;316(20):2115-2125. [doi: 10.1001/jama.2016.16201] [Medline: 27893129]
5. Barnes GD, Lucas E, Alexander GC, Goldberger ZD. National Trends in Ambulatory Oral Anticoagulant Use. Am J Med 2015 Dec;128(12):1300-5.e2 [FREE Full text] [doi: 10.1016/j.amjmed.2015.05.044] [Medline: 26144101]
6. Budnitz DS, Pollock DA, Mendelsohn AB, Weidenbach KN, McDonald AK, Annest JL. Emergency department visits for outpatient adverse drug events: demonstration for a national surveillance system. Ann Emerg Med 2005 Feb;45(2):197-206. [doi: 10.1016/j.annemergmed.2004.09.020] [Medline: 15671977]
7. Field TS, Gurwitz JH, Harrold LR, Rothschild JM, Debellis K, Seger AC, et al. Strategies for detecting adverse drug events among older persons in the ambulatory setting. J Am Med Inform Assoc 2004;11(6):492-498 [FREE Full text] [doi: 10.1197/jamia.M1586] [Medline: 15299000]
8. Gurwitz JH, Field TS, Harrold LR, Rothschild J, Debellis K, Seger AC, et al. Incidence and preventability of adverse drug events among older persons in the ambulatory setting. JAMA 2003 Mar 05;289(9):1107-1116. [Medline: 12622580]
9. Garcia DA, Witt DM, Hylek E, Wittkowsky AK, Nutescu EA, Jacobson A, Anticoagulation Forum. Delivery of optimized anticoagulant therapy: consensus statement from the Anticoagulation Forum. Ann Pharmacother 2008 Jul;42(7):979-988. [doi: 10.1345/aph.1L098] [Medline: 18559954]
10. Nutescu EA, Wittkowsky AK, Burnett A, Merli GJ, Ansell JE, Garcia DA. Delivery of optimized inpatient anticoagulation therapy: consensus statement from the anticoagulation forum. Ann Pharmacother 2013 May;47(5):714-724 [FREE Full text] [doi: 10.1345/aph.1R634] [Medline: 23585642]

11. Spyropoulos AC, Viscusi A, Singhal N, Gilleylen J, Kouides P, Howard M, et al. Features of electronic health records necessary for the delivery of optimized anticoagulant therapy: consensus of the EHR Task Force of the New York State Anticoagulation Coalition. Ann Pharmacother 2015 Jan;49(1):113-124. [doi: 10.1177/1060028014555176] [Medline: 25325906]
12. Dreijer AR, Kruip MJHA, Diepstraten J, Polinder S, Brouwer R, Leebeek FWG, et al. Antithrombotic stewardship: a multidisciplinary team approach towards improving antithrombotic therapy outcomes during and after hospitalisation: a study protocol. BMJ Open 2016 Dec 20;6(12):e011537 [FREE Full text] [doi: 10.1136/bmjopen-2016-011537] [Medline: 27998897]
13. Padron M, Miyares MA. Development of an anticoagulation stewardship program at a large tertiary care academic institution. J Pharm Pract 2015 Feb;28(1):93-98. [doi: 10.1177/0897190013514091] [Medline: 24326411]
14. Reardon DP, Atay JK, Ashley SW, Churchill WW, Berliner N, Connors JM. Implementation of a Hemostatic and Antithrombotic Stewardship program. J Thromb Thrombolysis 2015 Oct;40(3):379-382. [doi: 10.1007/s11239-015-1189-3] [Medline: 25669625]
15. Wychowski MK, Ruscio CI, Kouides PA, Sham RL. The scope and value of an anticoagulation stewardship program at a community teaching hospital. J Thromb Thrombolysis 2017 Apr;43(3):380-386. [doi: 10.1007/s11239-016-1455-z] [Medline: 27878507]
16. National Action Plan for Adverse Drug Event Prevention. Washington, DC: U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 2014 URL: https://health.gov/hcq/pdfs/ade-action-plan-508c.pdf [WebCite Cache ID 6xy0ZdcZd]
17. Quality Improvement Organizations. American Health Quality Association URL: http://www.ahqa.org [WebCite Cache ID 6xy5M88cv]
18. Dubois V, Dincq A, Douxfils J, Ickx B, Samama C, Dogné JM, et al. Perioperative management of patients on direct oral anticoagulants. Thromb J 2017;15:14 [FREE Full text] [doi: 10.1186/s12959-017-0137-1] [Medline: 28515674]
19. Healey JS, Eikelboom J, Douketis J, Wallentin L, Oldgren J, Yang S, RE-LY Investigators. Periprocedural bleeding and thromboembolic events with dabigatran compared with warfarin: results from the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) randomized trial. Circulation 2012 Jul 17;126(3):343-348. [doi: 10.1161/CIRCULATIONAHA.111.090464] [Medline: 22700854]
20. Sherwood MW, Douketis JD, Patel MR, Piccini JP, Hellkamp AS, Lokhnygina Y, ROCKET AF Investigators. Outcomes of temporary interruption of rivaroxaban compared with warfarin in patients with nonvalvular atrial fibrillation: results from the rivaroxaban once daily, oral, direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and embolism trial in atrial fibrillation (ROCKET AF). Circulation 2014 May 06;129(18):1850-1859 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.113.005754] [Medline: 24552831]
21. Spyropoulos AC, Al-Badri A, Sherwood MW, Douketis JD. Periprocedural management of patients receiving a vitamin K antagonist or a direct oral anticoagulant requiring an elective procedure or surgery. J Thromb Haemost 2016 Dec;14(5):875-885 [FREE Full text] [doi: 10.1111/jth.13305] [Medline: 26988871]
22. Sié P, Samama CM, Godier A, Rosencher N, Steib A, Llau JV, Working Group on Perioperative Haemostasis, French Study Group on ThrombosisHaemostasis. Surgery and invasive procedures in patients on long-term treatment with direct oral anticoagulants: thrombin or factor-Xa inhibitors. Recommendations of the Working Group on Perioperative Haemostasis and the French Study Group on Thrombosis and Haemostasis. Arch Cardiovasc Dis 2011 Dec;104(12):669-676 [FREE Full text] [doi: 10.1016/j.acvd.2011.09.001] [Medline: 22152517]
23. Dincq A, Lessire S, Douxfils J, Dogné JM, Gourdin M, Mullier F. Management of non-vitamin K antagonist oral anticoagulants in the perioperative setting. Biomed Res Int 2014;2014:385014 [FREE Full text] [doi: 10.1155/2014/385014] [Medline: 25276784]
24. Douketis JD, Spyropoulos AC, Spencer FA, Mayr M, Jaffer AK, Eckman MH, et al. Perioperative management of antithrombotic therapy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest 2012 Feb;141(2 Suppl):e326S-e350S [FREE Full text] [doi: 10.1378/chest.11-2298] [Medline: 22315266]
25. Management of Anticoagulants in the Peri-Procedural Period (MAPPP). IPRO URL: http://mappp.ipro.org/ [WebCite Cache ID 6xy1jq0TZ]
26. Anticoagulation Management Discharge Communication (Audit Tool). Quality Innovation Network-Quality Improvement Organizations URL: http://qioprogram.org/campaign-meds-management-resource-page-20 [WebCite Cache ID 6xy1SSzqI]
27. MAQI2Anticoagulation Toolkit App. Michigan Anticoagulation Quality Improvement Initiative URL: http://anticoagulationtoolkit.org [WebCite Cache ID 71ZKlSeYa]
28. ManageAnticoag App. American College of Cardiology URL: https://www.acc.org/tools-and-practice-support/mobile-resources/features/manageanticoag [WebCite Cache ID 71ZLD8apZ]

Abbreviations

ADE: adverse drug event

CMS: Centers for Medicare & Medicaid Services
ED: emergency department
MAPPP: Management of Anticoagulation in the Periprocedural Period
NYSACC: New York State Anticoagulation Coalition
QIN-QIO: Quality Innovation Network-Quality Improvement Organization

Edited by G Eysenbach; submitted 18.05.18; peer-reviewed by S Kaatz, E Goldstein, S Wymer, J Syverson, K Cabral, K Bailey; comments to author 01.08.18; revised version received 10.08.18; accepted 14.08.18; published 21.12.18.

Please cite as:
Spyropoulos AC, Myrka A, Triller DM, Ragan S, York C, King JM, Lee TK
Uptake and Utilization of the Management of Anticoagulation in the Periprocedural Period App: Longitudinal Analysis
JMIR Mhealth Uhealth 2018;6(12):e11090
URL: http://mhealth.jmir.org/2018/12/e11090/
doi:10.2196/11090
PMID:30578235

Original Paper

Inferring Physical Function From Wearable Activity Monitors: Analysis of Free-Living Activity Data From Patients With Knee Osteoarthritis

Vibhu Agarwal[1], MS; Matthew Smuck[2], MD; Christy Tomkins-Lane[3], PhD; Nigam H Shah[1], MBBS, PhD

[1]Center for Biomedical Informatics Research, Stanford University, Stanford, CA, United States

[2]Stanford University Hospital and Clinics, Stanford, CA, United States

[3]Department of Health and Physical Education, Mount Royal University, Calgary, AB, Canada

Corresponding Author:
Vibhu Agarwal, MS
Center for Biomedical Informatics Research
Stanford University
Medical School Office Building, X225
1265 Welch Road
Stanford, CA, 94305
United States
Phone: 1 (650) 724 3979
Email: vibhua@stanford.edu

Abstract

Background: Clinical assessments for physical function do not objectively quantify routine daily activities. Wearable activity monitors (WAMs) enable objective measurement of daily activities, but it remains unclear how these map to clinically measured physical function measures.

Objective: This study aims to derive a representation of physical function from daily measurements of free-living activity obtained through a WAM. In addition, we evaluate our derived measure against objectively measured function using an ordinal classification setup.

Methods: We defined function profiles representing average time spent in a set of pattern classes over consecutive days. We constructed a function profile using minute-level activity data from a WAM available from the Osteoarthritis Initiative. Using the function profile as input, we trained statistical models that classified subjects into quartiles of objective measurements of physical function as measured through the 400-m walk test, 20-m walk test, and 5 times sit-stand test. Furthermore, we evaluated model performance on held-out data.

Results: The function profile derived from minute-level activity data can accurately predict physical performance as measured through clinical assessments. Using held-out data, the Goodman-Kruskal Gamma statistic obtained in classifying performance values in the first quartile, interquartile range, and the fourth quartile was 0.62, 0.53, and 0.51 for the 400-m walk, 20-m walk, and 5 times sit-stand tests, respectively.

Conclusions: Function profiles accurately represent physical function, as demonstrated by the relationship between the profiles and clinically measured physical performance. The estimation of physical performance through function profiles derived from free-living activity data may enable remote functional monitoring of patients.

(JMIR Mhealth Uhealth 2018;6(12):e11315) doi:10.2196/11315

KEYWORDS

physical function; passive monitoring; physical function profile; wearable activity data; statistical learning

Introduction

Physical function is an important indicator of physiological well-being. Recently, physical status has become an outcome of interest in most medical specialties [1-4] and is increasingly regarded as the “sixth vital sign” [5]. Attempts at arresting and managing the functional decline must start with an evaluation of the baseline functional status. For example, maximizing improvement in advanced osteoarthritis requires knowing a

patient's baseline function to detect any improvement. Therefore, valid metrics to monitor physical function are necessary [6-8]. The International Classification of Functioning, Disability and Health [9] characterizes physical function in 2 distinct categories—capacity and performance. Capacity is the capability of a person to complete a given task in a controlled environment (eg, a timed walking test or a sit-stand test), while performance is what a person does in his or her current environment (eg, real-life physical activity monitoring). Traditionally, physical performance is estimated by surveys and self-reported questionnaires. One example is the assessment of one's ability to complete the daily activities necessary to live independently (including bathing, dressing, toileting, transferring, maintaining bowel and bladder continence, and feeding), collectively referred to as the activities of daily living (ADLs) [10,11] and are typically measured by surveys. Disability indexes based on ADLs can differentiate healthy aging patients, patients with mild cognitive impairment, and patients with dementia [12]. However, ADL scores may have a response bias from self-reporting and low sensitivity to changes in high-functioning older adults [13]. In contrast, physical capacity measures (such as walking and sit-to-stand speeds and grip strength observed under supervision) capture variation across a wider range of physical function, including initial changes in the early stages of decline [14-16]. The main drawback of such capacity measures is that they require substantial time and effort from patients and researchers, as well as access to specialized facilities. The relationship between physical activity and physical function measures is a topic of active research [17-21].

Although the need to measure physical function is widely appreciated, self-reported assessments of physical performance are inadequate owing to poor discrimination and biases and difficulties in recalling historical activities; physical capacity measures require adherence to specific test protocols and are usually limited to research settings. Seeking a more simple and accurate measure, we have created a novel method for inferring physical function based on objective measurements of daily physical activity obtained from a wearable device. Our work enables quantitative monitoring of physical function—the first step toward improved precision in clinical research and practice.

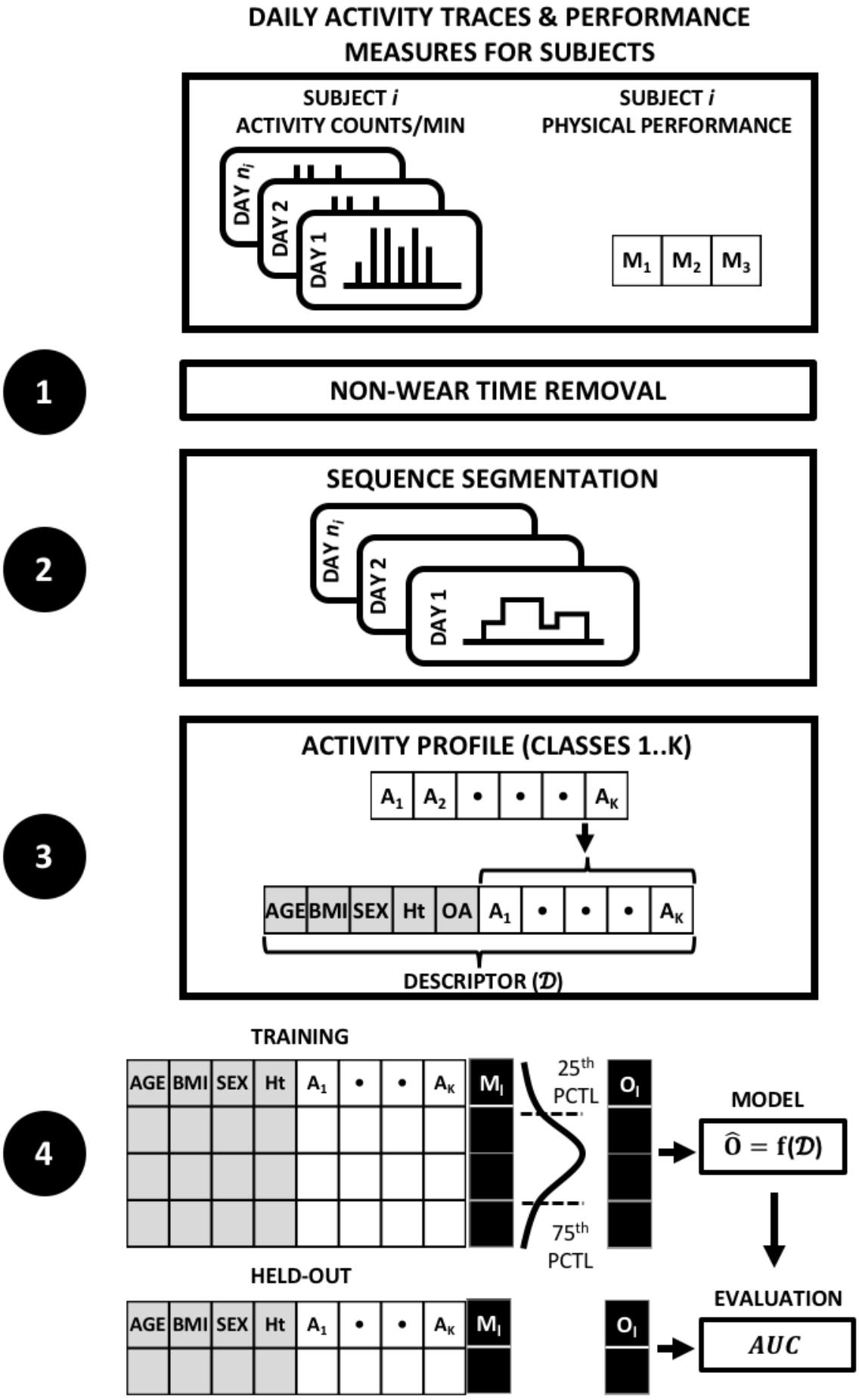
Wearable activity monitors (WAMs), typically equipped with one or more accelerometers, provide a convenient way to measure physical activity objectively [22-24]. However, attempts to use WAMs to link the measured physical activity and physical function have been limited by their reliance on traditional methods of analyzing WAM data [25-28]. Two research groups [20,29] have demonstrated that the measured physical activity and physical capacity are associated but independent domains of physical function. For example, a change in physical capacity (eg, on the 6-minute walk test) need not imply a corresponding change in real-life activity levels. Interestingly, both research groups concluded that differentiating physical activity into classes leads to a stronger association with physical function, compared with a univariate measure based on average acceleration. WAMs typically measure the aggregate velocity change over a period—which by itself was considered inadequate for distinguishing classes of activities. We hypothesized that higher-order patterns in daily activity recorded by a WAM would correlate with physical function. We defined pattern classes from daily activity data using an unsupervised approach and used this information to create a *function profile*, which represents the mean allocation of time to different pattern classes. Using machine learning techniques, we classified activity profiles into discrete quartiles of commonly used measures of physical function such as the 400-m walk test (400MWT).

Studies with WAMs have, thus far, focused on the following:

1. Evaluation of measurement reliability and validity and characterizing activity phenotypes by patterns in free-living activity data [20,30-33].
2. Developing models of isolated activities and postures using supervised learning [34-37].
3. Developing activity-based models of physical capacity wherein subjects undergo instrumented versions of various capacity tests as summarized in a recent literature review by Grimm and Bolink [38].

Furthermore, Gresham et al used daily activity metrics (steps, distance, and stairs) to compute correlations with the clinically measured performance status in patients with advanced cancer [39]. In a study on nursing home residents, Merilahti et al reported an association between features derived from daily free-living activity and patient-reported physical function [40]. However, none of the studies mentioned above has modeled physical function using daily free-living activity—a crucial step in medical applications that require passive monitoring of function. Our research contribution is to use WAM data to characterize daily free-living activity into pattern classes and infer physical function based on the pattern classes. This study demonstrates the feasibility of distinguishing physical function categories with high sensitivity and specificity, and discusses potential uses in medical research and treatment. Figure 1 illustrates our overall workflow. Daily activity, measured as counts per minute, was recorded for 2001 subjects in year 4 of the OAI study. For each subject, various objective measurements were obtained from which we selected results for the 400-m walk test (400MWT), average pace on the 20-m walk test (20MPACE), and 5 sit-stand time (5CS), labeled as M1, M2, and M3 in Figure 1. Thereafter, nonwear time was excluded from activity traces, daily activity count sequences were segmented, and a composite feature descriptor with a daily activity profile was constructed for each subject. Finally, quantitative response values were converted to ordinal values based on empirical quantiles obtained from the training partition and a classifier for the feature descriptor was trained on the training partition of the feature matrix (80%) and evaluated on the held-out partition.

Figure 1. Estimating physical function from daily activity traces (overall workflow). Ht: height; OA: osteoarthritis; PCTL: percentile; AUC: area under the receiver operating characteristic curve.



Methods

Data

We used publicly available data from the Osteoarthritis Initiative, (OAI) which follows a cohort of subjects who either had a clinical diagnosis (progression subcohort) of oseoarthritis or were at risk at baseline (incidence subcohort). The OAI has daily accelerometer measurements for subjects who participated in a physical activity study; these participants were instructed to use an ActiGraph GT1M uniaxial accelerometer (ActiGraph; Pensacola, FL, USA) continuously for up to 7 consecutive days, except during sleep and water activities. The ActiGraph GT1M is a compact, hip-worn device that measures dynamic acceleration in the range of 0.05-2.0 g; its validity and reliability have been established previously [41-43]. Participants maintained a daily log of water and cycling activities, as the accelerometer may not have been able to capture these accurately. A post facto analysis revealed that participants spent little time in these activities (median 0 minutes/day; interquartile range 0.0-3.4 minutes/day), indicating that little activity was missed by the monitors. Table 1 summarizes the key attributes of the physical activity study data (Multimedia Appendix 1).

The accelerometer data in OAI consist of activity counts per minute. An activity count is a weighted sum of discretely-sampled (30 Hz) values of one-dimensional acceleration. We used established guidelines to determine the wear time and valid days of activity monitoring, as reported previously [44]. Since 0 or low values of activity counts could also arise from nonwear time, we excluded nonwear periods. Continuous runs of 0 counts for >90 minutes (allowing for interruptions of up to 2 consecutive minutes with <100 counts) were discarded as nonwear periods. A day with a wear time of, at least, 10 hours was considered valid. Furthermore, objective, as well as patient-reported measures of physical function, were recorded during patient follow-up visits.

Objective Measures of Function

The Osteoarthritis Research Society International [45] recommends testing of activities that are typically affected by OA. We selected 3 OAI performance measures that had equivalents in Osteoarthritis Research Society International's recommended tests, which were as follows: the 400MWT, for which longer completion times are associated with a higher risk of mobility limitation and disability (adjusted hazard ratios, 4.43, P<.001), as well as a higher risk of death (adjusted hazard ratio, 3.23, P<.001) for subjects in the highest quartile [46]; the average pace in a 20-m walk test (20MPACE) is the closest available short walk-length evaluation in the OAI dataset that is used for gait speed assessment; the number of sit-to-stands per second measured over 5 repetitions (5CSPACE) as a measure of the sit-to-stand function, which has good test-retest reliability [47].

Relationship With Daily Activity

Physical function is defined as the repertoire and relative proportion of activities that a subject accomplishes in a given environment. We recovered segments representing homogenous activities from the daily sequences of counts per minute obtained from a WAM and defined *pattern classes* based on similar segments. A subject's function profile was computed as average minutes allocated to each pattern class. Finally, we inferred mappings from the function profile to the objective measurements of the function described in the earlier section using supervised learning.

Pattern Classes and Function Profile

We used the change-point analysis algorithm by James and Matteson [48] to segment counts-per-minute sequences; this algorithm searches for segment boundaries such that each segment represents a change in the distribution of the time-ordered counts with respect to preceding and subsequent segments. Figure 2 illustrates a counts-per-minute sequence for a typical subject on a given day and the segments that are recovered through change-point analysis (as shown by the horizontal red lines). Each segment is an instance of a pattern class whose mean and SD are estimated by the sample mean and SD of the segment.

Each segment was indexed using the mean and SD of the counts-per-minute values; this representation improves discrimination between classes of activity patterns (henceforth referred to as *pattern classes*) [49].

A *pattern class* is a bounded region in the segment feature space. Our feature space F consists of all (m, s) vectors: $m \in [0, M]$, $s \in [0, S]$, where M and S are the maximum mean and SD over all segments found through the segmentation. A pattern class is defined by a pair of intervals such as $[m_1, m_2)$ X $[s_1, s_2)$. A segment with mean x and SD y ($m_1 \leq x < m_2$, $s_1 \leq y < s_2$) is an instance of the pattern class so defined. Figure 3 illustrates such a segment represented in the mean-SD space spanned by all segments and its assignment to a pattern class $[m_1, m_2)$ X $[s_1, s_2)$, as shown by the shaded region. Based on the pattern classes obtained from partitioning F, we defined a *function profile* for each subject as the average time allocated to each pattern class per day. The function profile for a subject i is given by

$A_i=(a_{i1}, a_{i2}...a_{iJ})$ *where*

$a_{il}=(1/K_i)\sum_k t_{ilk}$

J: the number of pattern classes

$k=1...K_i$, *the number of days of observations for subject i*

t_{ijk}: *the number of minutes spent by subject i in pattern class j, on day k*

As seen in Figure 4, the number of instances of a pattern class decrease as the mean and SD increase resulting in a sparse daily activity profile. $D_i=(BMI_i, Age_i, Sex_i, Height_i, OA_i, A_i)$

Table 1. Key attributes of the knee osteoarthritis subjects providing physical activity data.

Characteristics	Value
Total number of subjects (N=2001), n (%)	
Incidence subcohort	1490 (74.46)
Progression subcohort	505 (25.24)
Control subcohort	6 (0.30)
Gender (male), n (%)	891 (44.53)
Body mass index, mean (SD)	28.52 (4.87)
Mean Comorbidity Index	0.52
Median days of activity	7

Figure 2. Segmentation of counts-per-minute sequences. Dur: duration.

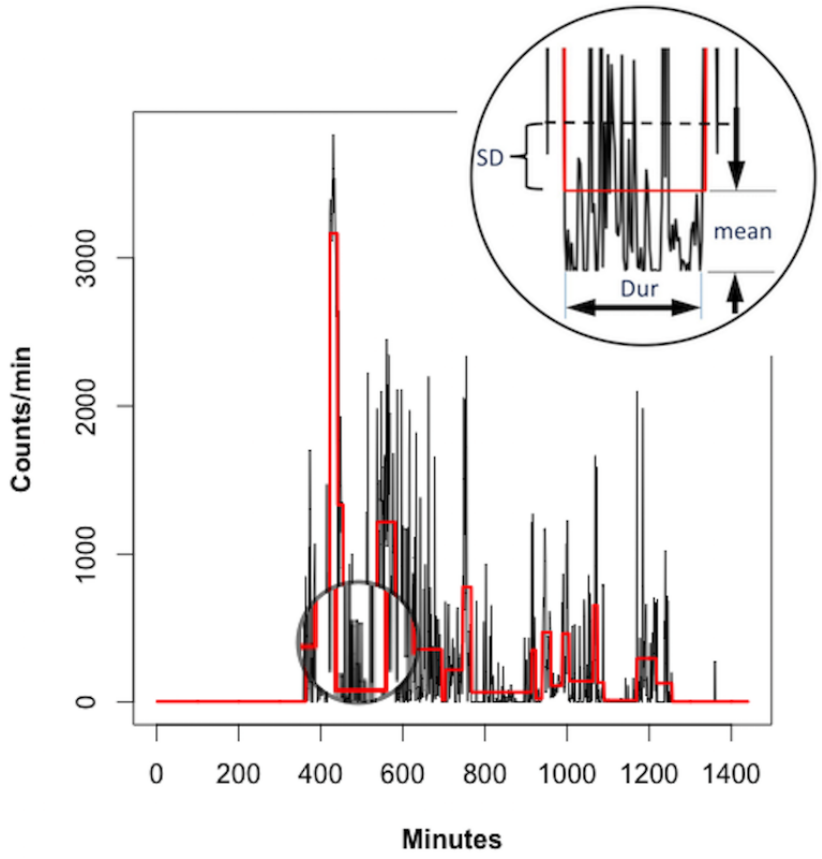


Figure 3. The mean and SD space containing all segments is partitioned into bounding regions, each defined by a mean and an SD interval.

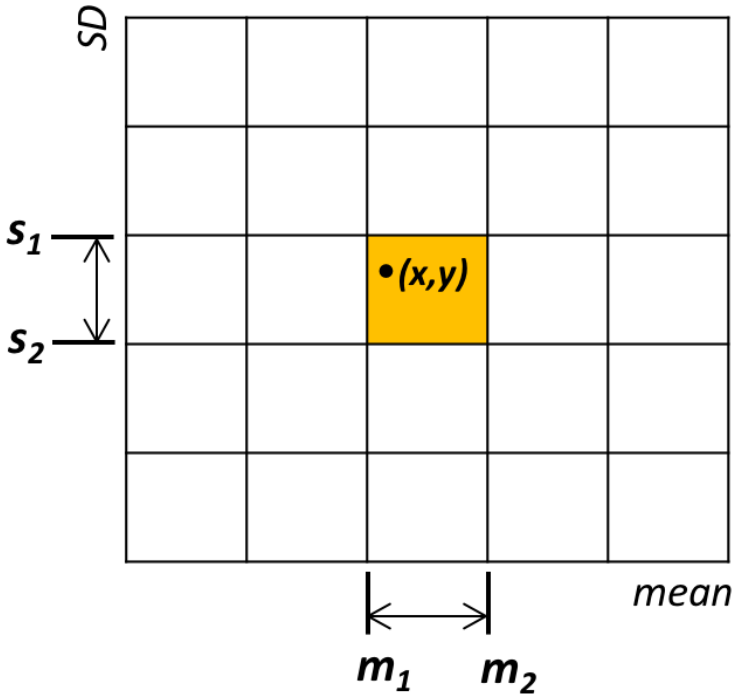
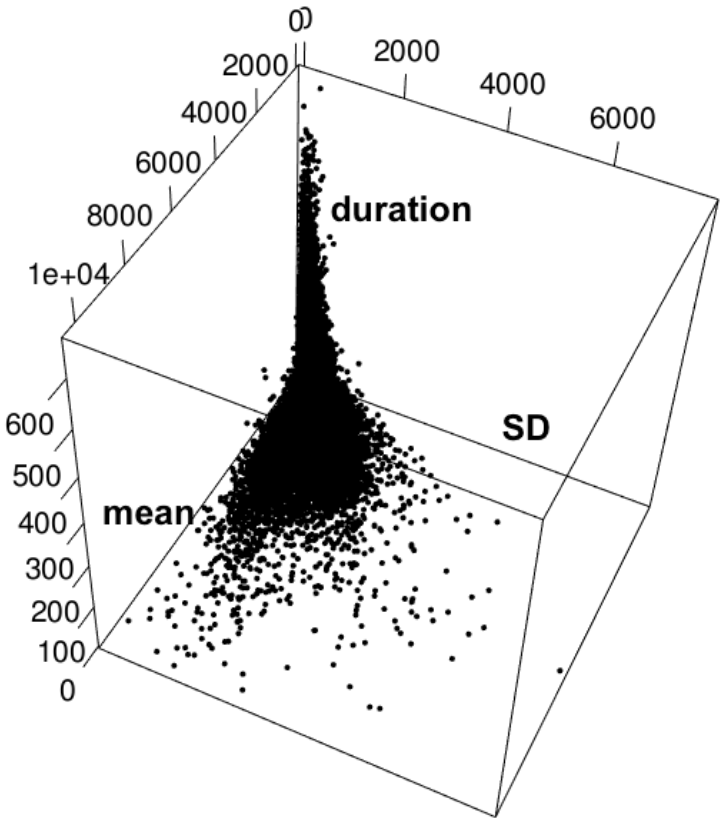


Figure 4. Scatter plot of the mean, SD, and duration of the segments.



Supervised Learning

We defined a composite descriptor D_i=(BMI_i, Age_i, Sex_i, $Height_i$, OA_i, A_i) for each subject i in our data, where OA_i refers to a subject's baseline status (healthy, at-risk, or progressive disease) and A_i is the function profile. A regression function f (D) that maps D_i to an objective measure of physical capacity can be obtained by minimizing the expected squared error loss.

Medical studies commonly group continuous variables into quantiles for ease of interpretation and analysis [46,50,51]. We, therefore, defined our response variable by grouping the objective measure of physical capacity into ordered categories

1<2<3. As shown in Figure 1, categories 1 and 3 represented values in the lowest and highest quartiles, respectively, and 2 represented values spanning the interquartile range for a specific response. Classes 1 and 3 correspond to the upper and lower quartiles on the physical capacity measurements and, therefore, contain only half as many observations as in class 2. To address the imbalance, each observation was weighted by its class prevalence in the fitting procedure.

Generalized Additive Models (GAM) can identify and characterize nonlinear regression effects through an additive specification of nonparametric functions of the predictors. We used GAMs because fits from quantitative regressions suggest that at higher values, linearity in the predictors may not be a justifiable modeling assumption (Multimedia Appendix 2). A GAM may be specified as follows:

$g(\mu(X)) = \alpha + f_1(X_1) + f_2(X_2) + ... + f_p(X_p)$ where

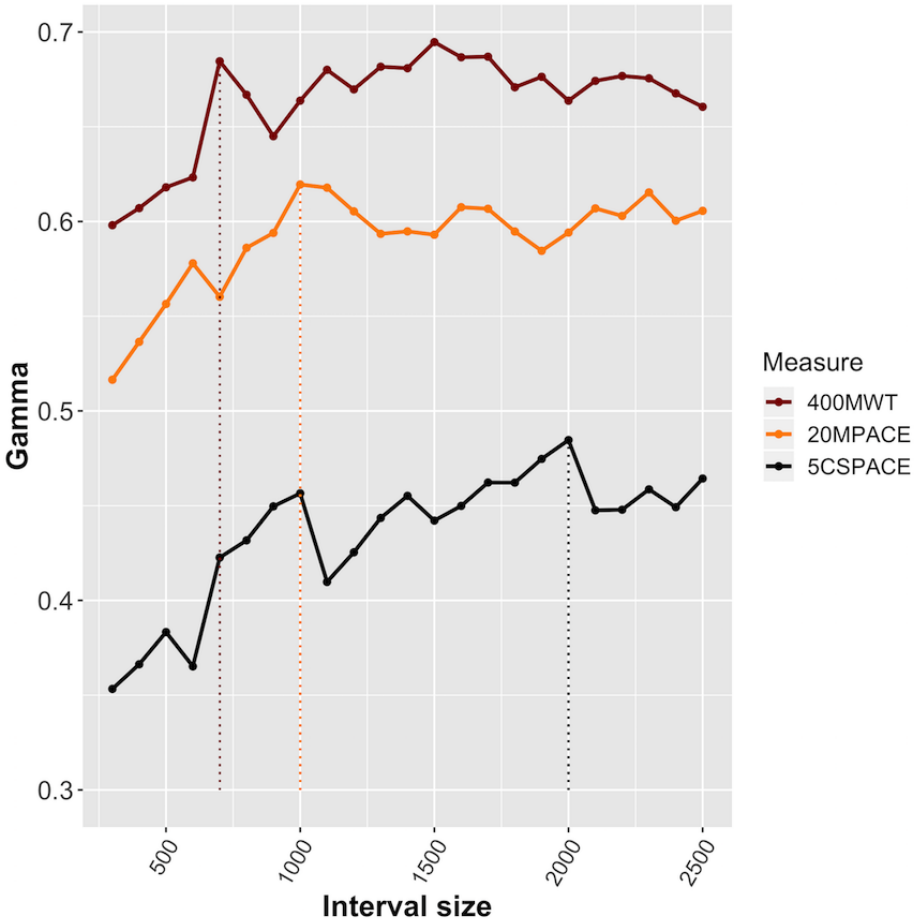
$\mu(X)$ denotes the conditional mean of the response, that is, E[Y|X]

$g(\mu(.)$ is the link function

$f_1 ... f_p$ are the unspecified smooth functions for each of the p predictors

Unspecified functions of predictors are smoothers (typically kernels or cubic splines) that are estimated simultaneously using a backfitting algorithm [52]. The estimated reveal the nature of the predictor-response relationship. The function profile depends on the pattern classes, which are defined as intervals in the mean-SD space covering all segments. The size of the 2D interval in feature space that defines our pattern classes is a tuning parameter. Small intervals allow instances from adjacent classes to be in close proximity, increasing the correlation between activity profile elements. We ascertained the optimal size of the 2D interval—with equal mean and SD intervals—through repeated 5-fold cross-validation on our training data, as shown in Figure 5 (dashed lines indicate the optimal region size). Our intuition for the different optimal size for the 5CS model is that daily activities that involve sit-stand transitions are subsumed in the function classes defined on wider intervals. For example, sit-stand-walk and walk-stand-sit (measured by per minute activity counts) are transitions to and from activities characterized by large mean counts, whereas sit-stand and stand-sit are transitions to and from low mean count activities. On the other hand, most daily activities require some level of lower extremity strength, balance and gait initiation, and control capability—each of which are necessary for walking. Therefore, it seems reasonable that a profile constructed from function classes that distinguish between such activities will have a high correlation with walking test results.

Figure 5. The mean Gamma (Goodman-Kruskal rank correlation between the predicted and true responses) in 5-fold cross-validation for 20MPACE (20-m walk test), 5CSPACE (number of sit-to-stands per second measured over 5 repetitions), and 400MWT (400-m walk test) models.



We evaluated cross-validation performance using the mean Goodman-Kruskal Gamma [53], which measures the rank correlation between the true and predicted categories (Multimedia Appendix 2). For optimal bin sizes, GAMs were refit using the full training data and features based on the optimal bin size, and ordered categorical for the response family using the mgcv package [54]. We evaluated the Goodman-Kruskal Gamma for the predicted and true classes, using the held-out data.

Results

Principal Results

As described in the Methods section, we found homogeneous segments from daily activity sequences of counts per minute and defined pattern classes based on similar segments. A subject's function profile was average daily minutes allocated to each pattern class. Finally, we learned mappings from the function profile to the objective measurements of physical capacity. Table 2 summarizes classifier performance for the GAMs on the held-out data using the function profiles based on the optimal interval sizes. The values in parentheses indicate improvement over baseline performance without function profile predictors.

Including the activity profile improved the held-out Gamma by 4%-10%, compared with classifiers in which the activity profile was excluded from the predictors, with higher improvement in classification of walking test results (Multimedia Appendix 3).

Predictors of Physical Function

GAMs fit smooth functions for each predictor in the model that additively contribute to the value of a latent variable. The model fitting algorithm [52] also estimates thresholds, whose values in relation to the latent variable computed from the smooth functions determine the ordered categorical response. Relationships between response and predictors in a GAM may be studied by plots of smoothers fitted by the GAMs. We studied predictors that were significant at P=.05 in the GAMs (Figure 6). Predictor-response relationships are shown by the smooth function plots arranged around the grid and linked to the corresponding predictor. We refer to a specific pattern class using the mean and SD interval pairs, as defined in the Methods section. The smooth function plots for pattern classes [0,700) X [701,1400) and [701,1400) X [701,1400) (numbered 2 and 3, respectively, on the mean-SD grid in Figure 6) show that up to 25 daily average minutes in these pattern classes were monotonically associated with improved response in the 400MWT and 20MPACE.

Most pattern classes with a higher mean and SD have low duration, resulting in fewer degrees of freedom for estimating the smooth functions at high values; this explains the wider confidence bands for the function estimates at higher values of daily average minutes. Inspection of a sample of instances from the class [701-1400) X [701-1400) revealed that long duration instances were typically spells of rest punctuated by frequent interruptions. A plausible explanation may be that such interruptions involve sit-stand transitions and, therefore, these instances are associated with the improved 5CSPACE response. Pattern classes with the mean interval [0-700) are not associated with the 5CSPACE response.

Pattern classes [2801,3500) X [0,700) and [2801,3500) X [1401,2100), numbered 4 and 5, respectively, are associated with the 400MWT response. The smooth function plots for these classes suggest that higher daily average minutes in both classes were associated with improved long-walk capacity. The association with increased completion times with >20 daily average minutes in the class [2801,3500) X [0,700) was due to instances comprising of mostly sedentary activity. Furthermore, infrequent occurrences of such instances resulted in wide estimate intervals for the smooth function.

Table 2. Gamma for generalized additive models evaluated on held-out data.

Predictors	Physical capacity measurement	Gamma
BMI[a], age, sex, height, OA[b] subcohort, function profile	400MWT[c]	0.62 (0.10)[d]
BMI, age, sex, height, OA subcohort, function profile	20MPACE[e]	0.53 (0.07)[d]
BMI, age, sex, height, OA subcohort, function profile	5CSPACE[f]	0.51 (0.04)[d]
BMI, age, sex, height, OA subcohort	400MWT	0.52
BMI, age, sex, height, OA subcohort	20MPACE	0.46
BMI, age, sex, height, OA subcohort	5CSPACE	0.47

[a]BMI: body mass index.

[b]OA: osteoarthritis.

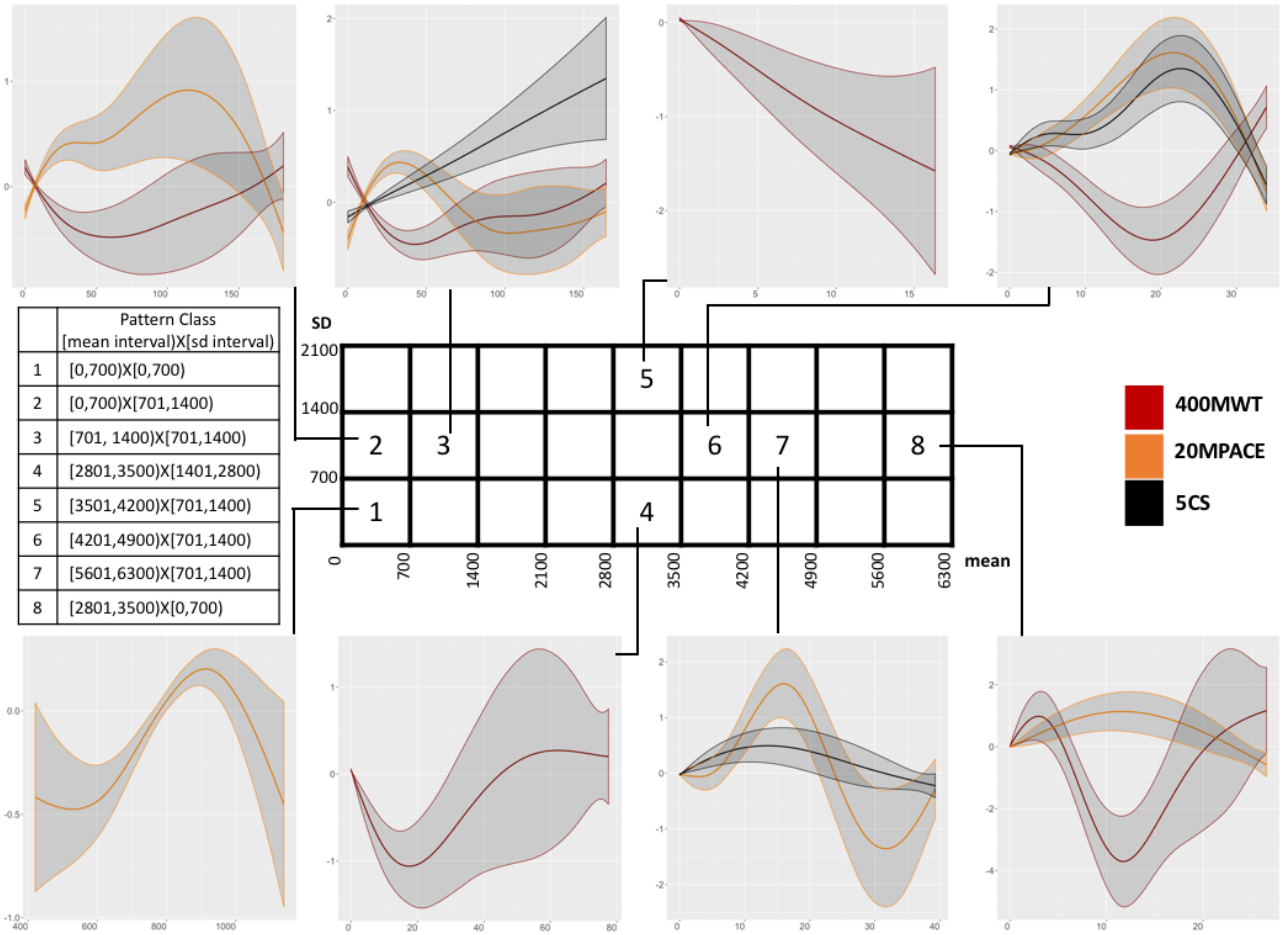
[c]400MWT: 400-m walk test

[d]The values in parentheses indicate improvement over baseline performance without function profile predictors.

[e]20MPACE: the average pace in a 20-m walk test.

[f]5CSPACE: number of sit-to-stands per second measured over 5 repetitions.

Figure 6. The grid boxes represent pattern classes labeled with the mean interval (x-axis) and the SD interval on the y-axis. 400MWT: 400-m walk test; 20MPACE: 20-m walk test; 5CSPACE: number of sit-to-stands per second measured over 5 repetitions.



Approximations of the distribution of activity counts in any given pattern class can be obtained by tail probability bounds. For example, use Chebyshev's inequality, $P(|X - \mu| > k\sigma) < (1/k^2)$ where μ and σ are the midinterval values of the mean and SD intervals, respectively, for a given pattern class. For the class [2801,3500) X [0,700), we obtained:

$P(|X - 3150| < k.\ 350) > (1 - 1/k^2)$ for k=1.8, implying that at least 70% of activity counts per minute were between 2520 and 3780. Thus, most of the activity in the class [2801,3500) X [0,700) was likely to be in the lower moderate-intensity range. Similarly, for the class [2801,3500) X [1401,2100), we noted that at least 70% of activity counts per minute were below 6300, indicating a mix of activity moderate and vigorous activity.

Moderate-to-Vigorous Activity With Knee Malalignment

In the pattern class [3501,4200) X [701,1400), numbered 6 in the mean-SD grid of Figure 6, an increase in daily average minutes was monotonically associated with improved responses in all 3 capacity measures up to 20 minutes/day. However, an increase of >20 minutes was associated with a decline. Unlike the classes with low mean and SD, instances >20 minutes did not represent sedentary activity. A drop in physical function with increased time in moderate-to-vigorous activity is counterintuitive. To understand this finding, we reviewed patient-reported outcomes on the Physical Activity Scale for the Elderly (PASE). The PASE measures engagement in different kinds of daily activities related to leisure, household, and occupational work in the elderly [55]. In addition, we reviewed joint exam results reporting varus (bow-legged) and valgus (knock-kneed) alignments for the same subjects; this information is summarized in Table 3. It suggests that subjects with >20 daily average minutes in the pattern class [3501,4200) X [701,1400) had a higher prevalence and severity of knee deformity, higher time in the pattern class (minutes as well as frequency), and fewer sitting hours along with more walking hours per week.

Table 3. Knee deformity and PASE results of subjects with at least one instance of the pattern class [3501-4200) X [701-1400).

Subject results	Daily average minutes	
	≥20	<20
Number of subjects with	12	255
Varus or valgus deformity in both knees	10	179
Varus or valgus deformity in either knee	11	221
Joint laxity (mild-severe) in either knee	8	109
Average number of days per week in activity	4.3	1.7
Average number of minutes per week in activity	161	43.4
Percentage with sitting hours <2, 2-4, >4 per day in last 7 days[a]	25, 50, 25	19, 55, 24
Percentage walking <1, 1-2, 2-4 h a day in last 7 days[a]	25, 33, 33	40, 40, 11

[a]Physical Activity Scale for the Elderly.

Studies have suggested that in subjects with knee malalignment or laxity, altered tibiofemoral loading could be responsible for biomechanical damage and OA progression [56-58]. A much debated view on the role of the quadriceps in OA is that the greater muscle strength in malaligned or lax knees increases the risk of OA progression [59,60]. If the relationship between the lower extremity strength and the risk of OA progression is confounded by the knee alignment status, a plausible explanation for the decreasing trend discussed above may be that regular investment in the pattern class [3501,4200) X [701,1400) promotes muscle strength but *advances* OA in subjects with malaligned knees. Though the current guidelines for knee OA management recommend muscle strengthening, our analysis highlights the need for a mechanistic investigation of greater power, given that muscle strength is a modifiable risk factor in OA.

Discussion

Principal Findings

To infer physical function from a daily activity trace, it is necessary to derive a representation that conveys information about the daily activity mix. We defined distinct segments from daily activity traces as instances of a set of pattern classes. Doing so transforms a sequence of activity counts into a sequence of pattern classes. Pattern classes provide an informative view of daily physical activity from the perspective of functional ability. Our approach of unsupervised segmentation and the subsequent definition of a set of pattern classes allows a function-based comparison among subjects without the overhead of obtaining annotated activity traces from subjects. This comparison is based on objective measurements and is, perhaps, the first effort to interpret functional outcomes based on pattern classes from free-living activity data, within a clinical research use case. Classifying physical function may be useful in several areas; for example, alternatives to outpatient physical therapy [61] are a topic of active research. Remote monitoring of physical function in daily living could allow rehabilitation programs to be evaluated in a site-less trial setting. We recognize that many clinical apps require a higher performance in physical function classification than obtained with our current models. Our results, however, suggest that this preliminary work may be advanced, potentially with higher resolution activity data.

Limitations

There are 2 main limitations of our methods. First, the mean and SD are likely to be inadequate representations of the activity-generating processes, as they ignore temporal relationships between activity counts. Modeling class instances as subsequences generated by a random process have been proposed [62], and may improve the detection of pattern classes. Second, our approach ignores time ordering between pattern class instances in the daily activity profile. One way to address these limitations may be to learn within- and interclass relationships for a set of daily activity sequences, as a single Bayesian network. In addition, methods to reliably estimate the function profile from missing activity data are needed as nonadherence is a well-known issue in most health studies with wearable devices.

Conclusions

An assessment of physical function based on the ability to perform routine tasks in daily life is desirable. Widely available wearable motion sensors can record daily activity objectively and unobtrusively. We have created an approach for deriving a function profile that represents time spent on various tasks encountered in daily living. Classifiers trained on the function profile were able to predict highest and lowest quartile results of clinically used physical capacity measures. We recovered associations between pattern classes and physical capacity measures, some of which corroborate prior OA research. The idea of representing physical function as a function profile derived from daily free-living activity may enable remote monitoring of patients' physical function.

Acknowledgments

This work was supported by gifts and unrestricted funds. In addition, the authors thank Jason Fries for his suggestions while editing the manuscript and Professor Trevor Hastie for his advice pertaining to data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The osteoarthritis initiative dataset.

[PDF File (Adobe PDF File), 146KB - mhealth_v6i12e11315_app1.pdf]

Multimedia Appendix 2

Quantitative & Ordinal Response Models.

[PDF File (Adobe PDF File), 126KB - mhealth_v6i12e11315_app2.pdf]

Multimedia Appendix 3

Model Evaluation.

[PDF File (Adobe PDF File), 37KB - mhealth_v6i12e11315_app3.pdf]

References

1. Stamm TA, Pieber K, Crevenna R, Dorner TE. Impairment in the activities of daily living in older adults with and without osteoporosis, osteoarthritis and chronic back pain: a secondary analysis of population-based health survey data. BMC Musculoskelet Disord 2016 Mar 28;17:139 [FREE Full text] [doi: 10.1186/s12891-016-0994-y] [Medline: 27020532]
2. Anand S, Johansen KL, Kurella TM. Aging and chronic kidney disease: the impact on physical function and cognition. J Gerontol A Biol Sci Med Sci 2014 Mar;69(3):315-322 [FREE Full text] [doi: 10.1093/gerona/glt109] [Medline: 23913934]
3. Dellapasqua S, Colleoni M, Goldhirsch A. Focusing on physical function limitations in elderly women surviving cancer: any opportunity for improvement? Nat Clin Pract Oncol 2006 Nov;3(11):598-599. [doi: 10.1038/ncponc0633] [Medline: 17080175]
4. Paterson DH, Warburton DE. Physical activity and functional limitations in older adults: a systematic review related to Canada's Physical Activity Guidelines. Int J Behav Nutr Phys Act 2010 May 11;7:38 [FREE Full text] [doi: 10.1186/1479-5868-7-38] [Medline: 20459782]
5. Bierman AS. Functional status: the six vital sign. J Gen Intern Med 2001 Nov;16(11):785-786 [FREE Full text] [Medline: 11722694]
6. Ferriolli E, Skipworth RJE, Hendry P, Scott A, Stensteth J, Dahele M, et al. Physical activity monitoring: a responsive and meaningful patient-centered outcome for surgery, chemotherapy, or radiotherapy? J Pain Symptom Manage 2012 Jun;43(6):1025-1035. [doi: 10.1016/j.jpainsymman.2011.06.013] [Medline: 22269181]
7. Hunter K, Lyon MG. Evaluation and management of polymyositis. Indian J Dermatol 2012 Sep;57(5):371-374 [FREE Full text] [doi: 10.4103/0019-5154.100479] [Medline: 23112357]
8. Kop WJ, Lyden A, Berlin AA, Ambrose K, Olsen C, Gracely RH, et al. Ambulatory monitoring of physical activity and symptoms in fibromyalgia and chronic fatigue syndrome. Arthritis Rheum 2005 Jan;52(1):296-303 [FREE Full text] [doi: 10.1002/art.20779] [Medline: 15641057]
9. World Health Orgainzation. Classification of Functioning, Disability and Health. Geneva: World Health Orgainization; 2017. The ICF: An Overview URL: http://www.who.int/classifications/icf/en/ [accessed 2018-11-06] [WebCite Cache ID 73joKQd6X]
10. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist 1970;10(1):20-30. [Medline: 5420677]
11. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of Illness in the Aged. The index of ADL: A Standardized Measure of Biological and Psychosocial Function. JAMA 1963 Sep 21;185:914-919. [Medline: 14044222]
12. Cornelis E, Gorus E, Beyer I, Bautmans I, De Vriendt P. Early diagnosis of mild cognitive impairment and mild dementia through basic and instrumental activities of daily living: Development of a new evaluation tool. PLoS Med 2017 Mar;14(3):e1002250 [FREE Full text] [doi: 10.1371/journal.pmed.1002250] [Medline: 28291801]
13. Reuben DB, Seeman TE, Keeler E, Hayes RP, Bowman L, Sewall A, et al. Refining the categorization of physical functional status: the added value of combining self-reported and performance-based measures. J Gerontol A Biol Sci Med Sci 2004 Oct;59(10):1056-1061. [Medline: 15528778]

14. Guralnik JM, Branch LG, Cummings SR, Curb JD. Physical performance measures in aging research. J Gerontol 1989 Sep;44(5):M141-M146. [Medline: 2768767]
15. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [Medline: 8126356]
16. Myers AM, Holliday PJ, Harvey KA, Hutchinson KS. Functional performance measures: are they superior to self-assessments? J Gerontol 1993 Sep;48(5):M196-M206. [Medline: 8366262]
17. Chmelo E, Nicklas B, Davis C, Miller GD, Legault C, Messier S. Physical activity and physical function in older adults with knee osteoarthritis. J Phys Act Health 2013 Aug;10(6):777-783 [FREE Full text] [Medline: 23307503]
18. Dunlop DD, Semanik P, Song J, Sharma L, Nevitt M, Jackson R, Osteoarthritis Initiative Investigators. Moving to maintain function in knee osteoarthritis: evidence from the osteoarthritis initiative. Arch Phys Med Rehabil 2010 May;91(5):714-721 [FREE Full text] [doi: 10.1016/j.apmr.2010.01.015] [Medline: 20434608]
19. Schlenk EA, Lias JL, Sereika SM, Dunbar-Jacob J, Kwoh CK. Improving physical activity and function in overweight and obese older adults with osteoarthritis of the knee: a feasibility study. Rehabil Nurs 2011;36(1):32-42 [FREE Full text] [Medline: 21290963]
20. Smuck M, Tomkins-Lane C, Ith MA, Jarosz R, Kao MCJ. Physical performance analysis: A new approach to assessing free-living physical activity in musculoskeletal pain and mobility-limited populations. PLoS One 2017;12(2):e0172804 [FREE Full text] [doi: 10.1371/journal.pone.0172804] [Medline: 28235039]
21. Thomas SG, Pagura SMC, Kennedy D. Physical activity and its relationship to physical performance in patients with end stage knee osteoarthritis. J Orthop Sports Phys Ther 2003 Dec;33(12):745-754. [doi: 10.2519/jospt.2003.33.12.745] [Medline: 14743988]
22. Berkemeyer K, Wijndaele K, White T, Cooper AJM, Luben R, Westgate K, et al. The descriptive epidemiology of accelerometer-measured physical activity in older adults. Int J Behav Nutr Phys Act 2016 Jan 07;13:2 [FREE Full text] [doi: 10.1186/s12966-015-0316-z] [Medline: 26739758]
23. Parker SJ, Strath SJ, Swartz AM. Physical activity measurement in older adults: relationships with mental health. J Aging Phys Act 2008 Oct;16(4):369-380 [FREE Full text] [Medline: 19033599]
24. Kahn TL, Schwarzkopf R. Does Total Knee Arthroplasty Affect Physical Activity Levels? Data from the Osteoarthritis Initiative. J Arthroplasty 2015 Sep;30(9):1521-1525. [doi: 10.1016/j.arth.2015.03.016] [Medline: 25882608]
25. Conway J, Tomkins CC, Haig AJ. Walking assessment in people with lumbar spinal stenosis: capacity, performance, and self-report measures. Spine J 2011 Sep;11(9):816-823 [FREE Full text] [doi: 10.1016/j.spinee.2010.10.019] [Medline: 21145292]
26. Agarwal V, Smuck M, Shah NH. Quantifying the relative change in physical activity after Total Knee Arthroplasty using accelerometer based measurements. In: AMIA Jt Summits Transl Sci Proc. 2017 Presented at: AMIA Joint Summits on Translational Science; March 27-30, 2017; San Francisco p. 463-472.
27. Smuck M, Muaremi A, Zheng P, Norden J, Sinha A, Hu R, et al. Objective measurement of function following lumbar spinal stenosis decompression reveals improved functional capacity with stagnant real-life physical activity. Spine J 2018 Dec;18(1):15-21. [doi: 10.1016/j.spinee.2017.08.262] [Medline: 28962914]
28. Yorston LC, Kolt GS, Rosenkranz RR. Physical activity and physical function in older adults: the 45 and up study. J Am Geriatr Soc 2012 Apr;60(4):719-725. [doi: 10.1111/j.1532-5415.2012.03906.x] [Medline: 22486736]
29. van Lummel RC, Walgaard S, Pijnappels M, Elders PJM, Garcia-Aymerich J, van DJH, et al. Physical Performance and Physical Activity in Older Adults: Associated but Separate Domains of Physical Function in Old Age. PLoS One 2015;10(12):e0144048 [FREE Full text] [doi: 10.1371/journal.pone.0144048] [Medline: 26630268]
30. Willetts M, Hollowell S, Aslett L, Holmes C, Doherty A. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. Sci Rep 2018 May 21;8(1):7961 [FREE Full text] [doi: 10.1038/s41598-018-26174-1] [Medline: 29784928]
31. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: A systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]
32. DasMahapatra P, Chiauzzi E, Bhalerao R, Rhodes J. Free-Living Physical Activity Monitoring in Adult US Patients with Multiple Sclerosis Using a Consumer Wearable Device. Digit Biomark 2018 Apr 13;2(1):47-63. [doi: 10.1159/000488040]
33. Paxton RJ, Melanson EL, Stevens-Lapsley JE, Christiansen CL. Physical activity after total knee arthroplasty: A critical review. World J Orthop 2015 Sep 18;6(8):614-622 [FREE Full text] [doi: 10.5312/wjo.v6.i8.614] [Medline: 26396937]
34. Foerster F, Smeja M, Fahrenberg J. Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. Computers in Human Behavior 1999 Sep;15(5):571-583. [doi: 10.1016/S0747-5632(99)00037-0]
35. van Kasteren TLM, Englebienne G, Kröse BJA. An activity monitoring system for elderly care using generative and discriminative models. Pers Ubiquit Comput 2010 Feb 26;14(6):489-498. [doi: 10.1007/s00779-009-0277-9]
36. Lara OD, Labrador MA. A Survey on Human Activity Recognition using Wearable Sensors. IEEE Commun. Surv. Tutorials 2013;15(3):1192-1209. [doi: 10.1109/SURV.2012.110112.00192]

37. Arif M, Kattan A. Physical Activities Monitoring Using Wearable Acceleration Sensors Attached to the Body. PLoS One 2015;10(7):e0130851 [FREE Full text] [doi: 10.1371/journal.pone.0130851] [Medline: 26203909]
38. Grimm B, Bolink S. Evaluating physical function and activity in the elderly patient using wearable motion sensors. British Editorial Society of Bone and Joint Surgery 2016 May;1(5):112-120 [FREE Full text] [doi: 10.1302/2058-5241.1.160022] [Medline: 28461937]
39. Gresham G, Hendifar AE, Spiegel B, Neeman E, Tuli R, Rimel BJ, et al. Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients. npj Digital Med 2018 Jul 5;1(1):27. [doi: 10.1038/s41746-018-0032-6]
40. Merilahti J, Viramo P, Korhonen I. Wearable Monitoring of Physical Functioning and Disability Changes, Circadian Rhythms and Sleep Patterns in Nursing Home Residents. IEEE J Biomed Health Inform 2016 Dec;20(3):856-864. [doi: 10.1109/JBHI.2015.2420680] [Medline: 25861091]
41. Eston RG, Rowlands AV, Ingledew DK. Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. J Appl Physiol (1985) 1998 Jan;84(1):362-371 [FREE Full text] [doi: 10.1152/jappl.1998.84.1.362] [Medline: 9451658]
42. Janz KF. Validation of the CSA accelerometer for assessing children's physical activity. Med Sci Sports Exerc 1994 Mar;26(3):369-375. [Medline: 8183103]
43. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? Med Sci Sports Exerc 2000 Feb;32(2):426-431. [Medline: 10694127]
44. Song J, Semanik P, Sharma L, Chang RW, Hochberg MC, Mysiw WJ, et al. Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010 Dec;62(12):1724-1732 [FREE Full text] [doi: 10.1002/acr.20305] [Medline: 20806273]
45. Dobson F, Hinman RS, Roos EM, Abbott JH, Stratford P, Davis AM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthritis Cartilage 2013 Aug;21(8):1042-1052 [FREE Full text] [doi: 10.1016/j.joca.2013.05.002] [Medline: 23680877]
46. Newman AB, Simonsick EM, Naydeck BL, Boudreau RM, Kritchevsky SB, Nevitt MC, et al. Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability. JAMA 2006 May 03;295(17):2018-2026. [doi: 10.1001/jama.295.17.2018] [Medline: 16670410]
47. Bohannon RW. Test-retest reliability of the five-repetition sit-to-stand test: a systematic review of the literature involving adults. J Strength Cond Res 2011 Nov;25(11):3205-3207. [doi: 10.1519/JSC.0b013e318234e59f] [Medline: 21904240]
48. James N, Matteson D. ecp : An R Package for Nonparametric Multiple Change Point Analysis of Multivariate Data. In: J. Stat. Soft. Innsbruck: An R Package for Nonparametric Multiple Change Point Analysis of Multivariate Data. 2013; 2014:1-31.
49. Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of novel techniques to classify physical activity mode using accelerometers. Med Sci Sports Exerc 2006 Sep;38(9):1626-1634. [doi: 10.1249/01.mss.0000227542.43669.45] [Medline: 16960524]
50. Chan BKS, Marshall LM, Winters KM, Faulkner KA, Schwartz AV, Orwoll ES. Incident fall risk and physical activity and physical performance among older men: the Osteoporotic Fractures in Men Study. Am J Epidemiol 2007 Mar 15;165(6):696-703. [doi: 10.1093/aje/kwk050] [Medline: 17194749]
51. Ekelund LG, Haskell WL, Johnson JL, Whaley FS, Criqui MH, Sheps DS. Physical fitness as a predictor of cardiovascular mortality in asymptomatic North American men. The Lipid Research Clinics Mortality Follow-up Study. N Engl J Med 1988 Nov 24;319(21):1379-1384. [doi: 10.1056/NEJM198811243192104] [Medline: 3185648]
52. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York: Springer; 2009.
53. Agresti A. An Introduction to Categorical Data Analysis. Hoboken, NJ: Wiley, John & Sons, Incorporated; 2007.
54. Wood S, Goude Y, Shaw S. Generalized additive models for large data sets. J. R. Stat. Soc. C 2014 May 27;64(1):139-155. [doi: 10.1111/rssc.12068]
55. Washburn RA, McAuley E, Katula J, Mihalko SL, Boileau RA. The physical activity scale for the elderly (PASE): evidence for validity. J Clin Epidemiol 1999 Jul;52(7):643-651. [Medline: 10391658]
56. Brouwer GM, van Tol AW, Bergink AP, Belo JN, Bernsen RMD, Reijman M, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. Arthritis Rheum 2007 Apr;56(4):1204-1211 [FREE Full text] [doi: 10.1002/art.22515] [Medline: 17393449]
57. Cerejo R, Dunlop DD, Cahue S, Channin D, Song J, Sharma L. The influence of alignment on risk of knee osteoarthritis progression according to baseline stage of disease. Arthritis Rheum 2002 Oct;46(10):2632-2636 [FREE Full text] [doi: 10.1002/art.10530] [Medline: 12384921]
58. Felson DT, Niu J, Gross KD, Englund M, Sharma L, Cooke TDV, et al. Valgus malalignment is a risk factor for lateral knee osteoarthritis incidence and progression: findings from the Multicenter Osteoarthritis Study and the Osteoarthritis Initiative. Arthritis Rheum 2013 Feb;65(2):355-362 [FREE Full text] [doi: 10.1002/art.37726] [Medline: 23203672]
59. Sharma L, Hayes KW, Felson DT, Buchanan TS, Kirwan-Mellis G, Lou C, et al. Does laxity alter the relationship between strength and physical function in knee osteoarthritis? Arthritis Rheum 1999 Jan;42(1):25-32 [FREE Full text] [doi: 10.1002/1529-0131(199901)42:1<25::AID-ANR3>3.0.CO;2-G] [Medline: 9920010]

60. Mikesky AE, Mazzuca SA, Brandt KD, Perkins SM, Damush T, Lane KA. Effects of strength training on the incidence and progression of knee osteoarthritis. Arthritis Rheum 2006 Oct 15;55(5):690-699 [FREE Full text] [doi: 10.1002/art.22245] [Medline: 17013851]
61. Han ASY, Nairn L, Harmer AR, Crosbie J, March L, Parker D, et al. Early rehabilitation after total knee replacement surgery: a multicenter, noninferiority, randomized clinical trial comparing a home exercise program with usual outpatient care. Arthritis Care Res (Hoboken) 2015 Feb;67(2):196-202 [FREE Full text] [doi: 10.1002/acr.22457] [Medline: 25220488]
62. Mannini A, Sabatini AM. Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors (Basel) 2010;10(2):1154-1175 [FREE Full text] [doi: 10.3390/s100201154] [Medline: 22205862]

Abbreviations

ADL: activities of daily living
GAM: Generalized Additive Models
OA: osteoarthritis
OAI: Osteoarthritis Initiative
PASE: Physical Activity Scale for the Elderly
WAM: wearable activity monitor

Edited by G Eysenbach; submitted 18.06.18; peer-reviewed by I Clay, WN Lee, E Yom-Tov; comments to author 26.07.18; revised version received 20.09.18; accepted 01.10.18; published 18.12.18.

Please cite as:
Agarwal V, Smuck M, Tomkins-Lane C, Shah NH
Inferring Physical Function From Wearable Activity Monitors: Analysis of Free-Living Activity Data From Patients With Knee Osteoarthritis
JMIR Mhealth Uhealth 2018;6(12):e11315
URL: http://mhealth.jmir.org/2018/12/e11315/
doi:10.2196/11315
PMID:30394876

Original Paper

Adherence to a Mindfulness and Relaxation Self-Care App for Cancer Patients: Mixed-Methods Feasibility Study

Michael Mikolasek[1], MSc; Claudia M Witt[1,2,3], MBA, MD; Jürgen Barth[1], PhD

[1]Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland

[2]Institute for Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, Berlin, Germany

[3]Center for Integrative Medicine, University of Maryland School of Medicine, Baltimore, MD, United States

Corresponding Author:
Michael Mikolasek, MSc
Institute for Complementary and Integrative Medicine
University Hospital Zurich and University of Zurich
Sonneggstrasse 6
Zurich, 8091
Switzerland
Phone: 41 44 255 51 49
Fax: 41 44 255 43 94
Email: michael.mikolasek@usz.ch

Abstract

Background: Cancer is highly prevalent worldwide and can cause high levels of distress in patients, which is often neglected in medical care. Smartphone apps are readily available and therefore seem promising to deliver distress-reducing interventions such as mindfulness and relaxation programs.

Objective: This study aimed to evaluate the feasibility of a mindfulness and relaxation app for cancer patients. We looked at characteristics of participating patients in a mobile health (mHealth) study, including adherence to the app intervention, predictors for adherence, and patients' feedback regarding the app.

Methods: In this prospective observational study with a mixed-methods approach, cancer patients received a mindfulness and relaxation self-care app. Cancer patients were recruited online and through hospitals in Switzerland. We assessed self-reported measures (eg, quality of life, anxiety, depressive symptoms, openness to experience, resistance to change) at baseline, and the app gathered data on patients' practicing time. With 8 semistructured interviews, we obtained patients' feedback about the app and recommendations for improvements. We looked at 3 dimensions of the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework (reach, adoption, and maintenance) and analyzed data for adherence for the first 10 weeks of the app intervention. We report descriptive statistics for patient characteristics and app use. For the prediction of adherence, we used Kaplan-Meier analyses with log-rank tests and a Cox proportional hazards regression.

Results: Data from 100 cancer patients (74 female) showed that 54 patients were using the app exercises continuously until week 10. In continuous app users, the median number of exercises per week dropped from 4 (interquartile range, IQR 1-7) at week 1 to a median of 2 (IQR 1-4) at week 10. Our analyses revealed 4 significant predictors for better adherence: female gender, higher openness to experience, higher resistance to change, and more depressive symptoms. Interviews revealed that the patients generally were satisfied with the app but also made suggestions on how to improve it.

Conclusions: Our study indicates that a mindfulness and relaxation mHealth intervention for cancer patients is feasible with acceptable adherence and largely positive feedback from patients.

Trial Registration: German Clinical Trials Register DRKS00010481; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00010481 (Archived by WebCite at http://www.webcitation.org/73xGE1B0P)

(JMIR Mhealth Uhealth 2018;6(12):e11271) doi:10.2196/11271

KEYWORDS

mobile app; mindfulness; relaxation; cancer; patient compliance

Introduction

Background

Cancer is highly prevalent worldwide, with an estimated 14 million newly diagnosed patients per year [1]. According to the World Health Organization, cancer is the second leading cause of death, with an increasing economic impact over recent years [2]. For patients, the diagnosis of cancer and subsequent treatment (eg, radiation or chemotherapy) can cause high levels of distress [3,4]. About every second cancer patient has clinically relevant distress, with elevated levels of depression or anxiety [5]. However, psychological support of patients is often not implemented in standard medical care [6-8]. In addition, many patients neglect their distress and do not seek psychosocial support [9]. However, untreated distress can reduce quality of life as well as lower adherence with recommended medical care and, therefore, negatively affect patients' recovery [8,10]. Thus, a variety of treatments such as counseling and Mind Body Medicine (MBM) interventions have been suggested to reduce cancer patients' distress during initial care and rehabilitation [11-13].

MBM focuses on the interactions between psychological and biological processes and their impact on health [14,15] and has shown beneficial effects in reducing cancer patients' distress [13,16]. MBM usually combines a variety of interventions, such as exercise, Qigong, relaxation, and mindfulness meditation [14]. Some of these interventions, for example, mindfulness and relaxation, are also commonly used on their own and have been studied extensively with promising effects in both healthy and patient populations [15,17-20]. In addition, an increasing number of cancer patients are interested in or use mindfulness or relaxation interventions [21].

Regular practice is crucial for the effect of mindfulness and relaxation-based interventions, which can be difficult to achieve due to lack of motivation, time constraints, as well as limited access to interventions [22]. Further restrictions for regular practice and access-limiting factors include geographical distance, financial constraints, lack of treatment providers or lack of knowledge thereof [8,9,23]. For cancer patients, regular practice might additionally be hindered due to restrictions caused by cancer (eg, fatigue and nausea) and its comprehensive treatments.

Mobile health (mHealth) interventions might overcome some of the restrictions of face-to-face interventions. The access to interventions can be easier due to a large and increasing number of smartphone owners [24]. In 2017, more than 32% of the world population and more than 60% of the population in Western Europe and North America owned a smartphone [25]. In addition, mHealth interventions have some specific advantages compared with face-to-face interventions. These advantages include easy and pervasive access to information (ie, psychoeducation), engaging audio and/or visual material, potential customization of the app according to client's preferences and needs, provision of regular feedback, reminders, and reduced perceived stigmatization due to potentially less therapist contact [24,26]. mHealth interventions can also be a good support for patients' self-care [26]. Such self-care interventions can have beneficial effects on cancer patients' distress and quality of life [27] and can be implemented via an app using audio instructions.

To date, mHealth interventions using a mindfulness or relaxation intervention strategy have been under-researched, with the focus of studies primarily on Web-based electronic Health (eHealth) interventions [28]. For eHealth interventions, studies indicate that mindfulness- and relaxation-based interventions can have beneficial effects on health outcomes in various populations, including cancer patients [28-30]. Beneficial effects of eHealth were reported for stress, well-being, anxiety, depression, and mindfulness. The majority of available primary studies in these reviews focused on eHealth interventions, with a partial emphasis on Web-based patient-therapist interactions. However, less is known about the feasibility and effectiveness of mHealth interventions, and certain disadvantages (eg, technical problems, concerns about data security) are well known [26]. Eysenbach [31] coined the term "law of attrition," which emphasizes that early and rapid attrition rates are an inherent problem in technology-delivered interventions. Especially in self-care interventions with regular exercises, good adherence itself often becomes an intervention goal. Although recent eHealth studies report acceptable rates of adherence (eg, 60% completed 4 or more out of 6 modules [32] and 71% practiced more than 50% of the days during 8 weeks [33]), little is known about the adherence to mindfulness and relaxation mHealth programs for cancer patients. Therefore, when setting up a self-care mHealth intervention, it is important to know which factors might influence the patient adherence.

Objective

The aim of this mHealth study, then, was to evaluate the feasibility of a mindfulness and relaxation app for cancer patients and its impact on health outcomes according to the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework [34]. In this analysis, we looked at the characteristics of patients who participated in this mHealth study, adherence and predictors for adherence, as well as patients' feedback regarding the mHealth intervention from interviews.

Methods

Study Design

We performed a prospective observational study using a mixed-methods approach. Quantitative data consisted of 4 paper-and-pencil questionnaires sent to cancer patients at baseline, weeks 4, 10, and 20. Qualitative data consisted of semistructured interviews with 8 cancer patients. Corresponding to the principles of theoretical sampling [35], we recruited the interviewees based on the sample distributions of gender and intervention dropouts versus continuous app users. We conducted individual qualitative interviews over the telephone with these patients; selecting 4 of them who used the app on a regular basis and 4 who did not use the app regularly. We conducted qualitative interviews with these patients; 4 of them used the app on a regular basis and 4 did not use the app regularly. This study was guided by the RE-AIM evaluation framework [34], which consists of the following 5 dimensions:

reach, effectiveness, adoption, implementation, and maintenance. For this analysis, we focused on 3 dimensions, namely reach, adoption, and maintenance during the first 10 weeks of the intervention. The dimensions effectiveness and implementation as well as results about the entire 20 weeks will be reported in an upcoming paper. The cantonal ethics committee granted ethical approval for the study (BASEC-Nr. 2016-00258) in April 2016, and the study was positively audited within the regular ICH-GCP audit of the University Hospital Zurich in August 2016. We registered the study in the German Clinical Trials Register (DRKS00010481).

Eligibility Criteria

We included female and male cancer patients (18 years or older) with any cancer diagnosis at any stage of cancer, who owned either an iOS- or Android-based smartphone with at least a weekly connection to the internet. We excluded patients if they had suicidal ideation or insufficient German language skills. Furthermore, patients who intended to move to another country and patients with insufficient knowledge on how to use a smartphone were excluded.

Recruitment

We recruited cancer patients in 2 different settings: (1) cancer patients who participated in a supportive MBM treatment (either as individual session or as a 10-week group treatment) or (2) cancer patients without an MBM treatment.

For setting 1, cancer patients were recruited at the Institute for Complementary and Integrative Medicine, University Hospital Zurich (ICI). All available cancer patients in an MBM group treatment (between June 2016 and December 2017; 12 groups with a total of 81 patients) were invited at the third session of the course to participate in the study. Therefore, enrolled patients from setting 1 were using the app partially in parallel to the MBM group treatment. In addition, we asked the health professionals of the ICI to distribute leaflets during individual MBM consultations with cancer patients.

For setting 2, patients were recruited through the University Hospital Zurich (ICI, cancer center, Department of Radiotherapy, and University Hospital Facebook page), University Hospital Basel, and the Cantonal Hospital Aarau. Cancer patients were informed of the study using leaflets in the waiting areas or during consultations. In addition, we informed cancer patients through the Swiss Cancer League via leaflets and their Facebook page, as well as through the Cancer League of Zurich.

Interested patients initially contacted a researcher at the ICI by phone or email and made an appointment for a 10-min telephone screening interview. During the screening interview, the researcher explained the study and assessed the eligibility of the patient. In addition, the researcher recommended that the patient carry out 1 of the 3 exercises of the app at least once a day on 5 different days per week during the 20-week intervention. However, the researcher also stated that patients were free to choose when, where, and how often they practiced. After the researcher provided all information and if the patient met the eligibility criteria, the researcher asked for contact details of the patient. Subsequently, the patient received the written study information with the informed consent form, as well as the first questionnaire by mail. We sent an email to every included patient with a code to activate the app. Thereafter, patients were able to use the app free of charge. The date of the code distribution was considered as the start of the intervention for each patient. No other verbal contact between the researcher and the patients took place after inclusion of patients.

Intervention

The mindfulness and relaxation app comprised 2 main features: (1) mindfulness and relaxation exercises guided by audio instructions and (2) a notification feature. The first feature of the app contained 3 exercises and was the main component of the app. The exercises were mindfulness meditation, guided imagery, and progressive muscle relaxation audio files with a duration of about 15 min each. Every exercise was guided by a narrator with either a male or female voice.

The second feature of the app was a notification feature, which reminded the patient to practice daily. The patient could set the time of notification according to individual preferences. The reminder to practice popped up as a push notification on the mobile device every day at the time set by the patient. The concept of the app built on previously developed relaxation study apps of an affiliated group [36], which were designed for patients with chronic low back pain (Relaxback) and chronic neck pain (Relaxneck).The app was developed by the software company Smart Mobile Factory GmbH (Berlin, Germany). After thorough testing, the app was released in June 2016 on the Apple iTunes Store and on the Google Play Store for Android devices. After the release, the content of the app was not changed. Screenshots of the app are available in Multimedia Appendices 1 and 2.

Outcomes

Reach

For the dimension reach, we looked into which and how many cancer patients participated in the study. We present baseline characteristics to describe participating patients: type of cancer, status of cancer treatment, sociodemographic data (gender, age, and highest education), distress (Distress-Thermometer [37,38]), quality of life (FACT-G, Functional Assessment of Cancer Therapy-General [39]), and anxiety and depression (HADS, Hospital Anxiety Depression Scale [40]).

The Distress-Thermometer consists of 1 item with a scale from 0 to 10 and assesses experienced distress in the past week. A score between 5 or higher is considered as clinically relevant distress [41]. The FACT-G consists of 27 items, which assess the 4 subscales: physical well-being (Cronbach alpha=.851), social well-being (Cronbach alpha=.760), emotional well-being (Cronbach alpha=.702), and functional well-being (Cronbach alpha=.794). Each item is rated on a 5-point scale (0-4), resulting in a score range of 0 to 108, with a higher score indicating a better quality of life. The HADS consists of 14 items, with 7 items for each subscale, that is, anxiety (Cronbach alpha=.787) and depression (Cronbach alpha=.667). Each item is rated on a 4-point scale (0-3), leading to a maximum score of 21 for each subscale. A score between 0 and 7 is considered normal, whereas

a score between 8 and 11 is considered as borderline, and a score above 11 as caseness.

Adoption

For the dimension adoption, we looked at indicators of patients' adoption of the app intervention into their regular life, adherence, and information about barriers and facilitators for regular use. For this purpose, we analyzed the use of the app during the first 10 weeks. We derived app use data from tracking the practicing time with the audios (start and end time and type of exercise used). This information was visible only for the research team (as an XML log file through the backend) and was not displayed to users. In addition, we analyzed results from interviews with patients regarding their adoption of the app intervention.

As a first indicator for app intervention adoption, we report the number of completed app exercises per week. We considered an exercise as completed if the patient used the exercise for at least 10 min (out of 15 min). As a second indicator, we report the number of intervention dropouts versus number of continuous app users per week. Intervention dropouts were defined as enrolled patients who never completed an exercise or did not complete an exercise during 4 consecutive weeks after initial practice. A patient counted as an intervention dropout in the first of the 4 weeks, in which he or she did not complete any exercise. According to our definition, a patient who never completed an exercise is an intervention dropout at week 1. Patients not classified as intervention dropouts were defined as continuous app users. Consequently, continuous app use was defined as at least weekly use of 1 or more app exercises. We also report results from 8 semistructured patient interviews, in which we inquired about patients' general impression regarding the app, app usage, and suggestions for improvements (for interview guideline, see Multimedia Appendix 3).

Maintenance

For the dimension maintenance, we looked into predictors for continuous app use. First, we assumed that patients with higher openness to experience are more often continuous app users. Second, we assumed that patients with higher resistance to change are less often continuous app users. In addition, we tested in explorative analyses if quality of life (FACT-G), anxiety (HADS anxiety), depression (HADS depression) at baseline and sociodemographic data (gender and age), as well as setting are associated with continuous app use. During the interviews, we also explored possible reasons for continuous app use and intervention dropout.

We measured openness to experience with the respective subscale of the NEO 5-Factor Inventory (NEO-FFI [42]) using the 5-item short version (Cronbach alpha=.755). Each item is rated on a 4-point scale (0-4), leading to a score with a range from 0 to 20. A higher score indicates greater openness to experience. We also used the Resistance to Change (RTC) Scale [43], which consists of 17 items (Cronbach alpha=.839). Each item is rated on a 6-point scale (1-6), resulting in a score with a range from 17 to 102. A higher score indicates greater resistance to change.

Sample Size

In this feasibility study, the sample size is an outcome in itself (ie, dimension reach in the evaluation framework). Therefore, we did not perform an a priori sample size calculation, but the aim was to recruit about 100 patients to conduct explorative analyses about the feasibility of the app with sufficient precision.

Data Analysis

Quantitative Data

Trained researchers entered data from printed case report forms using REDCap electronic data capture tools [44] hosted at the University Hospital Zurich. Analyses were conducted using SPSS version 25.0 (IBM Corp, Armonk, NY, USA) [45].

For the reach analyses, we used descriptive statistics (frequencies and percentages for categorical and dichotomous variables, mean and SD for continuous variables) for baseline data on sociodemographic characteristics (gender, age, and education), health status (type of cancer, status of cancer treatment, FACT-G, HADS, and Distress-Thermometer), and the setting of the enrolled patients. For the adoption analyses, we used boxplots to report median and interquartile range (IQR) of the number of completed exercises per week (week 1 to 10) for all enrolled patients, as well as for continuous app users during the 10-week intervention. In addition, we used a Kaplan-Meier plot to visualize the number of dropouts per week. For the maintenance analyses, we used Kaplan-Meier analyses with a log-rank test to compare continuous app users (ie, reversed rate of attrition) according to different baseline variables. As predictors, we used the following categorical variables: gender, setting, age groups (18-40, 41-55, 56+), high versus low well-being (FACT-G median split at 76.83), high versus low openness to experience (NEO-FFI-O median split at 17.00), high versus low resistance to change (RTC median split at 51.00), normal versus suggestive or higher anxiety or depression (HADS anxiety or depression scores of 0-7 vs 8 or higher). Subsequently, we performed a Cox proportional hazards regression with all significant predictors in the log-rank test in the Kaplan-Meier analyses.

For missing data, we used multiple imputation to conduct the Cox proportional hazards regression with a full dataset. We carried out imputations for the sum scores of FACT-G, as missing single items are already considered in the calculation of FACT-G sum scores (FACT-G sum scores are not calculated if there are more than 50% items missing in a subscale). For HADS, NEO-FFI-O, and RTC, we imputed all items with 1 or more missing values. For all other analyses (ie, descriptive analyses for the dimensions reach and adoption, Kaplan-Meier analyses for the dimension maintenance), we used complete datasets.

Qualitative Data

For the interview analyses about the adoption and maintenance of the intervention, we transcribed the recorded interviews verbatim and used thematic coding for structuring the interviews using MAXQDA 11 (VERBI Software, Berlin, Germany). Thereafter, we used content analysis according to Mayring [46].

Results

Reach

During the recruitment phase between June 2016 and December 2017, a total of 118 patients expressed interest in participating in the study and were screened for eligibility. All of the 118 patients fulfilled the eligibility criteria and received the informed consent form. By the end of December 2017, 100 patients signed and returned the informed consent form and were enrolled in the study (see Figure 1).

Baseline characteristics of all enrolled patients (N=100), as well as of continuous app users (54/100, 54%) and intervention dropouts (46/100, 46%) are presented in Table 1. The majority of patients (83/100, 83%) were recruited independent of the MBM treatment (setting 2). Patients were 74% (74/100) female, and the mean age of all patients was 53.24 (SD 11.55) with a range of 23 to 84 years. The most common diagnosis was breast cancer (39/100, 39%). The majority of participants had completed higher education, whereas 41% (41/100) had completed secondary education and 33% (33/100) had obtained a university degree. The Distress-Thermometer indicated that the enrolled patients reported, on average, elevated and clinically relevant distress levels. The HADS scores indicated that the enrolled patients had, on average, normal scores of anxiety and depressive symptoms. Continuous app users and intervention dropouts differed in their gender, with 85% (46/85) female continuous app users versus 61% (28/61) female intervention dropouts.

Figure 1. Patient flowchart. FMI: Freiburg Mindfulness Inventory; PROMIS 29: Patient Reported Outcomes Measurement Information System 29; NEO-FFI-O: NEO Five-Factor Inventory - openness to experience; FACT-G: Functional Assessment of Cancer Therapy-General; RTC: Resistance to Change; FoP-Q-SF: Fear of Progression Questionnaire - Short Form; HADS: Hospital Anxiety Depression Scale.

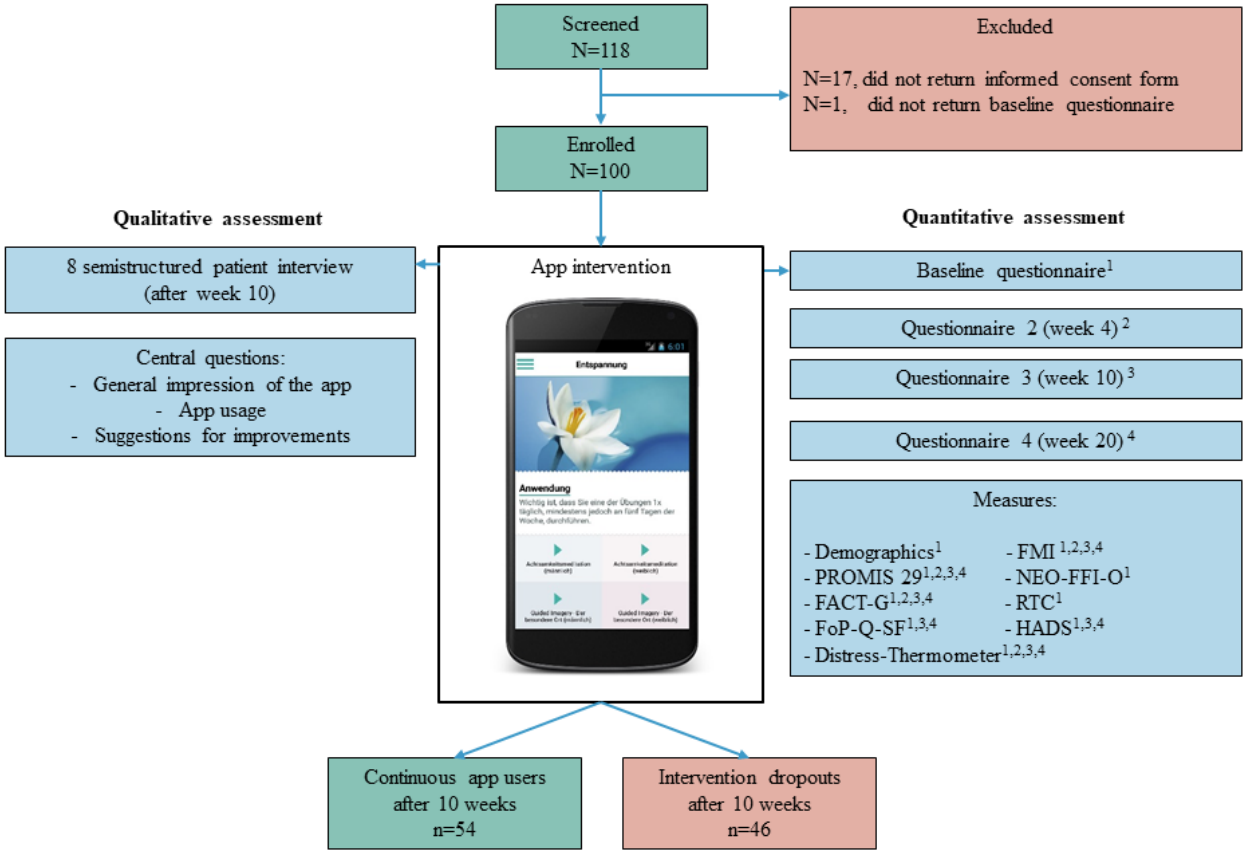


Table 1. Baseline characteristics of all enrolled patients, continuous app users, and intervention dropouts.

Baseline characteristics	Total (N=100)	Continuous app users (n=54)	Intervention dropouts (n=46)
Gender, n (%)			
Female	74 (74)	46 (85)	28 (61)
Male	26 (26)	8 (15)	18 (39)
Age (years), mean (SD)	53.24 (11.55)	54.77 (11.27)	51.45 (11.74)
Type of cancer, n (%)			
Breast cancer	39 (39)	26 (48)	13 (28)
Colon cancer	9 (9)	4 (7)	5 (11)
Ovarian or cervical cancer	6 (6)	6 (11)	0 (0)
Lung cancer	6 (6)	5 (9)	1 (2)
Others	40 (40)	13 (24.)	27 (59)
Status of cancer treatment, n (%)			
Total removal	46 (46)	26 (48)	20 (44)
Recurrence or incomplete removal	25 (25)	12 (22)	13 (28)
Uncertain	3 (3)	2 (4)	1 (2)
Others	26 (26)	14 (26)	12 (26)
Highest education, n (%)			
Primary school	3 (3)	2 (4)	1 (2)
Apprenticeship	22 (22)	10 (19)	12 (26)
Secondary education	41 (41)	24 (44)	17 (37)
University degree	33 (33)	17 (32)	16 (35)
Unknown	1 (1)	0 (0)	0 (0)
Setting, n (%)			
Setting 1[a]	17 (17)	9 (17)	8 (17)
Setting 2[b]	83 (83)	45 (83)	38 (83)
Distress-Thermometer, mean (SD)	5.29 (2.31)	5.36 (2.47)	5.22 (2.14)
FACT-G[c] Quality of life, mean (SD)	75.54 (13.85)	76.56 (14.08)	74.33 (13.63)
HADS[d] anxiety, mean (SD)	6.88 (3.50)	7.17 (3.60)	6.53 (3.38)
HADS depression, mean (SD)	4.96 (2.78)	5.37 (3.05)	4.48 (2.37)

[a]Setting 1: cancer patients with a supportive Mind Body Medicine treatment.

[b]Setting 2: cancer patients without a supportive Mind Body Medicine treatment.

[c]FACT-G: Functional Assessment of Cancer Therapy-General.

[d]HADS: Hospital Anxiety Depression Scale.

Adoption

The number of app exercises completed within the first 10 weeks of the intervention across all patients is presented in Figure 2. During the first week, the median of completed exercises was at 2 with an IQR of 0 to 6, that is, 50% of patients completed 2 or more exercises per week. Over the course of 10 weeks, the median dropped to 0 with an IQR of 0 to 2.5.

The median of app exercises completed across the first 10 weeks of the intervention for continuous app users is presented in Figure 3. During the first week, the median of completed exercises was 4 (IQR 1-7) and dropped down to a median of 2 (IQR 1-4) in week 10.

The Kaplan-Meier survival curve of continuous app users is presented in Figure 4. During the first week, 14/100 (14%) patients never started or stopped using the app exercises on a regular basis and were categorized at week 1 as intervention dropouts. At the end of the intervention, 54/100 (54%) patients were using the app exercises on a regular basis, and between week 1 and week 10, the decline can be regarded as continuous without any specific sensitive weeks to drop out.

Figure 2. Completed app exercises by all patients who were enrolled in the study (N=100) per week (median, interquartile range).

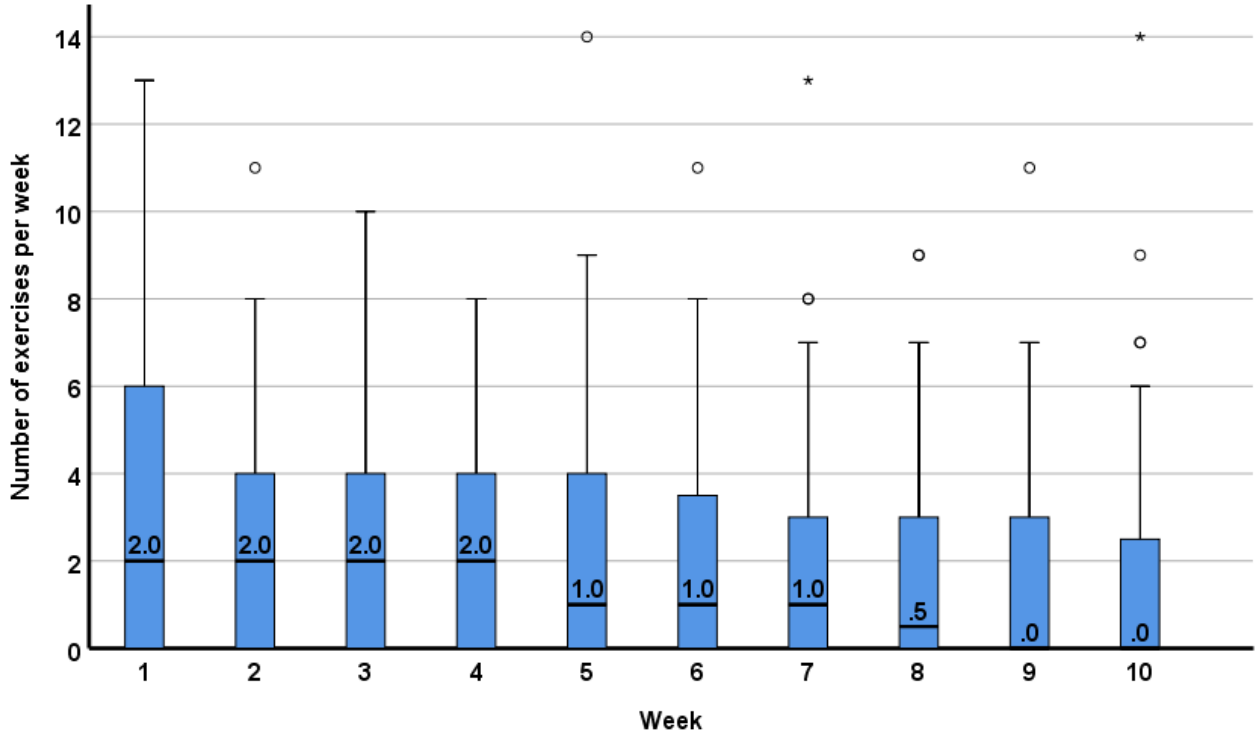


Figure 3. Completed app exercises by continuous app users within a 10-week app intervention (n=54) per week (median, interquartile range).

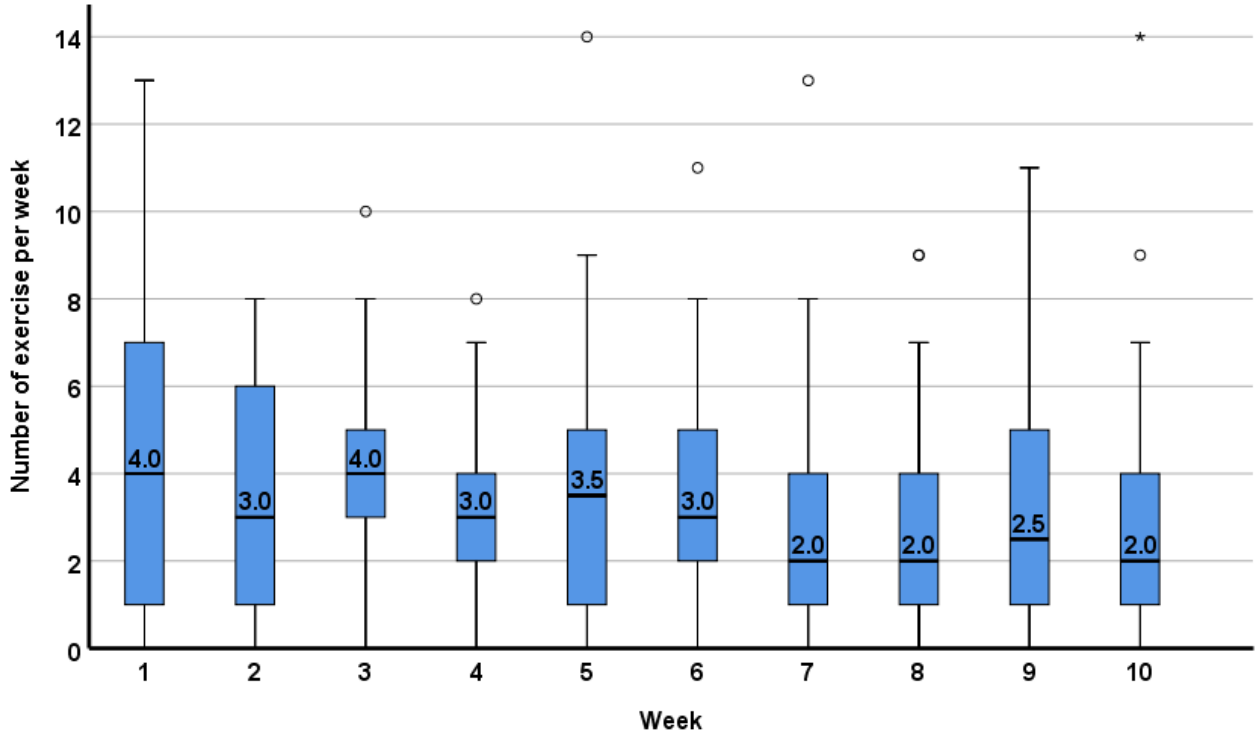
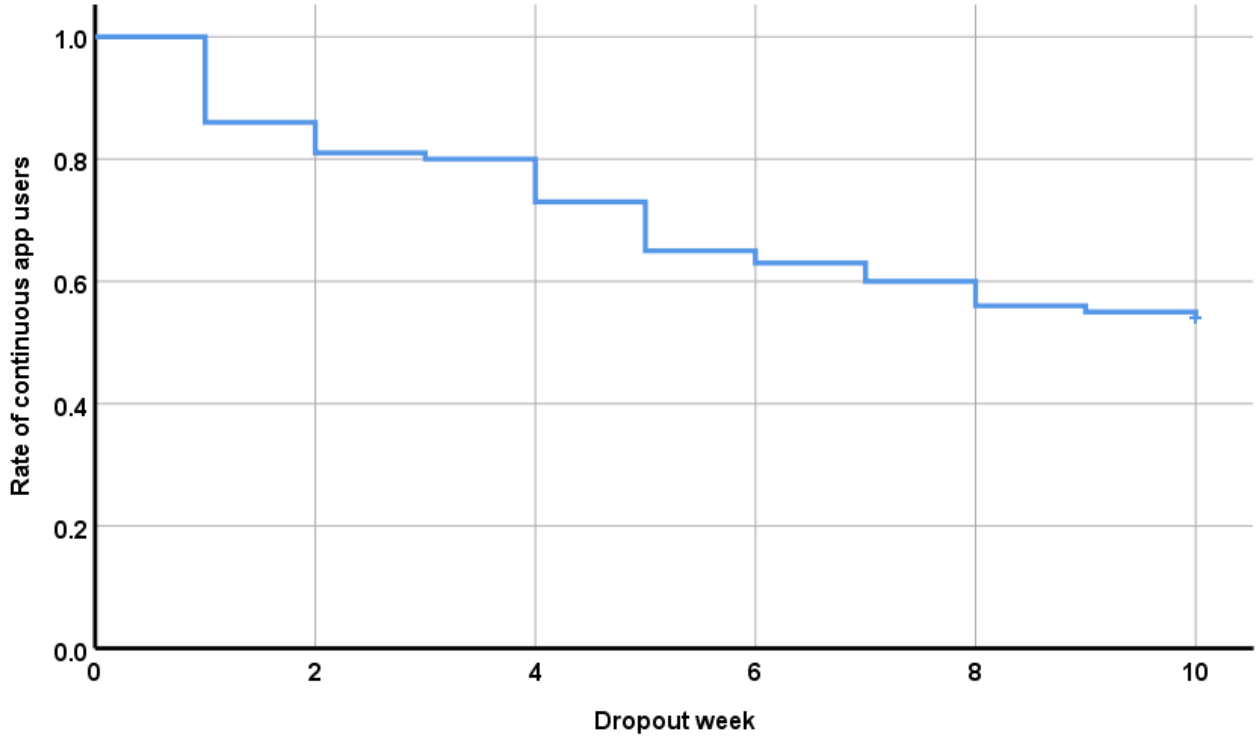


Figure 4. Kaplan-Meier survival curve of all enrolled patients (N=100) over 10 weeks.



Maintenance

The Kaplan-Meier plots for intervention dropouts by gender, setting, age groups, and well-being are presented in Figure 5. The Kaplan-Meier plots for intervention dropouts by openness to experience, resistance to change, anxiety, and depression are presented in Figure 6. Log-rank tests indicated 4 significant predictors for continuous app users, namely gender, openness to experience, resistance to change, and depression.

At week 10, 62% (46/74) of the female patients were still using the app continuously, whereas only 31% (8/26) of the male patients were using the app continuously. Therefore, females had a better adherence to use the app continuously over time than men (P=.005). In the high openness to experience group (NEO-FFI-O), 67% (28/42) of patients still used the app continuously through week 10. In the NEO-FFI-O low openness group, 44% (24/54) used the app continuously through week 10. Thus, patients with high openness to experience had a better adherence than patients with low openness to experience over time (P=.044). In patients with normal HADS depression values, only 49% (39/80) used the app exercises continuously compared with 75% (15/20) in the HADS suggestive or higher depression group (P=.046). In patients with high RTC, 65% (28/43) used the app exercises continuously, but in the low RTC group, only 44% (23/52)of patients used the app exercises continuously through week 10. Therefore, patients in the high RTC group had a better adherence in continuous app use (P=.03). For the factors setting, age groups, well-being (FACT-G), anxiety (HADS anxiety), log-rank tests did not result in significant group differences.

The 4 significant factors of the univariate log-rank test (gender, NEO-FFI-O, RTC, HADS depression) for the prediction of continuous app users went into the multivariate Cox proportional hazards regression. The multivariate analysis indicated solely gender as an independent factor for continuous app use, with an odds ratio (OR) of 2.16 (95% CI 1.09 to 4.27), with a higher chance for attrition in male cancer patients (P=.01). The 3 other factors (NEO-FFI-O, RTC, and HADS depression) did not contribute significantly in this analysis after controlling for gender: high openness to experience was associated with lower odds for attrition (OR 0.96; 95% CI 0.89 to 1.04; P=.30); high RTC with lower odds for attrition (OR 0.98, 95% CI 0.95 to 1.01; P=.17); more depressive symptoms with lower odds for attrition (OR 0.92, 95% CI: 0.80 to 1.03; P=.13).

Figure 5. Kaplan-Meier survival curves for continuous app users by gender, setting, age groups and well-being. FACT-G: Functional Assessment of Cancer Therapy-General.

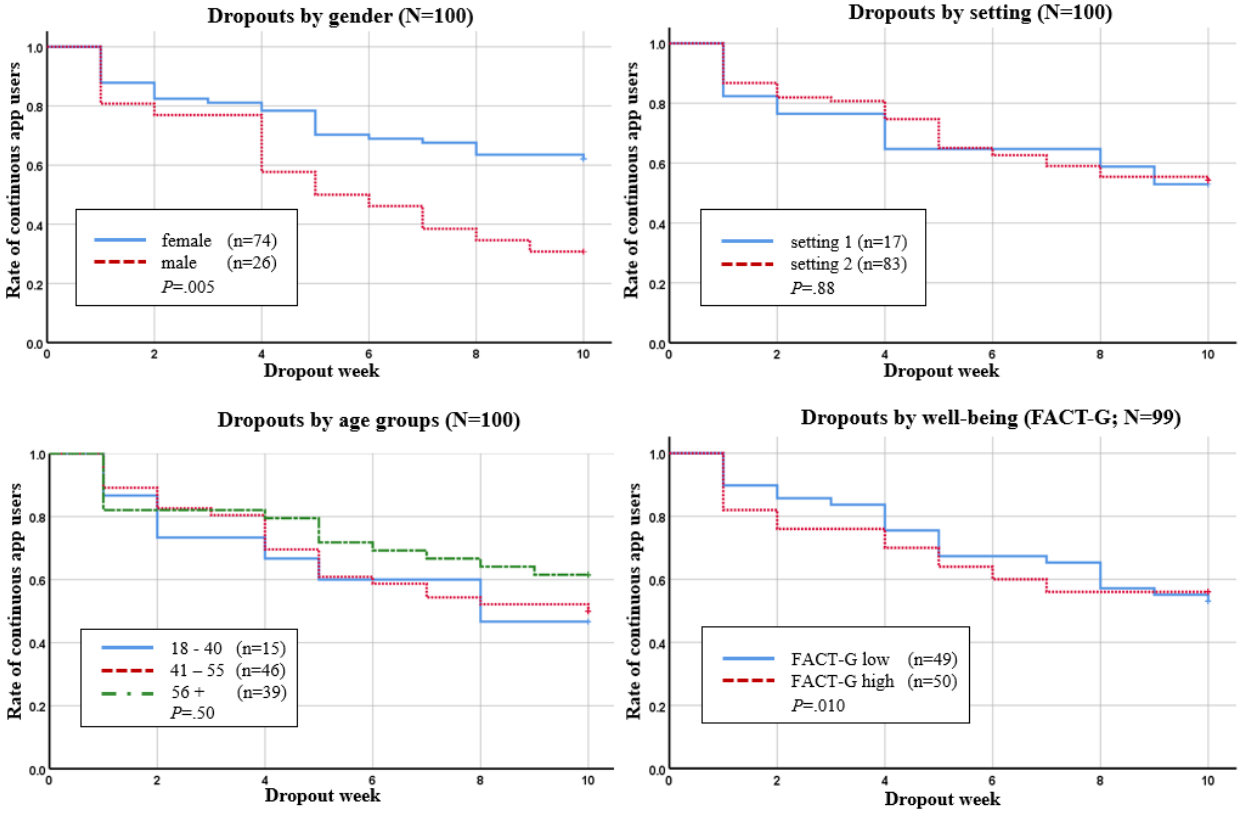
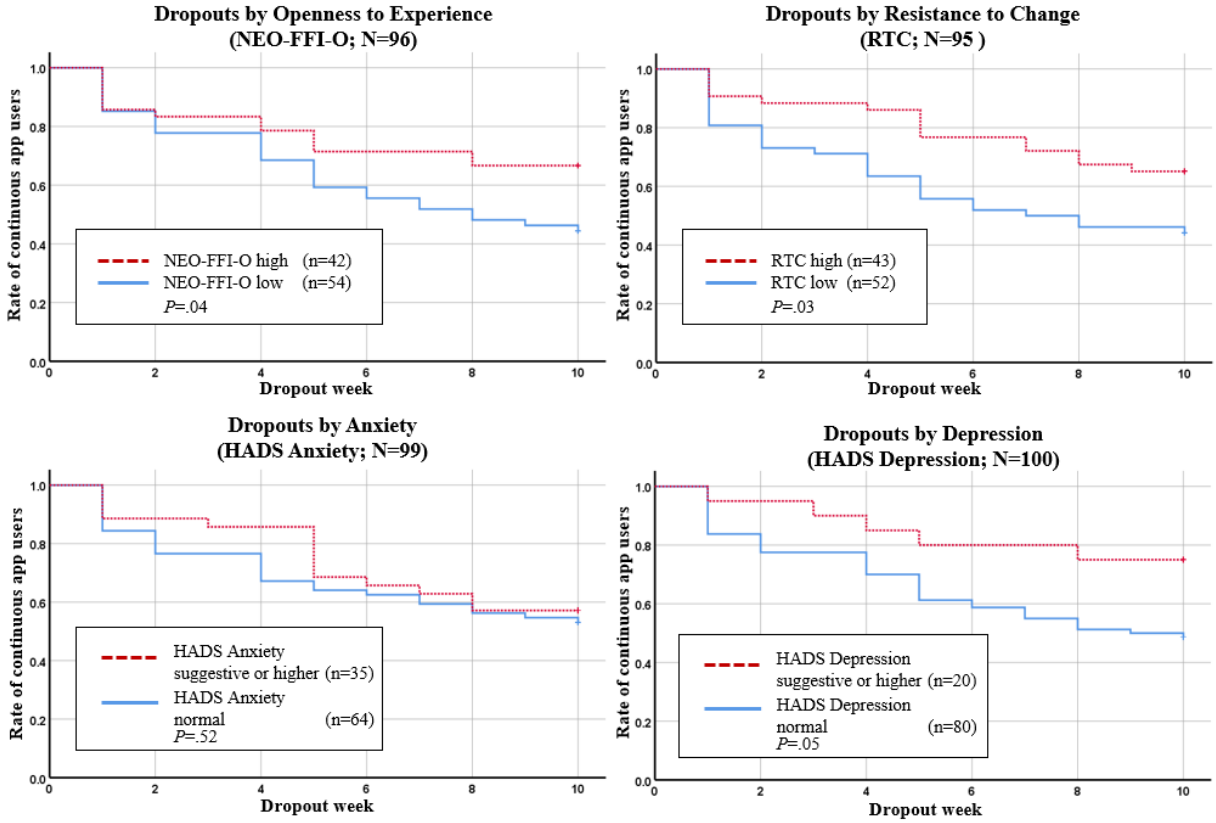


Figure 6. Kaplan-Meier survival curves for continuous app users by openness to experience, resistance to change, anxiety and depression. NEO-FFI-O: NEO Five-Factor Inventory - openness to experience; RTC: Resistance to Change; HADS: Hospital Anxiety Depression Scale.



Qualitative Results

We invited 8 patients (2 from setting 1, 7 female, mean age 50.70 years (SD 15.06), 3 with breast cancer) to an interview, and all agreed to take part. Interviews were conducted between October 2016 and April 2017 and lasted on average 23 min. The qualitative analysis of the interviews yielded 4 themes which were as follows: (1) general feedback regarding the app, (2) suggestions for improvement, (3) personal preferences, and (4) reasons for app use and nonuse.

General feedback about the app was predominantly positive. The interviewed patients appreciated the simplicity of the app and the easy-to-use interface. One patient stated the following regarding the design:

> *It was great. It was very simple, very self-explanatory. You didn't need to look around a lot and it also looked good. Yes, in any case, well designed.* [Female, 35 years old]

Two patients who attended an MBM course evaluated the app as a good addition to the face-to-face MBM course. The feedback about the number of exercises was mixed: Some patients regarded the implemented 3 exercises as sufficient, whereas others would welcome a larger selection of exercises. Most patients interviewed used and appreciated the reminder in the app. Some patients mentioned that they would have been less compliant without the reminders and, therefore, perceived the reminder as helpful for a continuous app use. For instance, 1 patient stated the following:

> *Yes, [the reminder function] was very good. A couple of times this was very good. I would have forgotten it a couple of times, if I wouldn't have had this reminder.* [Female, 63 years old]

Patients offered various suggestions for improvement. Several patients mentioned that they would like exercises with background music. One patient explained it as follows:

> *I think it is also precisely the high art of meditation or relaxation that you can relax as much as possible while not falling asleep. Some need chimes, while others need absolute silence to be able to do this. And I realize that when it is absolutely silent, either I fall asleep or I start to contemplate. When I have some music or chimes, it works better for me personally.* [Female, 42 years old]

As stated above, some patients also would welcome a larger variety of exercises (eg, autogenic training) or variations in the duration of exercises. Another patient stated also that the recordings of the exercises were too clean (ie, no noises from breathing), as the exercises were recorded in a studio. This led to the patient being startled when the narrator continued with the instruction after a moment of silence. One patient mentioned she did not set up the reminder during the first time she used the app and later forgot about the reminder function. Therefore, this patient suggested that the reminder function could be placed more prominently in the app instead of the options menu. The interviewer also inquired if the patients would appreciate a feedback system in the form of exercise statistics. The majority of interviewed patients had the opinion that such a feature would not be helpful. Some patients stated that statistics might even be stressful, as it might lead to a guilty conscience if the patient is not using the exercises as often as planned. One patient suggested that statistics might be added to the app as an optional feature. Only 2 patients thought that such a feature might be helpful.

The third topic that emerged from the qualitative analysis was personal preferences. Most interviewed patients mentioned that they developed some form of preference regarding the app use (eg, preference for a specific exercise, gender of narrator, time of day when using the exercises) during the intervention while they were trying out what suits them best. One patient stated the following:

> *Right at the beginning I tried [to do the exercises] before I went to bed. But I'm not a fan of having my cellphone, when I fall asleep, next to my head for the entire night. For this reason I changed it to lunchtime.* [Female, 31 years old]

The fourth topic that emerged from the qualitative analysis involved reasons for app use and nonuse. As a reason for using the app, patients mentioned that the exercises were beneficial and helped them to relax. One patient stated the following:

> *[The app exercises] have been good for me. I will continue to do my exercises. […] I believe I benefit [from the exercises]. It also makes you happy.* [Female, 63 years old]

As reasons for nonuse, 3 patients mentioned that they had previous experience with meditation or relaxation exercises. Therefore, these patients were already used to exercise routines, which differed from the instructions or the manner in general of the app exercises. One of these patients mentioned that she had learned and was used to silent meditation, and therefore the guided exercises in the app were more distracting than helpful to her. Another patient mentioned that she had experience in guided meditation and relaxation exercises, which differed linguistically and in form of conduct compared with the app exercises. This patient mentioned that she was unable to get used to these new exercises and was repeatedly comparing the app exercises with the already known exercises. Therefore, this patient could not relax as intended during the app exercises. A third patient mentioned that she was used to exercises with more guidance and described her experience with the app as follows:

> *Maybe because [the instructions in the app exercises] were different from what I was used to do by myself, where [the exercise] was guided the entire time. […] I did consider it more bothersome that…[…] your thoughts drift away because you get the feeling that [the exercise] should continue.* [Female, 49 years old]

As a further reason for nonuse, 1 patient mentioned that she was distracted by the choice of words and expressions in the app exercises. This patient mentioned that she had studied linguistics and had learned to closely scrutinize language. This caused her to be distracted during the app exercises, which is why she stopped using the app. Another patient mentioned that she suffered from cancer-related fatigue and that she was not able to complete an exercise when she was unduly fatigued.

Furthermore, 1 patient stated he had technical problems with his smartphone and therefore was not able to use the app during the entire 10 weeks.

Discussion

Summary of Findings

mHealth interventions with the aim of reducing distress in cancer patients seem promising due to easy access and potential positive effects for patients. To our knowledge, this is one of the first studies looking in detail at characteristics of users, adherence rates, and possible predictors for adherence in a mindfulness and relaxation mHealth study for cancer patients. This feasibility study showed that adherence to the mHealth intervention during the first 10 weeks was acceptable, with 54% of patients still using the app regularly in week 10 with a median of completed exercises ranging from 2 to 4 per week. Therefore, our study does not confirm the concern that adherence in mHealth interventions is in general poor, which would limit treatment implementation. The adherence of our patients is also comparable with recent research on adherence to e- and mHealth interventions for cancer patients [32,33]. A study by Beatty el al [32] reported that 60% of cancer patients completed 4 or more modules of an eHealth intervention with 6 modules, which aimed at reducing distress in cancer patients. A mindfulness app study for cancer patients and caregivers [33] reported that 71% of the participants practiced with the app on more than half of the days throughout 8 weeks.

The uptake of our intervention was good, with 117 screened and eligible patients, of whom 100 patients returned the informed consent form. In addition, 74%, mainly female, patients enrolled in this study, which is consistent with characteristics of mHealth users in other studies with 84% female patients [47] and 54% female patients [48]. The mean age of participating patients was 53 years; this is comparable with other face-to-face mindfulness and relaxation interventions [49] or Web-based interventions for cancer patients [32]. The interviews showed that the patients were satisfied with the app in general. However, several and sometimes contradictory suggestions were made for improving the app, such as less versus more guidance in the exercises and larger variety in exercises versus the notion that 3 exercises are sufficient.

Predictors for Adherence

Of a total of 8 investigated predictors for continuous app use, 4 turned out to be statistically significant. The strongest predictor was gender, with higher adherence in female cancer patients. Beyond the higher interest of female cancer patients to participate in this mindfulness and relaxation mHealth study, they were also more adherent after starting with the exercises. This result is in line with a study by Ruland et al [50], in which an analysis of use patterns in an eHealth intervention to support cancer patients' illness management revealed that female patients used the system almost twice as often as male patients. However, a study by Duman-Lubberding et al [51] investigated the feasibility of a Web-based self-management app and did not find a gender difference in adherence. Therefore, it seems likely that the type of intervention (eg, relaxation and mindfulness meditation) might be relevant for gender differences in adherence, which is also in line with studies about the use of complementary and alternative medicine, where users tend to be more often female [52,53].

A second predictor for continuous app use was the personality trait openness to experience, whereby higher openness to experience predicted more continuous app use. This result fits with previous research, which has shown that openness to experience predicts the use of complementary and alternative medicine, including mindfulness and relaxation [54,55]. Our study confirms that higher openness to experience still predicts the adherence to a mindfulness and relaxation intervention, even if the intervention is delivered through an app.

A third predictor for continuous app use was a higher score in resistance to change. This finding is contrary to our hypothesis, as we assumed that higher resistance to change would be associated with less adherence as the intervention promotes a new health behavior. However, our results indicate the opposite. When a patient has decided to follow a new exercise routine (ie, mindfulness and relaxation mHealth intervention), a higher resistance to change actually promotes continuous app use. To our knowledge, the Resistance to Change Scale had not previously been used to predict adherence to mHealth interventions for cancer patients. However, a study conducted in China by Deng et al [56] showed that resistance to change is negatively related to the intention to use mHealth services. Another study showed that resistance to change is negatively related to perceived usefulness of mHealth in elderly people in China [57]. Therefore, on the one hand, resistance to change might be a barrier for the uptake of an mHealth intervention, but on the other, it might be supportive in adhering to a new commitment, such as the regular use of a mindfulness and relaxation app.

A fourth predictor for continuous app use was higher depressive symptoms. This finding is surprising, as depressive symptoms are associated with decreased motivation and reduced activity [58]. In line with these corollaries, a study investigating a mindfulness-based cancer recovery program [49] reported a negative correlation of depressive symptoms and practicing time of yoga at home. Another study [59] reported that moderate to severe depressive symptoms predicted lower adherence to adjuvant cancer therapies. However, a study by Børøsund et al [60] found that high levels of depression were associated with high use of components of a Web-based illness management program in breast cancer patients. As depressed patients are oftentimes troubled with motivational deficits and face difficulties to stay active, the development of effective interventions with a good adherence in depressed patients is highly relevant. Our study indicates that mindfulness and relaxation mHealth interventions seem a feasible tool as supportive interventions for cancer patients with elevated depressive symptoms. This finding might also indicate that some patients adhere better to mindfulness and relaxation (ie, patients with higher depressive symptoms), whereas other patient groups with lower levels of distress are not in need of such interventions or might prefer other intervention types.

Limitations

This study has some limitations. First, the number of potentially interested patients for this intervention could not be assessed. Therefore, we were not able to calculate the rate of the total number of eligible cancer patients compared with the number of cancer patients with interest in a mindfulness and relaxation mHealth intervention. Second, for our definition of continuous app, we had no empirical data because the necessary dose for clinically significant improvements is still unclear for this kind of mHealth intervention. Instead, we opted for a clinical and rational justification, in which the term "continuous" use was operationalized as an at least weekly use of 1 or more app exercises. Third, the use of generated categories for age and the median split for other variables as predictors can be challenged. For age, we chose 3 age categories that represent patients of younger (18-40), middle (41-55), and older (56 plus) age. The use of median split variables has been critically discussed in the literature (see eg, Iacobucci et al [61]), with a major critique being the loss of information. In our case, the loss of information can be justified with the illustrative capacity of Kaplan-Meier survival curves and the following use of a multivariate Cox proportional hazards regression. Fourth, as the sample size was an outcome in itself, we did not perform an a priori sample size calculation. With a sample size (N) of 100, we had a power of 0.63 in the Cox proportional hazards regression for the main effect (OR 2.16) of gender as a predictor. For a power of 0.8, a sample size (N) of 150 would be necessary.

Conclusions and Future Research

The acceptable adherence to the intervention and the generally positive feedback by patients indicate that this app intervention is feasible. Suggestions for improvement by patients indicate that patients' needs are heterogeneous, which should be taken into account when developing other mHealth interventions. Due to the acceptable adherence and positive feedback by cancer patients, mindfulness and relaxation mHealth interventions might be promising supportive interventions, also for cancer patients with elevated depressive symptoms.

To further prove the importance of mindfulness and relaxation mHealth interventions for cancer patients, future research needs to investigate their effectiveness. As the dose potentially influences the effectiveness of mindfulness and relaxation interventions, future research should also look into dose-response relationships between the time spent exercising with the app and health outcomes. Knowledge of such a dose-response relationship could be of use to guide subsequent studies regarding intervention duration and practice recommendation for patients. This study suggests that variability across patients in weekly app use is large. About half of the patients used the app exercises continuously over 10 weeks and therefore adhered to the intervention. These interindividual differences in the use of app exercises underline the importance to take adherence into account when analyzing effectiveness data. Furthermore, these interindividual differences on adherence bring up the question if mindfulness- and relaxation-based mHealth interventions are better suited for specific patient groups (eg females, patients with higher depressive symptoms). In turn, male patients or patients with less distress might not be in need of such interventions or might require additional motivational interventions.

Acknowledgments

This study was funded by the Swiss Cancer League (KLS-3564-02-2015). The authors acknowledge and thank the contribution of Isabelle Werninger, who supported the project in both administration and data entry.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the CanRelax App exercises.

[PNG File, 2MB - mhealth_v6i12e11271_app1.png]

Multimedia Appendix 2

Screenshots of the CanRelax App notification feature.

[PNG File, 1MB - mhealth_v6i12e11271_app2.png]

Multimedia Appendix 3

Interview guideline.

[PDF File (Adobe PDF File), 12KB - mhealth_v6i12e11271_app3.pdf]

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 01;136(5):E359-E386 [FREE Full text] [doi: 10.1002/ijc.29210] [Medline: 25220842]
2. World Health Organization. Cancer. 2018 URL: http://www.who.int/en/news-room/fact-sheets/detail/cancer [accessed 2018-05-15] [WebCite Cache ID 6zR1qARmJ]
3. Mehnert A, Brähler E, Faller H, Härter M, Keller M, Schulz H, et al. Four-week prevalence of mental disorders in patients with cancer across major tumor entities. J Clin Oncol 2014 Nov 01;32(31):3540-3546. [doi: 10.1200/JCO.2014.56.0086] [Medline: 25287821]
4. Zabora J, BrintzenhofeSzoc K, Curbow B, Hooker C, Piantadosi S. The prevalence of psychological distress by cancer site. Psychooncology 2001;10(1):19-28. [Medline: 11180574]
5. Carlson LE. Distress Management Through Mind-Body Therapies in Oncology. J Natl Cancer Inst Monogr 2017 Nov 01;2017(52). [doi: 10.1093/jncimonographs/lgx009] [Medline: 29140490]
6. Faller H, Koch U, Brähler E, Härter M, Keller M, Schulz H, et al. Satisfaction with information and unmet information needs in men and women with cancer. J Cancer Surviv 2016 Feb;10(1):62-70. [doi: 10.1007/s11764-015-0451-1] [Medline: 25956402]
7. Burg MA, Adorno G, Lopez EDS, Loerzel V, Stein K, Wallace C, et al. Current unmet needs of cancer survivors: analysis of open-ended responses to the American Cancer Society Study of Cancer Survivors II. Cancer 2015 Feb 15;121(4):623-630 [FREE Full text] [doi: 10.1002/cncr.28951] [Medline: 25581252]
8. Carlson LE, Angen M, Cullum J, Goodey E, Koopmans J, Lamont L, et al. High levels of untreated distress and fatigue in cancer patients. Br J Cancer 2004 Jun 14;90(12):2297-2304 [FREE Full text] [doi: 10.1038/sj.bjc.6601887] [Medline: 15162149]
9. Waller A, Williams A, Groff SL, Bultz BD, Carlson LE. Screening for distress, the sixth vital sign: examining self-referral in people with cancer over a one-year period. Psychooncology 2013 Feb;22(2):388-395. [doi: 10.1002/pon.2102] [Medline: 22135205]
10. Ryan H, Schofield P, Cockburn J, Butow P, Tattersall M, Turner J, et al. How to recognize and manage psychological distress in cancer patients. Eur J Cancer Care (Engl) 2005 Mar;14(1):7-15. [doi: 10.1111/j.1365-2354.2005.00482.x] [Medline: 15698382]
11. Faller H, Schuler M, Richard M, Heckl U, Weis J, Küffner R. Effects of psycho-oncologic interventions on emotional distress and quality of life in adult patients with cancer: systematic review and meta-analysis. J Clin Oncol 2013 Feb 20;31(6):782-793. [doi: 10.1200/JCO.2011.40.8922] [Medline: 23319686]
12. Sheard T, Maguire P. The effect of psychological interventions on anxiety and depression in cancer patients: results of two meta-analyses. Br J Cancer 1999 Aug;80(11):1770-1780 [FREE Full text] [doi: 10.1038/sj.bjc.6690596] [Medline: 10468295]
13. Carlson LE, Zelinski E, Toivonen K, Flynn M, Qureshi M, Piedalue K, et al. Mind-Body Therapies in Cancer: What Is the Latest Evidence? Curr Oncol Rep 2017 Aug 18;19(10):67. [doi: 10.1007/s11912-017-0626-1] [Medline: 28822063]
14. Monti DA, Sufian M, Peterson C. Potential role of mind-body therapies in cancer survivorship. Cancer 2008 Jun 01;112(11 Suppl):2607-2616 [FREE Full text] [doi: 10.1002/cncr.23443] [Medline: 18428193]
15. Astin JA, Shapiro SL, Eisenberg DM, Forys KL. Mind-Body Medicine: State of the Science, Implications for Practice. The Journal of the American Board of Family Medicine 2003 Mar 01;16(2):131-147. [doi: 10.3122/jabfm.16.2.131]
16. Greenlee H, DuPont-Reyes MJ, Balneaves LG, Carlson LE, Cohen MR, Deng G, et al. Clinical practice guidelines on the evidence-based use of integrative therapies during and after breast cancer treatment. CA Cancer J Clin 2017 May 06;67(3):194-232 [FREE Full text] [doi: 10.3322/caac.21397] [Medline: 28436999]
17. Huang H, He M, Wang H, Zhou M. A meta-analysis of the benefits of mindfulness-based stress reduction (MBSR) on psychological function among breast cancer (BC) survivors. Breast Cancer 2016 Jul;23(4):568-576. [doi: 10.1007/s12282-015-0604-0] [Medline: 25820148]
18. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. Journal of Psychosomatic Research 2004 Jul;57(1):35-43. [doi: 10.1016/s0022-3999(03)00573-7]
19. Dimidjian S, Arch JJ, Schneider RL, Desormeau P, Felder JN, Segal ZV. Considering Meta-Analysis, Meaning, and Metaphor: A Systematic Review and Critical Examination of “Third Wave” Cognitive and Behavioral Therapies. Behav Ther 2016 Dec;47(6):886-905. [doi: 10.1016/j.beth.2016.07.002] [Medline: 27993339]
20. Blanck P, Perleth S, Heidenreich T, Kröger P, Ditzen B, Bents H, et al. Effects of mindfulness exercises as stand-alone intervention on symptoms of anxiety and depression: Systematic review and meta-analysis. Behav Res Ther 2018 Mar;102:25-35. [doi: 10.1016/j.brat.2017.12.002] [Medline: 29291584]
21. King N, Balneaves LG, Levin GT, Nguyen T, Nation JG, Card C, et al. Surveys of Cancer Patients and Cancer Health Care Providers Regarding Complementary Therapy Use, Communication, and Information Needs. Integr Cancer Ther 2015 Nov;14(6):515-524. [doi: 10.1177/1534735415589984] [Medline: 26069277]
22. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. J Behav Med 2008 Feb;31(1):23-33. [doi: 10.1007/s10865-007-9130-7] [Medline: 17899351]

23. 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]
24. Luxton D, McCann R, Bush N, Mishkind M, Reger G. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Prof Psychol Res Pr 2011;42(6):505-512. [doi: 10.1037/a0024485]
25. Statista. Smartphone user penetration as percentage of total global population from to 2021. 2014. 2018 URL: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/ [accessed 2018-05-22] [WebCite Cache ID 6zbNphoVM]
26. 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;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
27. 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 Jan;61(1):50-62 [FREE Full text] [doi: 10.3322/caac.20093] [Medline: 21205833]
28. Mikolasek M, Berg J, Witt CM, Barth J. Effectiveness of Mindfulness- and Relaxation-Based eHealth Interventions for Patients with Medical Conditions: a Systematic Review and Synthesis. Int J Behav Med 2017 Jul 27. [doi: 10.1007/s12529-017-9679-7] [Medline: 28752414]
29. Fish J, Brimson J, Lynch S. Mindfulness Interventions Delivered by Technology Without Facilitator Involvement: What Research Exists and What Are the Clinical Outcomes? Mindfulness (N Y) 2016 Jun;7(5):1011-1023 [FREE Full text] [doi: 10.1007/s12671-016-0548-2] [Medline: 27642370]
30. Spijkerman MPJ, Pots WTM, Bohlmeijer ET. 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]
31. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
32. Beatty L, Kemp E, Binnion C, Turner J, Milne D, Butow P, et al. Uptake and adherence to an online intervention for cancer-related distress: older age is not a barrier to adherence but may be a barrier to uptake. Support Care Cancer 2017 Dec;25(6):1905-1914. [doi: 10.1007/s00520-017-3591-1] [Medline: 28155018]
33. Kubo A, Altschuler A, Kurtovich E, Hendlish S, Laurent C, Kolevska T, et al. A Pilot Mobile-Based Mindfulness Intervention for Cancer Patients and Their Informal Caregivers. Mindfulness 2018:2018-2010. [doi: 10.1007/s12671-018-0931-2]
34. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322]
35. Glaser B, Strauss A. Discovery of grounded theory: Strategies for qualitative review. Chicago, IL: Aldine De Gruyter; 1967.
36. Blödt S, Pach D, Roll S, Witt CM. Effectiveness of app-based relaxation for patients with chronic low back pain (Relaxback) and chronic neck pain (Relaxneck): study protocol for two randomized pragmatic trials. Trials 2014 Dec 15;15:490 [FREE Full text] [doi: 10.1186/1745-6215-15-490] [Medline: 25511185]
37. National Comprehensive Cancer Network. Distress Management Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2003 Jul 01;1(3):344-344. [doi: 10.6004/jnccn.2003.0031]
38. Mehnert A, Müller D, Lehmann C, Koch U. Die deutsche Version des NCCN Distress-Thermometers. Zeitschrift für Psychiatrie, Psychologie und Psychotherapie 2006 Jan;54(3):213-223. [doi: 10.1024/1661-4747.54.3.213]
39. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol 1993 Mar;11(3):570-579. [doi: 10.1200/JCO.1993.11.3.570] [Medline: 8445433]
40. 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]
41. Jacobsen PB, Donovan KA, Trask PC, Fleishman SB, Zabora J, Baker F, et al. Screening for psychologic distress in ambulatory cancer patients. Cancer 2005 Apr 01;103(7):1494-1502 [FREE Full text] [doi: 10.1002/cncr.20940] [Medline: 15726544]
42. Costa P, McCrae R. Neo PI/FFI manual supplement for use with the NEO Personality Inventory and the NEO Five-Factor Inventory. Odessa, FL: Psychological Assessment Resources; 1989.
43. Oreg S. Resistance to change: Developing an individual differences measure. Journal of Applied Psychology 2003;88(4):680-693. [doi: 10.1037/0021-9010.88.4.680]
44. 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]
45. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp; 2017.
46. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. 11th edition. Basel: Beltz Pädagogik; 2010.

47. Elavsky S, Smahel D, Machackova H. Who are mobile app users from healthy lifestyle websites? Analysis of patterns of app use and user characteristics. Transl Behav Med 2017 Dec;7(4):891-901 [FREE Full text] [doi: 10.1007/s13142-017-0525-x] [Medline: 28929368]
48. Wang BR, Park J, Chung K, Choi IY. Influential Factors of Smart Health Users according to Usage Experience and Intention to Use. Wireless Pers Commun 2014 Apr 24;79(4):2671-2683. [doi: 10.1007/s11277-014-1769-0]
49. Tamagawa R, Speca M, Stephen J, Pickering B, Lawlor-Savage L, Carlson LE. Predictors and Effects of Class Attendance and Home Practice of Yoga and Meditation Among Breast Cancer Survivors in a Mindfulness-Based Cancer Recovery (MBCR) Program. Mindfulness 2015 Jan 9;6(5):1201-1210. [doi: 10.1007/s12671-014-0381-4]
50. Ruland CM, Maffei RM, Børøsund E, Krahn A, Andersen T, Grimsbø GH. Evaluation of different features of an eHealth application for personalized illness management support: cancer patients' use and appraisal of usefulness. Int J Med Inform 2013 Jul;82(7):593-603. [doi: 10.1016/j.ijmedinf.2013.02.007] [Medline: 23507561]
51. Duman-Lubberding S, van Uden-Kraan CF, Jansen F, Witte BI, van der Velden LA, Lacko M, et al. Feasibility of an eHealth application "OncoKompas" to improve personalized survivorship cancer care. Support Care Cancer 2016 May;24(5):2163-2171 [FREE Full text] [doi: 10.1007/s00520-015-3004-2] [Medline: 26563178]
52. Bishop FL, Lewith GT. Who Uses CAM? A Narrative Review of Demographic Characteristics and Health Factors Associated with CAM Use. Evid Based Complement Alternat Med 2010 Mar;7(1):11-28 [FREE Full text] [doi: 10.1093/ecam/nen023] [Medline: 18955327]
53. Reid R, Steel A, Wardle J, Trubody A, Adams J. Complementary medicine use by the Australian population: a critical mixed studies systematic review of utilisation, perceptions and factors associated with use. BMC Complement Altern Med 2016 Jun 11;16:176 [FREE Full text] [doi: 10.1186/s12906-016-1143-8] [Medline: 27289517]
54. Thomson P, Jones J, Browne M, Leslie SJ. Psychosocial factors that predict why people use complementary and alternative medicine and continue with its use: a population based study. Complement Ther Clin Pract 2014 Nov;20(4):302-310. [doi: 10.1016/j.ctcp.2014.09.004] [Medline: 25456023]
55. Honda K, Jacobson JS. Use of complementary and alternative medicine among United States adults: the influences of personality, coping strategies, and social support. Prev Med 2005 Jan;40(1):46-53. [doi: 10.1016/j.ypmed.2004.05.001] [Medline: 15530580]
56. Deng Z, Mo X, Liu S. Comparison of the middle-aged and older users' adoption of mobile health services in China. Int J Med Inform 2014 Mar;83(3):210-224. [doi: 10.1016/j.ijmedinf.2013.12.002] [Medline: 24388129]
57. Guo X, Sun Y, Wang N, Peng Z, Yan Z. The dark side of elderly acceptance of preventive mobile health services in China. Electron Markets 2012 Dec 11;23(1):49-61. [doi: 10.1007/s12525-012-0112-4]
58. Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5). 5th edition. Washington, DC: American Psychiatric Association; 2016.
59. Barber B, Dergousoff J, Nesbitt M, Mitchell N, Harris J, O'Connell D, et al. Depression as a predictor of postoperative functional performance status (PFPS) and treatment adherence in head and neck cancer patients: a prospective study. J Otolaryngol Head Neck Surg 2015 Sep 18;44:38 [FREE Full text] [doi: 10.1186/s40463-015-0092-4] [Medline: 26385356]
60. Børøsund E, Cvancarova M, Ekstedt M, Moore SM, Ruland CM. How user characteristics affect use patterns in web-based illness management support for patients with breast and prostate cancer. J Med Internet Res 2013 Mar 01;15(3):e34 [FREE Full text] [doi: 10.2196/jmir.2285] [Medline: 23454601]
61. Iacobucci D, Posavac SS, Kardes FR, Schneider MJ, Popovich DL. The median split: Robust, refined, and revived. Journal of Consumer Psychology 2015 Oct;25(4):690-704. [doi: 10.1016/j.jcps.2015.06.014]

Abbreviations

CRF: case report form
FACT-G: Functional Assessment of Cancer Therapy-General
FMI: Freiburg Mindfulness Inventory
FoP-Q-SF: Fear of Progression Questionnaire - Short Form
HADS: Hospital Anxiety Depression Scale
ICI: Institute for Complementary and Integrative Medicine, University Hospital Zurich
IQR: interquartile range
MBM: Mind Body Medicine
mHealth: mobile health
NEO-FFI-O: NEO Five-Factor Inventory - openness to experience
OR: odds ratio
PROMIS 29: Patient Reported Outcomes Measurement Information System 29
RTC: Resistance to Change

Edited by G Eysenbach; submitted 19.06.18; peer-reviewed by E Børøsund, M Schuler, P Butow; comments to author 15.08.18; revised version received 15.09.18; accepted 19.09.18; published 06.12.18.

Please cite as:
Mikolasek M, Witt CM, Barth J
Adherence to a Mindfulness and Relaxation Self-Care App for Cancer Patients: Mixed-Methods Feasibility Study
JMIR Mhealth Uhealth 2018;6(12):e11271
URL: https://mhealth.jmir.org/2018/12/e11271/
doi:10.2196/11271
PMID:30522990

Original Paper

Predictors of Engagement, Response to Follow Up, and Extent of Alcohol Reduction in Users of a Smartphone App (Drink Less): Secondary Analysis of a Factorial Randomized Controlled Trial

Claire Garnett[1], PhD; Olga Perski[1,2], PhD; Ildiko Tombor[1], PhD; Robert West[1], PhD; Susan Michie[2], DPhil; Jamie Brown[1], PhD

[1]Department of Behavioural Science and Health, University College London, London, United Kingdom

[2]Department of Clinical, Education and Health Psychology, University College London, London, United Kingdom

Corresponding Author:
Claire Garnett, PhD
Department of Behavioural Science and Health
University College London
1-19 Torrington Place
London, WC1E 7HB
United Kingdom
Phone: 44 2076791742 ext 41742
Email: c.garnett@ucl.ac.uk

Abstract

Background: Digital interventions for alcohol can help achieve reductions in hazardous and harmful alcohol consumption. The Drink Less app was developed using evidence and theory, and a factorial randomized controlled trial (RCT) suggested that 4 of its intervention modules may assist with drinking reduction. However, low engagement is an important barrier to effectiveness, and low response to follow up is a challenge for intervention evaluation. Research is needed to understand what factors influence users' level of engagement, response to follow up, and extent of alcohol reduction.

Objective: This study aimed to investigate associations between user characteristics, engagement, response to follow up, and extent of alcohol reduction in an app-based intervention, Drink Less.

Methods: This study involved a secondary data analysis of a factorial RCT of the Drink Less app. Participants (N=672) were aged 18 years or older, lived in the United Kingdom, and had an Alcohol Use Disorders Identification Test score >7 (indicative of excessive drinking). Sociodemographic and drinking characteristics were assessed at baseline. Engagement was assessed in the first month of use (number of sessions, time on app, number of days used, and percentage of available screens viewed). Response to follow up and extent of alcohol reduction (change in past week consumption) were measured after 1 month. Associations were assessed using unadjusted and adjusted linear or logistic regression models.

Results: Age (all unstandardized regression coefficients [B] >.02, all P<.001) and post-16 educational qualifications (all B>.18, all P<.03) were positively associated with all engagement outcomes. Age (odds ratio [OR] 1.04, P<.001), educational qualifications (OR 2.11, P<.001), and female gender (OR 1.58, P=.02) were positively associated with response to follow up. Engagement outcomes predicted response to follow up (all OR>1.02, all P<.001) but not the extent of alcohol reduction (all −.14<B<−.06, all P>.07). Baseline drinking characteristics were the only variables associated with the extent of alcohol reduction among those followed up (all B>.49, all P<.001).

Conclusions: Users of the alcohol reduction app, Drink Less, who were older and had post-16 educational qualifications engaged more and were more likely to respond at 1-month follow up. Higher baseline alcohol consumption predicted a greater extent of alcohol reduction among those followed up but did not predict engagement or response to follow up. Engagement was associated with response to follow up but was not associated with the extent of alcohol reduction, which suggests that the Drink Less app does not have a dose-response effect.

Trial Registration: International Standard Randomised Controlled Trial Number ISRCTN40104069; http://www.isrctn.com/ISRCTN40104069 (Archived by WebCite at http://www.webcitation.org/746HqygIV)

(JMIR Mhealth Uhealth 2018;6(12):e11175) doi:10.2196/11175

KEYWORDS

smartphone; mobile phone; alcohol drinking; mobile apps; engagement

Introduction

Background

Excessive alcohol consumption is a priority for public health and has a large economic impact on society because of lost productivity, crime, and health care costs [1-3]. Digital behavior change interventions (DBCIs) focused on alcohol reduction show promise as they can help achieve reductions in hazardous and harmful alcohol consumption [4]; can improve the accessibility of support; have a low incremental cost (once developed); are anonymous, and avoid potential stigma associated with seeking help in person. Smartphone apps ("apps") have the added advantage of being almost constantly available and, therefore, able to provide support when and where needed. Drink Less is an alcohol-reduction app aimed at those who consume alcohol excessively (defined as an Alcohol Use Disorders Identification Test [AUDIT] score of 8 and above [5]). The Drink Less app was developed systematically and in line with the principles of Open Science and consisted of 1 core module and 5 experimental intervention modules (described in detail elsewhere [6]). A first-phase factorial randomized controlled trial (RCT) suggested that combinations of 4 of the intervention modules assisted short-term drinking reduction; however, it also showed that the app suffered from high rates of attrition [7].

Eysenbach's law of attrition distinguishes between 2 types of attrition: nonusage and dropout [8]. Nonusage attrition refers to the complete lack of, or low, engagement with a DBCI. Engagement with DBCIs can be defined as "the extent of DBCI use (eg, amount, depth, frequency, duration)" [9] and can be measured through automatic recording of DBCI use. A minimum level of engagement with an intervention is assumed to be necessary for that intervention to have its desired effect, although there is no research on what constitutes the minimum required. However, observed levels of engagement with available DBCIs have often been considered too limited to support behavior changes [10]. The second type of attrition—dropout—refers to participants being lost to follow up. A low rate of response to follow up is a major methodological challenge in intervention evaluation [11] because it reduces statistical power and, therefore, the ability to estimate effectiveness accurately [12]. Trials of DBCIs—especially those involving remote recruitment—appear particularly vulnerable to low response rates to follow up [8,13,14].

The law of attrition also proposes that engagement and response to follow up are positively associated: if users stop engaging with an intervention, then they are unlikely to respond to follow up [8]. This positive association between engagement and response to follow up was found in a systematic review of Web-based health interventions [15] and in each trial arm of an RCT of a Web-based alcohol intervention [16]. Although the relationship may depend upon the intervention and context in which it is being studied, there have been other reports of higher follow-up responses in the control condition if the intervention arm was particularly demanding [17,18]. It may be that users become fatigued in the experimental condition and decide too much time has already been dedicated to the trial.

To improve the likelihood of behavior change and the validity of DBCI trial's results, research is needed to understand whether certain users are less likely to engage with the intervention, respond to follow up, or change behavior. The identification of predictors of engagement, response to follow up, and behavior change could inform the development of tailored strategies for specific user groups in DBCI trials. The relationship between engagement and response to follow up has not yet been evaluated in an app-based alcohol intervention.

Predictors of Engagement and Response to Follow Up

Existing literature indicates that being female, older, and better educated predicts higher engagement and greater response to follow up in Web-based alcohol interventions [16,19-21]. Drinking characteristics tend to have an impact on engagement and response to follow up in opposite directions; people at less risk of alcohol harm [20,21] and consuming fewer units a week [16,21,22] are more likely to respond to follow up, though they show lower levels of engagement [16,22-24]. However, some of these studies involved the specific population of students [22-24], and there are also inconsistencies in the evidence when studying problem drinkers [25,26]. Hence, these findings may not be generalizable to the general population, and current evidence regarding user characteristics that may predict engagement and response to follow up in Web-based alcohol interventions is ambiguous. There is also a lack of evidence specifically relating to app-based alcohol interventions, which may differ from the equivalent Web-based intervention in terms of user characteristics. A recent study comparing users of app and Web-based versions of the same Drinks Meter intervention found that app users were younger and had higher levels of alcohol consumption compared with website users [27].

Predictors of Alcohol Reduction

Drinkers in England who report an attempt to reduce their drinking are more likely to be older, female, of higher socioeconomic status, have high levels of alcohol consumption, and less likely to be white [28]. Attempts and success in those attempts are distinct and likely to be independently predicted, and there is a lack of research on the predictors of successful attempts to reduce drinking, particularly in the context of DBCIs. Understanding which users are less likely to be successful in reducing their drinking when using a DBCI can inform the development of additional strategies to support them and help identify those who are more likely to require face-to-face support. It is also important to understand the relationship between engagement with a DBCI and behavior change and to establish whether there is a threshold level of engagement required to achieve the intended outcomes of the DBCI [29]. However, some studies, conducted across behavioral domains and study settings, have found a positive association between engagement and successful behavior change, suggesting a

dose-response relationship may exist [29-32]. There is no research to date on the relationship between engagement and alcohol reduction in app-based interventions. It is unwise to assume that these relationships will be consistent across different behaviors and types of digital interventions; a recent study found different levels of engagement with a digital stress management intervention, depending on whether it was delivered by an app or a website [33].

This Study

This study investigated the associations between user characteristics, engagement, response to follow up, and extent of alcohol reduction in an app-based intervention, Drink Less, and addressed the following research questions:

1. What associations, if any, exist between user characteristics (sociodemographic and drinking) measured when drinkers register with the app and (a) engagement, (b) likelihood of response to follow up, and (c) extent of alcohol reduction at follow up among those followed up?
2. What associations, if any, exist between engagement and (a) likelihood of response to follow up and (b) extent of alcohol reduction at follow up among those followed up?

Methods

Design

The study design is secondary data analysis from a factorial RCT of the Drink Less app between May and August 2016 (reported in full elsewhere [7]). This analysis was not planned before the factorial RCT. Ethical approval was provided by the University College London's ethics committee ("Optimisation and implementation of interventions to change health-related behaviours" project [CEHP/2013/508]).

Sample and Recruitment

Participants were eligible if they were: aged 18 years or above; lived in the United Kingdom; had an AUDIT score of 8 or above (indicative of excessive alcohol consumption warranting intervention [5]); confirmed that they were interested in reducing their drinking; provided an email address; and downloaded a *trial version* of the app. People who downloaded the app more than once were removed, with the first case of download retained for the trial. The app was listed in the iTunes store and was promoted through organizations such as Public Health England and Cancer Research UK. The sample size of 672 was prespecified and calculated, so there was more than 80% power to detect a mean change in alcohol consumption of 5 units between the intervention modules.

Intervention

Drink Less was designed as a stand-alone intervention available to anyone seeking digital support for reducing excessive alcohol consumption. It is centered on a goal-setting module with 5 intervention modules: (1) normative feedback, (2) cognitive bias retraining, (3) self-monitoring and feedback, (4) action planning, and (5) identity change. The app also contains standard features such as the AUDIT questionnaire and feedback on users' results, the UK drinking guidelines, and links for additional support. Usability testing was conducted during the original app development to understand the user experience and refine the app [34]. Each intervention module existed in 2 versions: *enhanced* (the hypothesized active ingredients for reducing alcohol consumption) and *minimal* (the control). The intervention involved 32 possible options (2^5: 2 versions of the 5 intervention modules) that users could be randomly allocated to. Users completed the AUDIT questionnaire, sociodemographic assessment, and normative feedback module in a tunneled approach before arriving to the main dashboard. Users were then provided with a stepped guide to aid them in exploring the app, although this was optional and users were free to navigate the app as they wished. Full details on the intervention are reported elsewhere [35], and the app is freely available on the iTunes store [36].

Procedures

Data collection began on May 18, 2016, and ended on August 28, 2016. On first opening the app, each user was provided with a participant information sheet and asked to provide consent to participate in the trial. Users who consented were asked to complete the AUDIT and a sociodemographic questionnaire, indicate whether they were interested in drinking less alcohol, and provide their email address for follow up. Users were then given their AUDIT score and informed of their *AUDIT risk zone*. At this point, users who met inclusion criteria were randomized to 1 of the 32 experimental conditions in a block randomization method by the app. The follow-up questionnaire consisted of the AUDIT and usability measures and was conducted 1 month (28 days) after first using the app by means of an in-app questionnaire or by a Web-based survey (Qualtrics) that was distributed by email.

Measures

Engagement was measured as a continuous variable through automatic recording of the extent of DBCI use in terms of amount, depth, frequency, and duration in the 28-day period following registration [9]: number of sessions (ie, frequency of use), a new session was defined as a new screen view after 30 minutes of inactivity [37]; time on app, in minutes (ie, amount); number of days used (ie, duration); and percentage of available screens viewed (ie, depth)—the percentage of unique screens viewed by the user, out of the number of screens available to view (differed depending on treatment group, ranged from 50 to 80).

Response to follow up was a binary (yes or no) measure of completion of the 1-month follow-up questionnaire. The extent of alcohol reduction was measured as the change in past week alcohol consumption (−90 to +90 units) derived from the Alcohol Use Disorders Identification Test-Consumption (AUDIT-C; a brief screening test for alcohol consumption) between time of registration and 1-month follow up.

User characteristic variables were measured at baseline, on first opening the app, and assessed: age (continuous); gender (male or female); employment status (dichotomized into employed vs not employed); ethnicity (dichotomized into white vs not white); education (dichotomized into pre-16 and post-16 educational qualifications); whether they were a current smoker (yes or no); past week alcohol consumption derived from the

AUDIT-C (ranging from 0 to 90 units), and full AUDIT score (ranging from 0 to 40).

Analyses

All analyses were conducted using R version 3.4.0, and the analysis plan was preregistered on the Open Science Framework [38]. Participants with missing data on any variable of interest were excluded from the analyses. The assumptions for a parametric test were assessed (eg, normality of the distribution of residuals), and if these assumptions were not met, the appropriate nonparametric test or transformation (eg, log() transformation for positively skewed data) was used. Descriptive statistics (mean [SD] or median [interquartile range, IQR] to account for a positive skew or n [%], as appropriate) were used to report on the variables included in the analyses (user characteristics, engagement measures, response to follow up, and extent of alcohol reduction).

Generalized linear modeling (linear or logistic, as appropriate) was used to examine the associations between user characteristics (predictor variable) and engagement, response to follow up, or extent of alcohol reduction (outcome variables). Both unadjusted (univariate) and fully adjusted (multivariable) regression models were reported. Treatment group was included in all adjusted analyses as it is a factor relating to the DBCI that may predict engagement [9], response to follow up [16,39], and extent of alcohol reduction [4]. Past week alcohol consumption was not included in this adjusted model because of anticipated high collinearity between past week alcohol consumption and full AUDIT score.

Generalized linear modeling (linear or logistic, as appropriate) was used to examine the associations between engagement (predictor variable) and response to follow up or extent of alcohol reduction (outcome variables). Both unadjusted and fully adjusted (for treatment group and any predictors of the outcome variables) regression models are reported.

Sensitivity analyses were conducted in which the engagement measures were dichotomized into high or low groups and entered in a logistic regression model to see whether the pattern of results differs.

Results

User Characteristics, Engagement Measures, Response to Follow Up, and Extent of Alcohol Reduction in Those Followed Up

A total of 672 participants were included. The mean age was 39.2 years; over half were females (377/672, 56.1%); and the majority were employed (581/672, 86.5%), white (640/672, 95.2%), and had post-16 years' educational qualifications (484/672, 72.0%). About a quarter of participants (165/672, 24.6%) were current smokers and participants consumed a mean of 39.9 units of alcohol in the past week and had a mean AUDIT score of 19.1, indicating excessive alcohol consumption.

Table 1 reports the user characteristics, engagement measures, response to follow up, and, among those who were followed up, extent of alcohol reduction. In the 28 days following registration, participants used the app a median of 5 times and the median number of days the app was used was 4. The median time on the app was 17 minutes 14 seconds, and participants viewed a mean of 39.0% of the available screens. In total, 26.6% of participants (179/672) responded to follow up and of these, 83.2% (149/179) responded through a Web-based survey. There was a mean 14.3 unit reduction in past week alcohol consumption among those participants who responded to follow up.

Associations Between User Characteristics and Measures of Engagement

Tables 2 and 3 report the linear regression models in which the engagement measures (number of sessions, time spent on the app, number of days used, and percentage of available screens viewed) were regressed onto the user characteristic variables. Overall, 3 engagement measures—number of sessions, time spent on the app, and number of days used—were log-transformed because of the non-normality of residuals.

Age was significantly positively associated with all 4 measures of engagement. Education level was significantly positively associated with all 4 measures of engagement: number of sessions, time spent on the app (only when adjusted for other user characteristics and treatment group), the number of days on which the app was used, and the percentage of available screens viewed. Older users and those with post-16 educational qualifications were more likely to have a greater number of sessions, spend more time on the app, use the app on a greater number of days, and view a larger percentage of available screens.

Gender was significantly associated with the percentage of available screens viewed by the user in both unadjusted and adjusted models; users who were female viewed a greater percentage of screens available to them. No other user characteristics were associated with engagement with the Drink Less app.

Associations Between User Characteristics and Likelihood of Response to Follow Up

Table 4 reports the unadjusted and adjusted logistic regression models assessing the association between user characteristics and likelihood of responding to follow up. Age, gender, and education were all significantly associated with likelihood of responding to follow up in both unadjusted and adjusted models. Users who were older, female, and had post-16 educational qualifications were more likely to respond to follow up. Current smoking status was significantly associated with the likelihood of responding to follow up in the unadjusted model, but not in the adjusted model. Users who were not current smokers were more likely to respond to follow up when other sociodemographic variables, AUDIT score, and treatment group were not adjusted for.

Table 1. User characteristics, engagement, response to follow up, and extent of alcohol reduction (N=672).

User characteristics	Statistics
Age (years), mean (SD)	39.2 (10.9)
Gender	
Female, n (%)	377 (56.1)
Employment status	
Employed, n (%)	581 (86.5)
Ethnicity	
White, n (%)	640 (95.2)
Education	
Post-16 years, n (%)	484 (72.0)
Current smoker	
Yes, n (%)	165 (24.6)
Past week alcohol consumption in units, mean (SD)	39.9 (27.3)
AUDIT[b] score, mean (SD)	19.1 (6.6)
Engagement measures (n=672)	
Number of sessions, median (IQR[a])	5 (2-17)
Time on app in min:s, median (IQR)	17:14 (8:53-37:19)
Number of days used, median (IQR)	4 (2-13)
Percentage of available screens viewed, mean (SD)	39.0 (13.3)
Response to follow up measure (n=672)	
Completion of 1-month follow up, n (%)	179 (26.6)
Extent of alcohol reduction in those followed up (n=179)	
Reduction in past week alcohol consumption in units, mean (SD)	14.3 (24.1)

[a]IQR: interquartile range.

[b]AUDIT: Alcohol Use Disorder Identification Test.

Table 2. The effect of user characteristics on measures of engagement (number of sessions and time on app).

User characteristics	Sessions, median (IQR[a])	Unadjusted simple linear regression		Adjusted[b] multiple regression		Time on app (min), median (IQR)	Unadjusted simple linear regression		Adjusted[b] multiple regression	
		B[c] (95% CI)	*P* value	B (95% CI)	*P* value		B (95% CI)	*P* value	B (95% CI)	*P* value
Age (years)	—[d]	.02 (0.02 to 0.03)	<.001	.03 (0.02 to 0.03)	<.001	—	.03 (0.02 to 0.03)	<.001	.03 (0.02 to 0.03)	<.001
Gender										
Male (reference; n=295)	7 (2 to 18)	—	—	—	—	16 (8 to 34)	—	—	—	—
Female (n=377)	5 (2 to 15)	.14 (−0.05 to 0.33)	.14	.12 (−0.07 to 0.31)	.21	18 (9 to 41)	.13 (−0.02 to 0.29)	.09	.11 (−0.03 to 0.26)	.13
Employment status										
Unemployed (reference; n=91)	5 (1 to 15)	—	—	—	—	15 (9 to 37)	—	—	—	—
Employed (n=581)	6 (2 to 17)	.23 (−0.04 to 0.51)	.10	.27 (−0.01 to 0.54)	.06	18 (9 to 37)	.06 (−0.16 to 0.28)	.61	.12 (−0.10 to 0.34)	.27
Ethnicity										
White (reference; n=640)	5 (2 to 17)	—	—	—	—	17 (9 to 38)	—	—	—	—
Not white (n=32)	6 (2 to 16)	−.09 (−0.54 to 0.35)	.68	.01 (−0.43 to 0.45)	.96	17 (10 to 27)	−.06 (−0.42 to 0.29)	.74	.04 (−0.30 to 0.39)	.81
Educational qualification										
Pre-16 (reference; n=188)	4 (2 to 13)	—	—	—	—	15 (8 to 34)	—	—	—	—
Post-16 (n=484)	6 (2 to 18)	.31 (0.10 to 0.52)	.004	.36 (0.15 to 0.57)	<.001	18 (9 to 41)	.12 (−0.05 to 0.28)	.18	.18 (0.02 to 0.16)	.03
Current smoker										
Yes (reference; n=165)	4 (2 to 16)	—	—	—	—	15 (8 to 29)	—	—	—	—
No (n=507)	6 (2 to 18)	.20 (−0.02 to 0.42)	.08	.01 (−0.22 to 0.23)	.96	18 (9 to 39)	.16 (−0.01 to 0.34)	.07	−.02 (−0.19 to 0.16)	.84
Past week alcohol consumption (units)[e]	—	0 (0)	.570	0 (0)	.75	—	0 (0)	.47	0 (0)	.37
Alcohol use (AUDIT[f] score)	—	−.01 (−0.02 to 0)	.19	−.01 (−0.02 to 0.01)	.42	—	0 (–0.01 to 0.01)	.64	0 (–0.01 to 0.01)	.94

[a]IQR: interquartile range.

[b]Adjusted for all sociodemographic variables, AUDIT score, and treatment group (unless otherwise specified).

[c]Unstandardized regression coefficient.

[d]Not applicable.

[e]Adjusted for all sociodemographic variables and treatment group (not AUDIT score).

[f]AUDIT: Alcohol Use Disorder Identification Test.

Table 3. The effect of user characteristics on measures of engagement (number of days used and % of screens viewed).

User characteristics	Days used, median (IQR[a])	Unadjusted simple linear regression		Adjusted[b] multiple regression		Screens viewed, % mean (SD)	Unadjusted simple linear regression		Adjusted[b] multiple regression	
		B[c] (95% CI)	*P* value	B (95% CI)	*P* value		B (95% CI)	*P* value	B (95% CI)	*P* value
Age (years)	—[d]	.02 (0.01 to 0.03)	<.001	.02 (0.02 to 0.03)	<.001	—	.25 (0.16 to 0.34)	<.001	.28 (0.19 to 0.38)	<.001
Gender										
Male (reference; n=295)	4 (2 to 11)	—	—	—	—	37.7 (13.54)	—	—	—	—
Female (n=377)	5 (2 to 13)	.12 (−0.05 to 0.30)	.16	.10 (−0.07 to 0.27)	.25	40.0 (13.54)	2.29 (0.27 to 4.32)	.03	1.99 (0.01 to 3.97)	.049
Employment status										
Unemployed (reference; n=91)	3 (1 to 11)	—	—	—	—	38.6 (13.97)	—	—	—	—
Employed (n=581)	4 (2 to 13)	.16 (−0.09 to 0.41)	.22	.18 (−0.08 to 0.43)	.17	39.0 (13.22)	.37 (−2.58 to 3.32)	.80	.07 (−2.88 to 3.02)	.96
Ethnicity										
White (reference; n=640)	4 (2 to 13)	—	—	—	—	39.1 (13.29)	—	—	—	—
Not white (n=32)	4 (1 to 13)	−.13 (−0.54 to 0.27)	.52	−.04 (−0.44 to 0.36)	.83	36.9 (13.84)	−2.12 (−6.86 to 2.61)	.38	−1.48 (−6.16 to 3.19)	.53
Education										
Pre-16 years (reference; n=188)	3 (1 to 9)	—	—	—	—	36.7 (13.27)	—	—	—	—
Post-16 years (n=484)	5 (2 to 14)	.23 (0.04 to 0.42)	.02	.28 (0.10 to 0.47)	.003	39.8 (13.24)	3.15 (0.91 to 5.38)	.006	4.04 (1.85 to 6.24)	<.001
Current smoker										
Yes (reference; n=165)	3 (1 to 12)	—	—	—	—	37.8 (12.72)	—	—	—	—
No (n=507)	5 (2 to 13)	.20 (0 to 0.39)	.05	.02 (−0.02 to 0.01)	.85	39.4 (13.49)	1.57 (−0.77 to 3.91)	.19	−.05 (−2.40 to 2.30)	.97
Past week alcohol consumption (units)[e]	—	0 (–0.01 to 0)	.33	0 (0)	.42	—	0 (−0.04 to 0.04)	.98	0 (−0.03 to 0.04)	.87
Alcohol use (AUDIT[f] score)	—	−.01 (−0.02 to 0)	.13	−.01 (−0.02 to 0.01)	.28	—	0 (−0.15 to 0.16)	.99	.02 (−0.13 to 0.17)	.82

[a]IQR: interquartile range.

[b]Adjusted for all sociodemographic variables, AUDIT score, and treatment group (unless otherwise specified).

[c]Unstandardized regression coefficient.

[d]Not applicable.

[e]Adjusted for all sociodemographic variables and treatment group (not AUDIT score).

[f]AUDIT: Alcohol Use Disorders Identification Test.

Table 4. The effect of user characteristics on response to follow up.

User characteristics	Completed follow up, n (%)	Unadjusted simple logistic regression		Adjusted[a] multiple logistic regression	
		OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value
Age (years)	—[c]	1.04 (1.02-1.05)	<.001	1.04 (1.02-1.06)	<.001
Gender					
Male (reference; n=295)	63 (21.4)	—	—	—	—
Female (n=377)	116 (30.8)	1.64 (1.15-2.34)	.006	1.58 (1.09-2.29)	.02
Employment status					
Unemployed (reference; n=91)	31 (34.1)	—	—	—	—
Employed (n=581)	148 (25.5)	0.66 (0.42-1.07)	.09	0.66 (0.40-1.12)	.12
Ethnicity					
White (reference; n=640)	172 (26.9)	—	—	—	—
Not white (n=32)	7 (21.9)	0.76 (0.30-1.70)	.53	0.74 (0.28-1.73)	.51
Educational qualification					
Pre-16 (reference; n=188)	36 (19.1)	—	—	—	—
Post-16 (n=484)	143 (29.5)	1.77 (1.18-2.70)	.007	2.11 (1.38-3.29)	<.001
Current smoker					
Yes (reference; n=165)	34 (20.6)	—	—	—	—
No (n=507)	145 (28.6)	1.54 (1.02-2.38)	.045	1.23 (0.79-1.95)	.37
Past week alcohol consumption (units)[d]	—	1.00 (0.99-1.01)	.92	1.00 (1.00-1.01)	.56
Alcohol use (AUDIT[e] score)	—	1.00 (0.97-1.02)	.72	1.00 (0.97-1.03)	.95

[a]Adjusted for all sociodemographic variables, AUDIT score, and treatment group (unless otherwise specified).

[b]OR: odds ratio.

[c]Not applicable.

[d]Adjusted for all sociodemographic variables and treatment group (not AUDIT score).

[e]AUDIT: Alcohol Use Disorder Identification Test.

Associations Between User Characteristics and Extent of Alcohol Reduction at Follow Up Among Those Followed Up

Table 5 reports the unadjusted and adjusted linear regression models assessing the association between user characteristics and extent of alcohol reduction among those followed up. Past week alcohol consumption and AUDIT score (at baseline) were significantly positively associated with extent of alcohol reduction, with those having a higher past week alcohol consumption and greater AUDIT scores at baseline reducing their alcohol consumption to a greater extent at follow up. No sociodemographic user characteristics were significantly associated with extent of alcohol reduction among those followed up.

Associations Between Measures of Engagement and Likelihood of Response to Follow Up

Table 6 reports the unadjusted and adjusted logistic regression models assessing the association between engagement and response to follow up. All engagement measures were significantly associated with likelihood of response to follow up, whereby greater engagement increased the likelihood of responding to follow up. This held true for the adjusted model when the known predictors of response to follow up were adjusted for.

Associations Between Measures of Engagement and Extent of Alcohol Reduction at Follow Up Among Those Followed Up

Table 7 reports the unadjusted and adjusted linear regression models assessing the association between engagement and extent of alcohol reduction. No association between engagement and the extent of alcohol reduction was detected in any of the unadjusted models or the adjusted model.

Sensitivity Analyses

Sensitivity analyses were conducted in which the engagement measures were dichotomized into high or low groups based on their median score (except for percentage of available screens viewed, which was dichotomized based on the mean score). The pattern of results remained the same.

Table 5. The effect of user characteristics on extent of alcohol reduction.

User characteristics	Mean (SD)	Unadjusted linear regression		Adjusted[a] multiple regression	
		B[b] (95% CI)	*P* value	B (95% CI)	*P* value
Age (years)	—[c]	−.13 (−0.44 to 0.18)	.42	−.11 (−19.71 to 27.33)	.53
Gender					
Male (reference; n=63)	14.4 (26.25)	—	—	—	—
Female (n=116)	14.2 (22.97)	−.20 (−7.67 to 7.26)	.96	.51 (−6.95 to 7.96)	.89
Employment status					
Unemployed (reference; n=31)	10.9 (21.4)	—	—	—	—
Employed (n=148)	15.0 (24.63)	4.05 (−5.35 to 13.45)	.40	5.07 (−4.73 to 14.89)	.31
Ethnicity					
White (reference; n=172)	14.5 (24.28)	—	—	—	—
Not white (n=7)	8.6 (19.91)	−5.92 (−24.29 to 12.45)	.53	−7.17 (−25.84 to 11.49)	.45
Educational qualification					
Pre-16 (reference; n=36)	17.2 (26.49)	—	—	—	—
Post-16 (n=143)	13.6 (23.51)	−3.58 (−12.46 to 5.30)	.43	−2.96 (−11.93 to 6.01)	.52
Current smoker					
Yes (reference; n=34)	15.9 (25.42)	—	—	—	—
No (n=145)	13.9 (23.86)	−1.97 (−11.05 to 7.11)	.67	.45 (−8.91 to 9.81)	.93
Past week alcohol consumption (units)[d]	—	.49 (0.37 to 0.61)	<.001	.49 (0.37 to 0.62)	<.001
Alcohol use (AUDIT[e] score)	—	1.01 (0.46 to 1.55)	<.001	.98 (0.40 to 1.55)	<.001

[a]Adjusted for all sociodemographic variables, AUDIT score, and treatment group (unless otherwise specified).

[b]Unstandardized regression coefficient.

[c]Not applicable.

[d]Adjusted for all sociodemographic variables and treatment group (not AUDIT score).

[e]AUDIT: Alcohol Use Disorder Identification Test.

Table 6. The association between engagement and response to follow up.

Engagement measures	Statistics	Unadjusted simple logistic regression		Adjusted[a] multiple logistic regression	
		OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value
Sessions, median (IQR[c])					
Did not respond (reference)	4 (2-11)	—[d]	—	—	—
Responded	19 (8-32)	1.08 (1.06-1.10)	<.001	1.08 (1.06-1.09)	<.001
Time on app, median (IQR)					
Did not respond (reference)	13.2 (7.4-26.0)	—	—	—	—
Responded	35.1 (18.9-70.9)	1.02 (1.02-1.03)	<.001	1.02 (1.02-1.03)	<.001
Days used, median (IQR)					
Did not respond (reference)	3 (1-8)	—	—	—	—
Responded	14 (6-24)	1.14 (1.11-1.17)	<.001	1.13 (1.11- 1.16)	<.001
Percentage of available screens viewed, mean (SD)					
Did not respond (reference)	35.6 (12.14)	—	—	—	—
Responded	48.3 (11.92)	1.09 (1.07-1.11)	<.001	1.09 (1.07-1.11)	<.001

[a]Adjusted for treatment group, age, gender, and education group (as significant predictors of response to follow up).

[b]OR: odds ratio.

[c]IQR: interquartile range.

[d]Not applicable.

Table 7. The association between engagement and extent of alcohol reduction.

Engagement measures	Unadjusted linear regression		Adjusted[a] multiple linear regression	
	B[b] (95% CI)	*P* value	B (95% CI)	*P* value
Sessions	−.16 (−0.37 to 0.05)	.13	−.14 (−0.34 to 0.07)	.19
Time on app	−.06 (−0.13 to 0.01)	.08	−.06 (−0.13 to 0.01)	.07
Days used	−.18 (−0.57 to 0.21)	.37	−.10 (−0.48 to 0.28)	.61
Available screens viewed	−.07 (−0.37 to 0.23)	.66	−.09 (−0.39 to 0.22)	.58

[a]Adjusted for treatment group and baseline AUDIT score (as a significant predictor of extent of alcohol reduction).

[b]Unstandardized regression coefficient.

Discussion

Summary of Principal Findings

Engagement and Response to Follow Up

Users who were older and had post-16 educational qualifications engaged with the Drink Less app to a greater extent, which was indicated by number of sessions, time on app, number of days used, and percentage of available screens viewed. Female users viewed a significantly greater percentage of available screens compared with male users. Users who were older, female, and had post-16 educational qualifications were also significantly more likely to respond to follow up. In line with previous literature from Web-based alcohol interventions [16,19-21], users who were female, older, and with post-16 educational qualifications engaged to a greater extent and were more likely to respond to follow up. This suggests that there are similarities in the user characteristics that are predictors of engagement and response to follow up between app- and Web-based alcohol interventions. However, this study found that gender was only associated with the percentage of available screens viewed (ie, depth of use) and was not associated with more typical measures of engagement such as amount or frequency of use.

All 4 measures of engagement were positively associated with the likelihood of responding to follow up, and this association remained when adjusting for the user characteristics that were significant predictors of response to follow up, which replicated previous findings [15,16]. There was no evidence that drinking characteristics were associated with engagement or the likelihood of response to follow up. This contradicts previous literature that has found that drinking characteristics are positively associated with engagement and negatively associated with response to follow up in Web-based alcohol interventions. A possible explanation could be the delivery modality (ie, app vs website) [33] or the intervention content. For example, Drink Less relies on a quick on-boarding process and involves no contact with health care professionals, which may have resulted in users feeling less stigmatized and more likely to return

irrespective of their drinking status. Alternatively, the content of the Drink Less app (eg, a game, a drinking diary) might have been sufficiently rewarding to promote engagement irrespective of drinking. Due to the lack of existing evidence relating to app-based interventions, it is not possible to draw any firm conclusions as to what may have caused the difference between the findings from this study and the association found between drinking characteristics and engagement and response to follow up in Web-based alcohol interventions.

Extent of Alcohol Reduction

Past week alcohol consumption and AUDIT score were both positively associated with the extent of alcohol reduction among those followed up. None of the sociodemographic characteristics were associated with the extent of alcohol reduction. There has been a lack of research on the predictors of alcohol reduction in the general population, particularly in app-based interventions. Our study found that drinking characteristics were positively associated with the extent of alcohol reduction. This is likely explained by regression to the mean, as measurements that differ substantially from the true mean tend to be followed by measurements closer to the true mean [40]. Further research is needed to elucidate whether this observation is driven by regression to the mean or whether app-based interventions tend to be more effective for people with higher levels of alcohol consumption. This could be tested with a confirmatory RCT comparing an optimized version of the app with a no-treatment control group to examine whether the same patterns of associations (ie, between drinking characteristics and alcohol reduction) are also observed in the control group.

No associations were detected between the engagement measures and the extent of alcohol reduction among those followed up, and these results were robust to a sensitivity analysis with the engagement measures as dichotomous variables. This means that we were not able to determine whether there is a threshold level of engagement with the app that would achieve users' intended reduction in alcohol consumption. These findings conflict previous findings of a positive association between engagement measures and successful behavior change [30-32,41]. This difference may be because of a different methodology used; in this study, we only analyzed the subsample of participants who completed the follow-up questionnaire, whereas many studies use an intent-to-treat approach [30,31,41]. Another possibility is that people use the Drink Less app in different ways because of the complexity of the app's design, meaning that threshold level of engagement differs across users. The strong association reported in previous studies might be driven by people who are unsuccessful in their behavior change, disengage with the app, and then do not complete the follow up. Nevertheless, a dose-response relationship between engagement and alcohol reduction would still be expected among those users who responded to follow up. The finding that engagement is not related to successful behavior change through a dose-response function is consistent with the findings from the factorial RCT that certain combinations of Drink Less modules were more effective than others [7]. An unplanned analysis found that greater engagement with the app mediated the effect of the self-monitoring module on reduction in AUDIT score for those users who received the combination of self-monitoring and action planning [42]. This mediation effect suggests that an engagement dose-response effect may depend on the intervention module. It is also possible that the threshold level of engagement for the intended outcomes of Drink Less was relatively low for all users (compared with other DBCIs), and a ceiling effect may have played a role in not detecting an overall dose-response effect between the extent of engagement and alcohol reduction.

Implications

Tailored strategies for younger male users with lower educational qualifications, who tend to have lower levels of engagement and response to follow up, could be codeveloped with these users to improve engagement and response to follow up. Users who were older and had post-16 educational qualifications engaged with the app to a greater extent in terms of number of sessions, time spent on the app, the number of days it was used for, and the depth of their use. The app was not designed for a specific age group (other than the adult population) and involved user testing with participants from disadvantaged groups who typically have poorer Web-based literacy to ensure it was usable and acceptable to these groups [34]. A possible explanation for the difference in engagement based on the user's age is that different age groups might differ in the ways in which they tend to engage with apps more generally (eg, younger users being less willing to spend a lot of time on apps). Future research should use the data available from alcohol reduction apps to investigate whether there are different user typologies and if these are categorized by age.

The finding that engagement measures were not associated with the extent of alcohol reduction suggests that engagement measures should not be used as a proxy for behavior change and that greater levels of engagement are not necessarily required to achieve a desired change in behavior. Therefore, tailored strategies for improving engagement and response to follow up will not necessarily result in the desired behavior change.

Strengths and Limitations

To our knowledge, this is the first study to investigate the predictors of engagement, response to follow up, and extent of alcohol reduction in an app-based intervention. Drink Less is freely available in the UK iTunes App Store and users were not directly recruited for a trial; instead, they downloaded the app and were then recruited to the trial. Therefore, this sample has high ecological validity and represents the real-world situation for most users of behavior change apps. This study had a modest sample size and could be repeated with a larger sample to assess whether the findings are replicable. A limitation of this study is that the measures of engagement used were summative and, therefore, could not be used to assess more specific patterns of engagement (eg, the order in which users engaged with the app's different components), which future research should look to investigate.

Conclusions

Users of an alcohol reduction app who were older and had post-16 educational qualifications engaged to a greater extent.

These characteristics and being female predicted users being more likely to respond to a follow-up questionnaire 1 month later. Higher baseline levels of alcohol consumption were predictive of a greater extent of alcohol reduction, but were not predictive of engagement or response to follow up. Engagement measures were significantly associated with response to follow up, in line with the law of attrition. Engagement measures were not associated with the extent of alcohol reduction, which suggests that there is no dose-response effect of the Drink Less app.

Acknowledgments

CG, IT, JB, and RW are funded by Cancer Research UK (CRUK: C1417/A22962). OP is funded by a grant from Bupa under its partnership with University College London. SM's contribution is funded by Cancer Research UK (CRUK) and the National Institute for Health Research (NIHR) School for Public Health Research (SPHR). Drink Less was funded by the NIHR SPHR, the UK Centre for Tobacco and Alcohol Studies (UKCTAS), the Society for the Study of Addiction, and CRUK. The views expressed are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. The research team is part of the UKCTAS, a UK Clinical Research Collaboration Public Health Research Centre of Excellence. Funding from the Medical Research Council, British Heart Foundation, Cancer Research UK, Economic and Social Research Council, and the National Institute for Health Research under the auspices of the UK Clinical Research Collaboration is gratefully acknowledged. The funders played no role in the design, conduct, or analysis of the study, or in the interpretation or reporting of study findings.

Conflicts of Interest

JB has received unrestricted research grants from Pfizer related to smoking cessation. RW has received research funding and has undertaken consultancy for companies that manufacture smoking cessation medications. CG, OP, IT, and SM have no competing interests.

References

1. World Health Organization. Global status report on alcohol and health, 2014. 2014. URL: http://www.who.int/sorry/ [accessed 2018-10-31] [WebCite Cache ID 73a2uD8O2]
2. Burton R, Henn C, Lavoie D, O'Connor R, Perkins C, Sweeney K, et al. A rapid evidence review of the effectiveness and cost-effectiveness of alcohol control policies: an English perspective. Lancet 2017 Apr 15;389(10078):1558-1580. [doi: 10.1016/S0140-6736(16)32420-5] [Medline: 27919442]
3. Anderson P. Alcohol in Europe: a public health perspective. A report for the European Commission. London: Institute of Alcohol Studies; 2006:1-432.
4. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev 2017 Dec 25;9:CD011479. [doi: 10.1002/14651858.CD011479.pub2] [Medline: 28944453]
5. Reinert DF, Allen JP. The alcohol use disorders identification test: an update of research findings. Alcohol Clin Exp Res 2007 Feb;31(2):185-199. [doi: 10.1111/j.1530-0277.2006.00295.x] [Medline: 17250609]
6. Garnett C, Crane D, Michie S, West R, Brown J. Evaluating the effectiveness of a smartphone app to reduce excessive alcohol consumption: protocol for a factorial randomised control trial. BMC Public Health 2016 Dec 8;16:536 [FREE Full text] [doi: 10.1186/s12889-016-3140-8] [Medline: 27392430]
7. 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]
8. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
9. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
10. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
11. 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]
12. Shadish W, Luellen J. Attrition. In: Everitt B, Howell DC, editors. Encyclopedia of Statistics in Behavioral Science. New Jersey: Wiley-Blackwell; 2005.

13. Bull S, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. AIDS Care 2004 Nov;16(8):931-943. [doi: 10.1080/09540120412331292507] [Medline: 15511725]
14. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. J Med Internet Res 2009 Apr 3;11(2):e9 [FREE Full text] [doi: 10.2196/jmir.1052] [Medline: 19403465]
15. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011 Feb 14;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
16. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013 Aug 30;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
17. Swartz LH, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
18. Blankers M, Smit ES, van der Pol P, de Vries H, Hoving C, van Laar M. The missing=smoking assumption: a fallacy in internet-based smoking cessation trials? Nicotine Tob Res 2016 Jan;18(1):25-33. [Medline: 25744969]
19. Bewick BM, West R, Gill J, O'May F, Mulhern B, Barkham M, et al. Providing web-based feedback and social norms information to reduce student alcohol intake: a multisite investigation. J Med Internet Res 2010 Dec 19;12(5):e59 [FREE Full text] [doi: 10.2196/jmir.1461] [Medline: 21169171]
20. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. J Med Internet Res 2007 May 8;9(2):e10 [FREE Full text] [doi: 10.2196/jmir.9.2.e10] [Medline: 17513281]
21. Radtke T, Ostergaard M, Cooke R, Scholz U. Web-based alcohol intervention: study of systematic attrition of heavy drinkers. J Med Internet Res 2017 Jun 28;19(6):e217 [FREE Full text] [doi: 10.2196/jmir.6780] [Medline: 28659251]
22. Bewick BM, Trusler K, Mulhern B, Barkham M, Hill AJ. The feasibility and effectiveness of a web-based personalised feedback and social norms alcohol intervention in UK university students: a randomised control trial. Addict Behav 2008 Sep;33(9):1192-1198. [doi: 10.1016/j.addbeh.2008.05.002] [Medline: 18554819]
23. Bewick BM, West RM, Barkham M, Mulhern B, Marlow R, Traviss G, et al. The effectiveness of a Web-based personalized feedback and social norms alcohol intervention on United Kingdom university students: randomized controlled trial. J Med Internet Res 2013 Jul 24;15(7):e137 [FREE Full text] [doi: 10.2196/jmir.2581] [Medline: 23883616]
24. Bingham CR, Barretto AI, Walton MA, Bryant CM, Shope JT, Raghunathan TE. Efficacy of a web-based, tailored, alcohol prevention/intervention program for college students: initial findings. J Am Coll Health 2010;58(4):349-356. [doi: 10.1080/07448480903501178] [Medline: 20159758]
25. Riper H, Kramer J, Conijn B, Smit F, Schippers G, Cuijpers P. Translating effective web-based self-help for problem drinking into the real world. Alcohol Clin Exp Res 2009 Aug;33(8):1401-1408. [doi: 10.1111/j.1530-0277.2009.00970.x] [Medline: 19413646]
26. Riper H, Kramer J, Keuken M, Smit F, Schippers G, Cuijpers P. Predicting successful treatment outcome of web-based self-help for problem drinkers: secondary analysis from a randomized controlled trial. J Med Internet Res 2008 Nov 22;10(4):e46 [FREE Full text] [doi: 10.2196/jmir.1102] [Medline: 19033150]
27. Garnett C, Crane D, West R, Michie S, Brown J, Winstock A. User characteristics of a smartphone app to reduce alcohol consumption. Transl Behav Med 2017 Dec;7(4):845-853 [FREE Full text] [doi: 10.1007/s13142-017-0477-1] [Medline: 28315201]
28. Beard E, Brown J, Kaner E, West R, Michie S. Predictors of and reasons for attempts to reduce alcohol intake: a population survey of adults in England. PLoS One 2017;12(3):e0173458 [FREE Full text] [doi: 10.1371/journal.pone.0173458] [Medline: 28278218]
29. Yardley L, Spring BJ, Riper H, Morrison LG, Crane D, 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]
30. Alexander GL, McClure JB, Calvi JH, Divine GW, Stopponi MA, Rolnick SJ, MENU Choices Team. A randomized clinical trial evaluating online interventions to improve fruit and vegetable consumption. Am J Public Health 2010 Feb;100(2):319-326 [FREE Full text] [doi: 10.2105/AJPH.2008.154468] [Medline: 20019315]
31. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216 [FREE Full text] [doi: 10.1080/14622200500055319] [Medline: 16036277]
32. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
33. Morrison LG, Geraghty AW, Lloyd S, Goodman N, Michaelides DT, Hargood C, et al. Comparing usage of a web and app stress management intervention: An observational study. Internet Interv 2018 Apr 21;12:74-82 [FREE Full text] [doi: 10.1016/j.invent.2018.03.006] [Medline: 30135771]

34. Crane D, Garnett C, Brown J, West R, Michie S. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: think aloud and interview studies. Front Public Health 2017 Apr 3;5:39 [FREE Full text] [doi: 10.3389/fpubh.2017.00039] [Medline: 28421175]
35. Garnett C, Crane D, West R, Brown J, Michie S. The development of Drink Less: an alcohol reduction smartphone app for excessive drinkers. Transl Behav Med 2018 May 4. [doi: 10.1093/tbm/iby043] [Medline: 29733406]
36. iTunes, Apple Inc. Drink Less on iTunes store URL: https://itunes.apple.com/gb/app/drink-less-get-help-reducing/id1020579244?mt=8 [accessed 2018-08-28] [WebCite Cache ID 720R3EpBJ]
37. Google Analytics. Google Inc. How a web session is defined in Analytics URL: https://support.google.com/analytics/answer/2731565?hl=en [accessed 2018-08-28] [WebCite Cache ID 720R5KJJK]
38. Garnett C. Open Science Framework. 2018. Drink Less: Predictors of engagement, follow up and extent of alcohol reduction URL: https://osf.io/htfrp/ [WebCite Cache ID 73WsVD1Kj]
39. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
40. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol 2005 Feb;34(1):215-220. [doi: 10.1093/ije/dyh299] [Medline: 15333621]
41. Tate DF, Wing RR, Winett RA. Using internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 7;285(9):1172-1177. [Medline: 11231746]
42. Crane D. Development and evaluation of a smartphone app to reduce excessive alcohol consumption: Self-regulatory factors. London: University College London (Doctoral thesis); 2017.

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
AUDIT-C: Alcohol Use Disorders Identification Test-Consumption
CRUK: Cancer Research UK
DBCI: digital behavior change intervention
IQR: interquartile range
NIHR: National Institute for Health Research
OR: odds ratio
SPHR: School for Public Health Research
RCT: randomized controlled trial
UKCTAS: UK Centre for Tobacco and Alcohol Studies

Edited by G Eysenbach; submitted 04.07.18; peer-reviewed by L Morrison, R Robinson; comments to author 01.08.18; revised version received 28.08.18; accepted 28.08.18; published 14.12.18.

Please cite as:
Garnett C, Perski O, Tombor I, West R, Michie S, Brown J
Predictors of Engagement, Response to Follow Up, and Extent of Alcohol Reduction in Users of a Smartphone App (Drink Less): Secondary Analysis of a Factorial Randomized Controlled Trial
JMIR Mhealth Uhealth 2018;6(12):e11175
URL: https://mhealth.jmir.org/2018/12/e11175/
doi:10.2196/11175
PMID:30552081

Original Paper

A Comprehensive Digital Program for Smoking Cessation: Assessing Feasibility in a Single-Group Cohort Study

Heather Patrick[1], MA, PhD; Craig A Fujii[1], BSc; Debra B Glaser[1], EdD; David S Utley[1], MD; Jennifer D Marler[1], MD

Carrot Inc, Redwood City, CA, United States

Corresponding Author:
Heather Patrick, MA, PhD
Carrot Inc
1600 Seaport Boulevard, Suite 150
Redwood City, CA, 94063
United States
Phone: 1 650 216 9689
Email: heather@carrot.co

Abstract

Background: Cigarette smoking remains the leading cause of preventable death and disease worldwide. Evidence-based approaches are available, but few people access them. Although digital solutions offer great promise for population reach, few multicomponent programs exist. Pivot is a comprehensive digital solution combining a Food and Drug Administration–cleared carbon monoxide (CO) breath sensor; cigarette logging; a 6-phase, app-delivered smoking cessation program based on the US Clinical Practice Guidelines; and dedicated human coaching via text-based chat.

Objective: The purpose of this study was to assess program engagement, changes in attitudes toward smoking, self-reported changes in smoking behavior, and program acceptability for the initial phase of Pivot: Explore.

Methods: A total of 48 participants enrolled, and 41 completed the study. About half the participants (54%, 22/41) were men, and the mean age was 43 years. Most (85%, 35/41) were daily smokers and smoked an average of 12 cigarettes per day. Explore includes CO breath sensing, logging cigarettes in-app, learning via in-app activities, and dedicated human coaching through a text messaging interface. Participants completed surveys at baseline and exit assessing attitudes toward quitting including readiness, perceived difficulty, and confidence in quit success. At exit, participants also completed a survey of changes in smoking behavior and ratings of program acceptability.

Results: More than 80% of participants (34-39 of 41) took ≥1 CO breath sample each day, and more than 55% (23-27 of 41) took ≥5 samples each day. More than 65% of participants (27-34 of 41) logged ≥1 cigarette using the in-app logging feature each day. All 9 in-app activities had completion rates ≥80% (33-40 of 41). Response to coach-initiated outreach was also high, with all contacts receiving ≥73% (30-39 of 41) response. In matched pair analyses, significant positive changes in mean attitudes toward quitting (scale 1-10) were evident from baseline (T1) to study exit (T2), including increased readiness to quit (T1 mean=6.1, T2 mean=7.4, P=.005), lower perceived difficulty (T1 mean=3.7, T2 mean=5.6, P=.001), and greater expectations of success (T1 mean=4.5, T2 mean=6.5, P<.001). At exit, 78% (32/41) of participants reported decreasing the number of cigarettes smoked per day during the study. Participants rated program quality and satisfaction very high (mean ≥8 for all items).

Conclusions: These results support the feasibility and acceptability of the initial 9-day phase of Pivot: Explore. Participants had high levels of engagement with sensing, logging, learning, and coaching. Attitudes toward quitting improved significantly, and the majority of users indicated decreasing smoking behavior. Explore was designed to raise smoker awareness and motivation. Additional research is underway to assess how users progress through the full Pivot smoking cessation program and determine the program's effectiveness for achieving sustained cessation.

(JMIR Mhealth Uhealth 2018;6(12):e11708) doi:10.2196/11708

KEYWORDS

smoking cessation; cell phone; health promotion; carbon monoxide

Introduction

Background

Although cigarette smoking has declined considerably over the past 50 years, it remains the leading cause of preventable death and disease in the United States and worldwide. In the United States, approximately 36.5 million adults smoke cigarettes, and smoking-related illnesses are responsible for nearly 500,000 premature deaths annually [1]. Most people who smoke (70%) say they want to quit, and nearly half of them will make at least one quit attempt in any given year [2,3]. However, most smokers who try to quit (65%-75%) do so on their own and do not receive the support and assistance that would raise their chances of success. As a result, these unsupported quit attempts yield very low success rates (2%-5%) [4,5].

Traditionally, smoking cessation programs have been offered in-person or telephonically. In-person services are challenged by limited scalability, and appointment scheduling and transportation may create insurmountable participant burden, thus limiting reach [6-8]. Telephonic programs (ie, quitlines) are cost-effective and scalable, and they have been shown to increase the odds of successful smoking cessation [5,9,10]. However, these services have been underutilized [11].

Digital health interventions may help to alleviate these issues of reach, scalability, and access, particularly given the ubiquity of mobile phones across diverse populations. As of February 2018, 77% of US adults own and use a mobile phone, and 20% use their mobile phone as their primary means of accessing the internet. Mobile phone ownership is consistent across racial and ethnic groups, with 75% of blacks and 77% of both Hispanics and whites owning mobile phones. Dependency on mobile phones for internet access is higher among racial and ethnic minorities (35% Hispanics and 24% blacks vs 17% whites). Among lower-income populations (ie, those making less than US $30,000 per year), 67% own and use a mobile phone, and 31% of these individuals rely on their mobile phone for internet access [12].

In addition to their potential to improve reach, an emerging evidence base supports the efficacy of digital solutions for smoking cessation. A recent Cochrane review examined the literature including randomized or quasi-randomized trials using any type of mobile phone–based intervention for smoking cessation. A total of 12 studies were identified that included 6-month smoking cessation outcomes. Interventions were primarily text messaging based, although several included initial assessments or in-person visits along with text messaging. Results indicated that the evidence supports the positive impact of mobile phone–based smoking cessation interventions, most notably text messaging programs, on 6-month cessation outcomes [13]. Haskins et al conducted a systematic review to examine both the strength of the scientific evidence for smoking cessation mobile apps and the extent to which scientifically supported apps have been made commercially available [14]. Their review identified 6 apps with some level of scientific support, ranging from high quality (43%; exploratory pilot randomized controlled trials) to low quality (57%; acceptability or usability studies, feature-level analysis, or being grounded in an evidence-based approach but not subjected to a study). In the examination of commercially available smoking cessation apps, Haskins et al identified 177 unique apps relevant to smoking cessation in the App Store for iPhone and 139 in Google Play for Android. Only 3 of the 6 scientifically vetted apps were available in these app stores. Of these 3, only 2 were listed among the top apps by at least one app store.

Traditional smoking cessation programs delivered in-person or telephonically have demonstrated efficacy but limited reach and utilization. Digital solutions have greater potential for reach, and there is some evidence that they are efficacious, but few scientifically vetted apps have been designed for commercialization. Thus, their potential for reach may not be fully realized. In addition, for both traditional and digital smoking cessation programs, one of the early activities (often the first) involves setting a quit date and working on developing a quit plan. Although these are critical, evidence-based elements that all smoking cessation programs should include, requiring participants to select a quit date and actively work toward quitting from program entry may limit reach only to those who are ready to quit and have a reasonable degree of confidence in their ability to do so. There is an opportunity to develop and deliver interventions that include some *runway* before actively working toward a quit attempt, during which users are able to engage in self-exploration (eg, through self-monitoring) and reflect on how smoking fits into the bigger picture of their life. Programs that allow users to ease into quitting rather than starting with quitting may have a greater potential to reach the population of smokers who do not engage in other cessation programs and put them on the path to quitting.

Finally, to date, most digital solutions for smoking cessation have leveraged only 1 form of technology—typically text messaging or an app—which fails to capitalize on the breadth of technologies available to engage, motivate, and ultimately improve cessation rates among the population of people who continue to smoke. One exception to this has been carbon monoxide (CO) breath sampling. Several published studies [15-18], as well as expert opinion [19,20], suggest that digital sensors that provide individuals with their CO breath sample values can be educational and motivational and may lead to attitude changes—including increased interest in seeking a quit program. CO sampling is most commonly done in health clinics, using equipment that is not conducive to daily, real-time usage. A small, 10-participant study in the United Kingdom permitted smokers to self-administer breath samples to measure CO [15], but it was not integrated into an overall, evidence-based smoking cessation program.

The Pivot Program

Pivot is a commercial-grade program designed for delivery in the context of employee wellness programs and health plans. It represents a comprehensive digital solution that brings together (1) the first Food and Drug Administration (FDA)–cleared (with over-the-counter labeling) CO breath sensor, which communicates via Bluetooth with a mobile phone and app; (2) a 6-phase mobile app delivering the US Clinical Practice Guidelines for Treating Tobacco Use and Dependence [5] and developed in collaboration with a team of scientific

advisors representing some of the world leaders in tobacco control and smoking cessation; and (3) dedicated human coaching, delivered one-on-one through a digital text messaging interface.

Pivot is a year-long program designed to support users along the spectrum of quitting, from being unsure or on the fence about quitting to maintaining a smoke-free life. Pivot begins with *Explore*, which is designed for anyone who smokes, to raise awareness and interest in moving forward. In Explore, users take samples with the Pivot Breath Sensor, log cigarettes, get to know their coach, and complete daily activities to understand their smoking patterns and explore how smoking affects their lives. The second phase of Pivot is *Build*, which is tailored to users' readiness, motivation, and confidence. Build culminates with users setting a quit date and building a quit plan. Next is *Mobilize*, which provides opportunities for users to put into practice individual elements of their quit plan, one at a time, in preparation for quit day. The fourth phase of Pivot is *Quit*, which begins on the user's selected quit day and continues through the first week of living smoke-free. Quit incorporates a daily check-in feature to allow users to track their progress and set daily goals to reinforce the idea of quitting as a process. *Secure* is a natural extension of Quit and focuses on supporting users in developing internal, sustainable motivation to stay smoke-free for good. With continued coaching support, self-monitoring, and practice, Pivot's newly smoke-free users learn to navigate the challenges that come in the first few months after quitting. The final phase of Pivot—*Sustain* —focuses on maintenance. Users continue to build skills and confidence and receive personal coaching designed to prevent relapse, so that they can remain smoke-free.

Study Overview

The purpose of this study was to examine feasibility and acceptability of the initial, 9-day phase of the Pivot program—Explore. One of the defining elements of Explore is that unlike most cessation interventions (digital or otherwise), the program does not begin with setting a quit date and building a quit plan. Instead, Explore is designed for all smokers, whether they are ready to quit or not. To that end, Explore consists of 4 key components: (1) use of a CO breath sensor; (2) in-app cigarette logging; (3) activities that encourage self-exploration; and (4) interaction with a dedicated human coach through a text messaging interface. These components are well grounded in the literature, and Pivot expands on existing digital interventions by bringing these components together in a single, integrated solution. As noted above, CO breath sampling has been shown to be educational and motivational and may be particularly useful early on in cessation programs and particularly for those individuals who are not yet ready to quit smoking [19,20]. Unlike sensors that have traditionally been used in clinical settings, Pivot's portable, Bluetooth-enabled breath sensor allows for daily sampling and feedback. Self-monitoring via cigarette logging is a commonly used evidence-based behavior change strategy implemented in many cessation programs [5]. In Explore, participants can easily log cigarettes in the Pivot app and see visual displays of their smoking patterns, including the amount and times of day when they are most likely to smoke. The activities in Explore leverage principles of motivational interviewing [21] to help participants move from being unsure or ambivalent about quitting smoking to being ready to work toward quitting. For example, some activities allow participants to explore how smoking affects their lives in terms of time and financial costs, including creating an opportunity to connect to their broader values by considering how they might otherwise spend those resources if not on smoking. Other activities encourage participants to consider how smoking might serve a purpose in their lives, including identifying their reasons for smoking. Finally, in Explore, participants work with a dedicated human coach via a text messaging interface. Behavioral counseling is a pillar of evidence-based smoking cessation programs [5]. However, behavioral counseling is often underutilized because of challenges with scheduling, transportation, or both. In addition, many telephonic programs do not allow participants to interact with the same coach over time, thus limiting the potential for participants to establish a strong therapeutic relationship with their cessation counselor. In Explore—and throughout Pivot—coaching is designed to directly address these barriers by (1) mitigating the need for scheduling by providing asynchronous chat, which allows participants to respond to coach-initiated outreach at their convenience, (2) eliminating both transportation and scheduling requirements by allowing participants to message their coach anytime and anywhere, and (3) creating the opportunity to cultivate a strong coach-participant relationship by assigning participants a dedicated coach so they are working with the same person each time they connect with their Pivot coach.

Methods

Study Design

This was an open-label, single-group, pretest-posttest study of the initial, 9-day phase of the Pivot program—Explore. The study was conducted as a feasibility and acceptability study to examine levels of engagement with program elements (ie, sensing, logging, learning, and coaching), analyze changes in attitudes toward quitting smoking from baseline (T1) to study exit (T2), and describe self-reported changes in smoking behavior and satisfaction with Pivot.

Recruitment and Study Population

Participants in the greater San Francisco Bay Area were recruited via Web-based advertisements and a clinical study recruiter. Participants completed a telephone screener to determine eligibility and receive a description of the study. Eligibility criteria included being aged 27 to 57 years, being able to speak and read English, smoking 5 or more cigarettes per day, owning and using a mobile phone (iPhone 5 and above, operating system iOS 9.0 and above, or Android 4.4 and above, operating system Android 4.4 and above), and using at least 1 app on their mobile phone. All participants indicated that they worked ≥30 hours per week. Thus, all participants were benefits-eligible (ie, eligible to receive insurance, wellness, and other benefits through their employer). Participants represented a range of employment sectors, including sales, warehouse management, human resources, forklift operation, guest services, administrative or secretarial, and education or teaching. According to 2015 data from the Society for Human Resource

Management [22], approximately 70% of US employers offer some form of wellness benefit to employees—either through a stand-alone wellness program or through insurance benefits. These wellness benefits are offered across a range of employment sectors, including those represented by this participant population. Thus, this sample is representative of the Pivot target market. Participants did not have to indicate an intent to quit smoking as a condition of study participation. Interested, eligible participants were invited to attend an on-site study appointment.

Figure 1 provides the CONSORT diagram for the study. As shown, 49 potential participants attended the study site visit and provided informed consent. One participant did not have a compatible phone; thus, 48 participants were enrolled, and 41 participants completed the study. Overall, 7 participants who were lost to follow-up were either not reachable after at least three contact attempts or did not agree to return. Moreover, 2 participants stated they smoked 5 or more cigarettes per day during the phone screening; however, they indicated that they smoked 4 cigarettes per day during registration. It is possible that these participants decreased smoking from screening to day 1.

As this was a feasibility and acceptability study, the purpose was to examine how people engaged with this first phase of the Pivot program, before developing the full Pivot experience, to inform the development of later phases of the program and to understand how and whether this first phase of the program was associated with shifts in attitude or behavior that might suggest the potential for participants to (1) engage in later phases of the Pivot program and (2) attempt to quit or quit smoking. As hypothesis testing was not a primary aim of this study, we did not conduct power analyses to determine sample size. Rather, we recruited and enrolled a sample size that allowed us to glean the necessary insights for further program development and additional research and ceased enrollment once we reached saturation. This is consistent with emerging standards for feasibility and pilot work [23].

Consent and Ethical Approval

All participants provided written and oral informed consent before participation. The study protocol was reviewed and approved by Solutions Institutional Review Board (Little Rock, AR, USA), Protocol #2017/04/12.

Procedure

During the on-site appointment, participants provided informed consent and then used a study laptop to complete the Web-based study registration and a baseline questionnaire to assess smoking history (ie, age when the participant started smoking, number of years smoked), number of cigarettes smoked per day, and attitudes toward quitting. They were then provided with the CO breath sensor and instructions to self-train on the use of the sensor and to download the Pivot app on their mobile phone. The CO breath sensor is FDA-cleared for single-user use by cigarette smokers in smoking cessation programs to inform the user about how breath CO levels are affected by smoking behavior. During the 9-day study period, participants were asked to engage with Explore, which comprised all features of the Pivot program: sensing, logging, learning, and coaching.

Sensing

Participants were instructed to use the CO breath sensor to complete hourly breath samples while awake. In the Pivot app, they were able to view their breath sample results, including the CO levels of each breath sample. Figure 2 shows the CO breath sensor and how it is used.

Figure 1. Consort (Consolidated standards of reporting trials) flow diagram of participants.

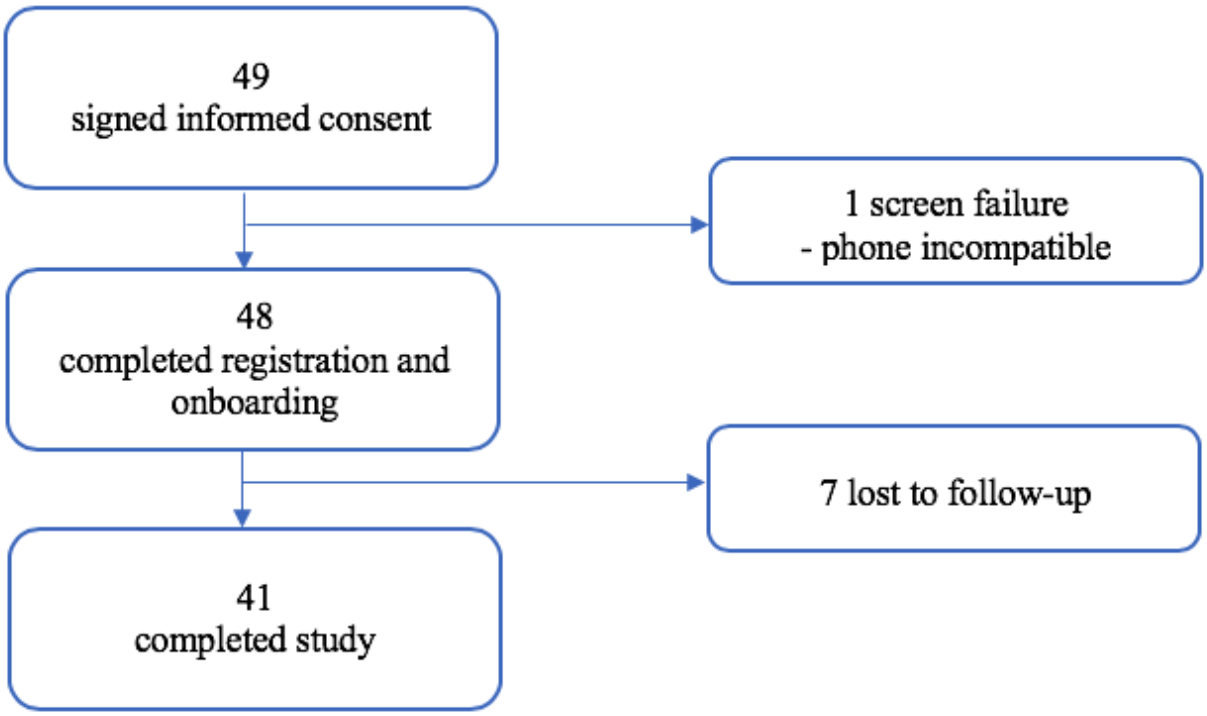


Figure 2. The Pivot carbon monoxide (CO) breath sensor.

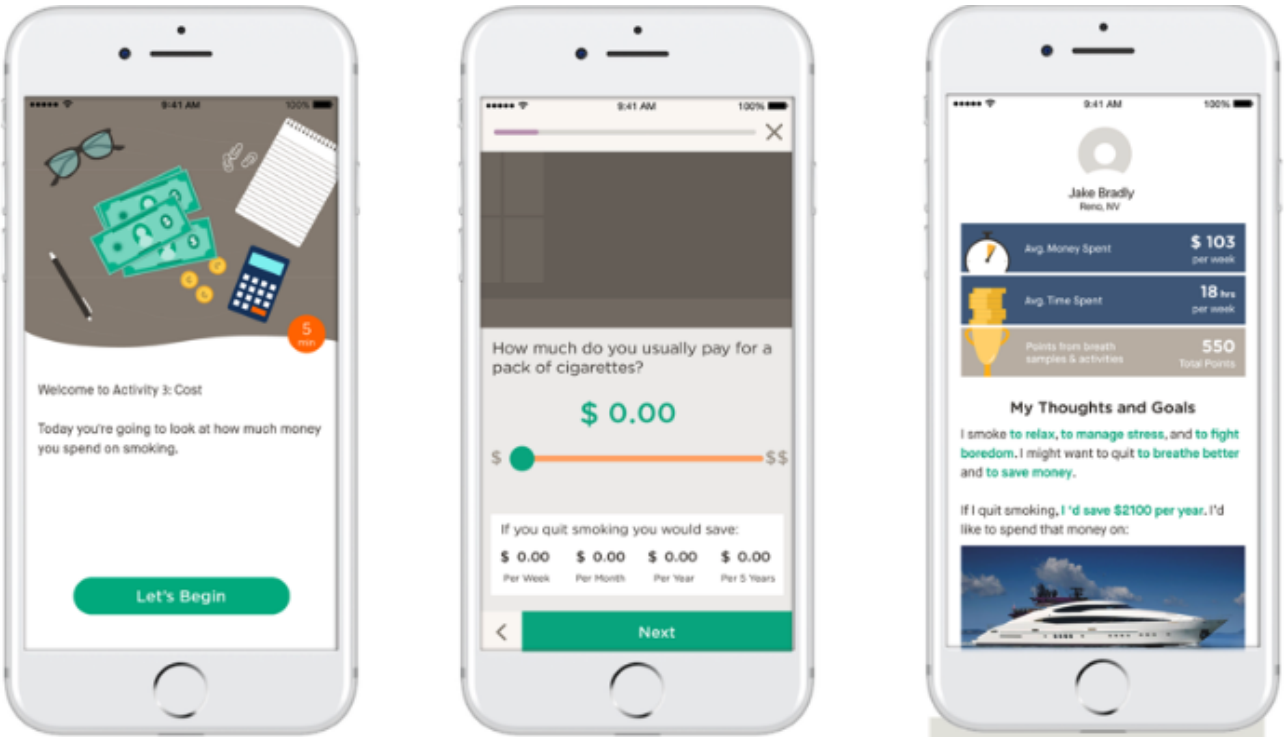
Logging

The Pivot app includes a cigarette-logging feature, and participants were asked to log all cigarettes smoked and were encouraged to log as soon after smoking as possible. The purpose of logging was to facilitate awareness of smoking behavior. The Pivot app provides participants with a visual display of smoking trends, including information about number, time, and location of cigarettes smoked.

Learning

The Explore phase of Pivot consists of 9 activities. One activity was *unlocked* each day, and participants received a daily push notification letting them know when a new activity had become available. The first 2 activities in Explore focus on CO and provide an overview of what CO is, how it is related to smoking, and why measuring it with the CO breath sensor is important. The remaining activities provided additional opportunities for participants to learn about how smoking fits into and influences their daily lives. This included a cost calculator, identifying both reasons for smoking and reasons for considering quitting, assessing level of addiction, exploring one's household influences, a calculation of time spent smoking, self-reflection on the importance of and confidence about quitting, and a summary of the participant's experience and activity in Explore. Figure 3 provides an example of one of the activities from Explore.

Figure 3. Example Pivot activity.



Coaching

Participants interacted with their dedicated coach via an asynchronous, text messaging interface. Coaches initiated 4 contacts during Explore: 1 after participants completed onboarding, 1 on day 3 or 4, 1 on day 5 or 6, and the last on day 8 or 9, before participants exited the study. Coaching conversations focused on supporting participants in reflecting on their experience in Explore, including what they were learning from using the CO breath sensor, logging cigarettes, and completing activities.

On day 9, participants returned to the study site to complete the final activity in the Pivot app and the end-of-study survey. The end-of-study survey included measures of attitudes toward quitting, self-reported changes in smoking behavior, and satisfaction with Pivot. Participants received a US $100 stipend on completion of the study. Participants were provided the stipend to compensate them for time and transportation costs associated with attending 2 onsite study visits and for the time spent participating during the 9 days of the study. All participants received this stipend, regardless of the level of engagement with Pivot.

Outcome Variables and Measurement

Engagement

Program engagement was assessed across all 4 elements of Explore: (1) sensing: use of the CO breath sensor (average daily breath samples and percentage of study population using the breath sensor); (2) logging: use of the in-app cigarette logging feature (average number of cigarettes logged in-app per day and percentage of study population using the logging feature); (3) learning: completion of daily in-app activities; and (4) coaching: response to coach-initiated outreach via texting interface.

Attitudes Toward Quitting Smoking

Participants answered 3 items to assess attitudes toward quitting smoking at baseline (T1) and study exit (T2), selecting a number using a 1 to 10 scale. Items included: *If you were to quit smoking right now, how difficult do you think it would be to stay smoke-free?* (1=really hard; 10=really easy); *If you were to quit smoking right now, how successful would you be?* (1=not successful at all; 10=completely successful); and *How ready are you to quit smoking?* (1=not ready at all; 10=completely ready).

Self-Reported Smoking Behavior

At study exit, participants answered several questions regarding whether and how their smoking behavior had changed over the course of the study. These items included the following: Do you feel that the amount you smoke has changed since your first study visit, 9 days ago? (Yes or No); (If Yes) Do you feel the amount decreased? (Yes or No); (If Yes) How did you decrease it? Select all that apply (I increased my time between cigarettes; I smoked less of each cigarette; I smoked fewer cigarettes per day; Other); (If Yes) What was the change due to? Read all options and select the single best answer (Being prompted to submit breath samples hourly [first 7 days of study]; Realizing how much money I spend on smoking; Seeing *the [CO] guy* fill up with red; Tracking my CO levels; Realizing the time I spend smoking; Tracking my cigarettes; Other).

Satisfaction With Pivot

At study exit, participants answered 3 items to indicate their satisfaction with Pivot. All items were answered by selecting a number from 0 to 10, as described below. Items included: *How informative did you find the program?* (0=not at all informative, 10=very informative); *How would you rank your satisfaction with the product?* (0=not at all satisfied, 10=very satisfied); and *How likely are you to recommend this program to a friend?* (0=not at all likely, 10=very likely). Likelihood of recommending Pivot to a friend was converted to a net promoter score (NPS). NPS is an industry indicator of participant loyalty to a product or service. NPS was calculated by subtracting the percentage of respondents who answered 6 or lower (detractors) from the percentage of respondents who answered 9 or 10 (promoters). Finally, participants indicated how many times per day they would be willing to use the breath sensor as part of a smoking cessation program.

Statistical Analyses

Statistical analyses were conducted using data from the 41 participants who completed the study. For engagement, data were collected through the Pivot app to describe sensing (eg, average daily breath samples), in-app cigarette logging, and completion of in-app activities. Response to coach-initiated contact was measured via the text messaging interface to indicate whether a participant responded to a coach-initiated message sent as part of the 4-coach touchpoint protocol described above. Analyses were conducted to calculate the mean (SD) for normally distributed variables or median (interquartile range) values in instances of non-normally distributed variables. Analyses involving attitudes toward quitting smoking were conducted with matched-pair *t* tests using SAS Version 9.4 (SAS Institute, Cary, NC, USA) to determine whether statistically significant changes in attitudes occurred from baseline (T1) to study exit (T2). Statistical significance was set at $P<.05$. Self-reported change in smoking behavior was analyzed as the percentage of study participants who indicated that they changed their smoking from the beginning to the end of the study. Finally, satisfaction with Pivot was analyzed descriptively to reflect mean and SD of responses for items on degree of program informativeness and program satisfaction. NPS was calculated for the item reflecting likelihood of recommending Pivot to a friend.

Results

Per-protocol analyses were conducted using the 41 participants who completed the study. More than half of study participants (54%, 22/41) were men, and the average age of participants was 43 years (SD 9 years). Most participants (61%, 25/41) used Android mobile phones, and most participants (85%, 35/41) smoked daily. Baseline characteristics of smoking history and experience are presented in Table 1.

Table 1. Baseline characteristics of smoking history and experience (n=41).

Characteristics	Mean (SD)
Age when started smoking	21.0 (10.2)
Years smoking	21.4 (13.0)
Cigarettes smoked per day	12.2 (6.0)

Engagement

Sensing

Figure 4 presents average daily breath sensor use for all participants (n=41) across days 1 to 8 of the study. On day 9, participants returned to the study site and therefore did not have a full day available for sensor use. As shown, self-monitoring via the breath sensor was reasonably consistent throughout the study, ranging from a high of 8.1 breath samples on day 2 (the first full day of participation) to a low of 5.9 breath samples on day 8 (the last full day of participation). The interquartile range (25%-75%) was 2 to 12 samples per day.

Figure 5 shows, for each day of the study, the percent of the study population who used the breath sensor 0, 1 to 4, and 5 or more times on each day of the study. Overall, daily breath sensor use was quite high, with 83% to 95% of participants (34-39 of 41) using the breath sensor at least once daily.

Figure 4. Average daily breath sensor usage.

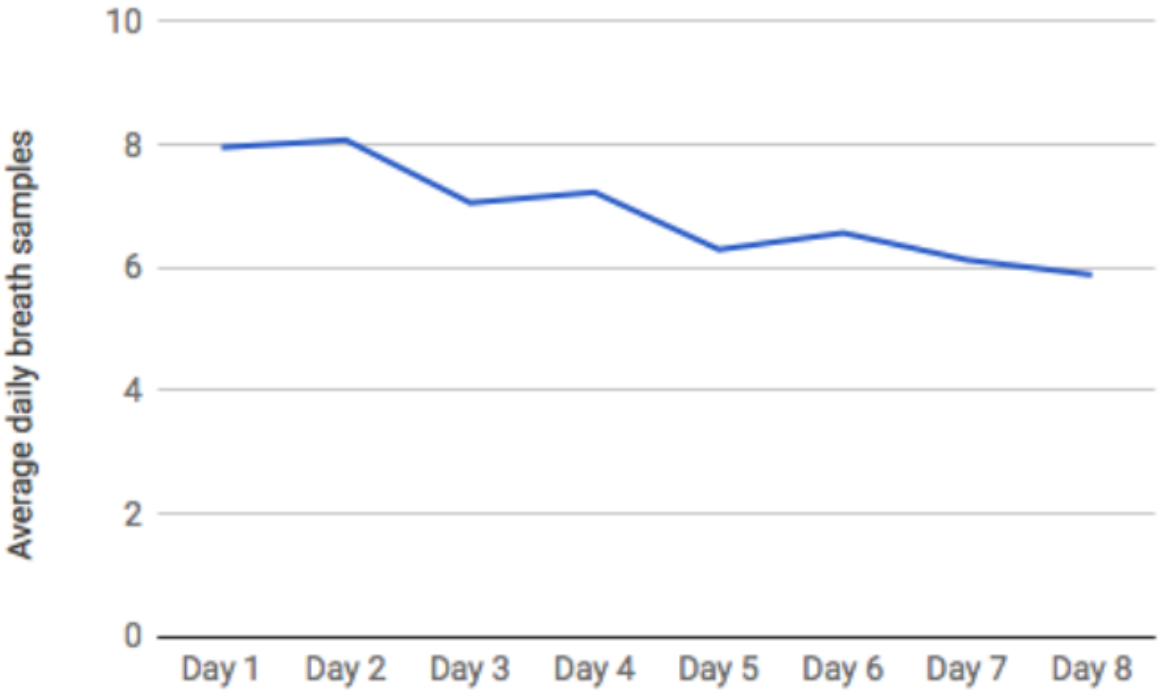
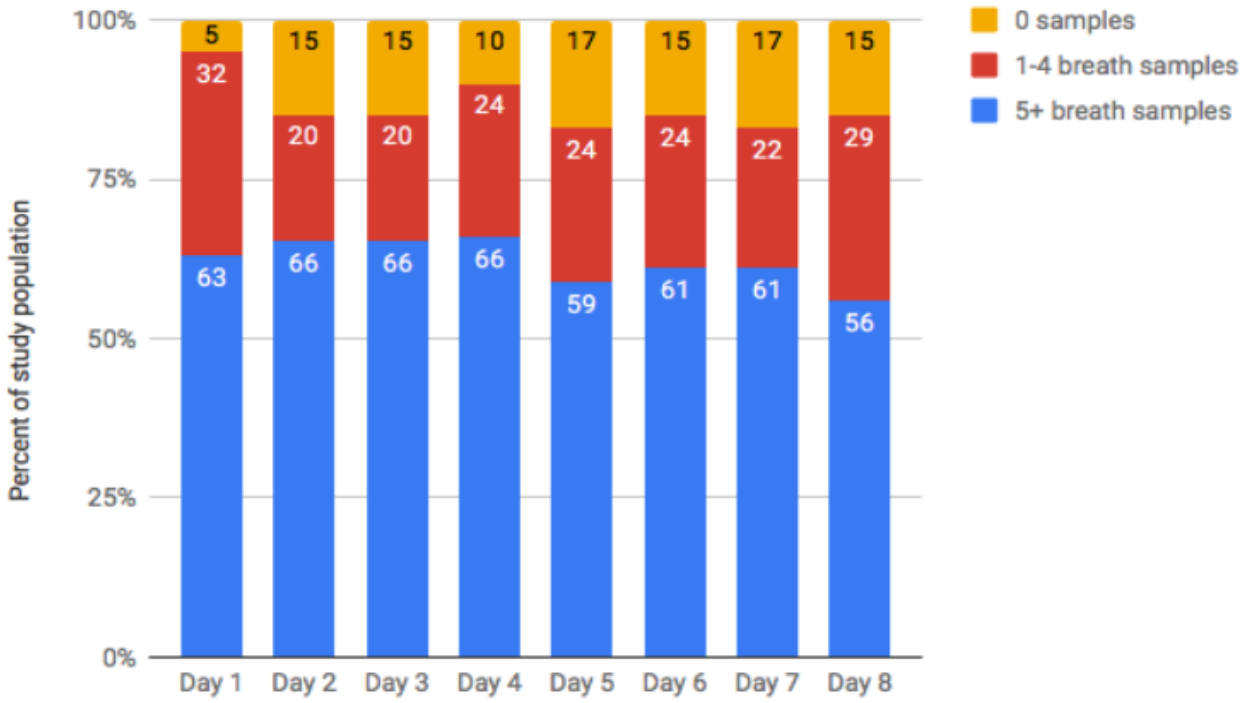


Figure 5. Percentage of participants using the breath sensor 0, 1 to 4, and more than 5 times each day.



Logging

Figure 6 presents the average number of cigarettes logged per day, using the in-app logging feature. Data include all participants (n=41) across days 1 to 8 of the study. As shown, beginning on day 2 (the first full day of study participation), use of the in-app logging feature was quite consistent, with participants logging 4.6 to 5.7 cigarettes per day, on average. The interquartile range (25%-75%) was 0 to 9 cigarettes logged per day. Figure 7 presents the percent of the study population who logged at least one cigarette each day using the in-app logging feature. For each day of the study, more than 65% of the study population (27-34 of 41) logged at least one cigarette using the in-app logging feature.

Figure 6. Average cigarettes logged in-app per person per day.

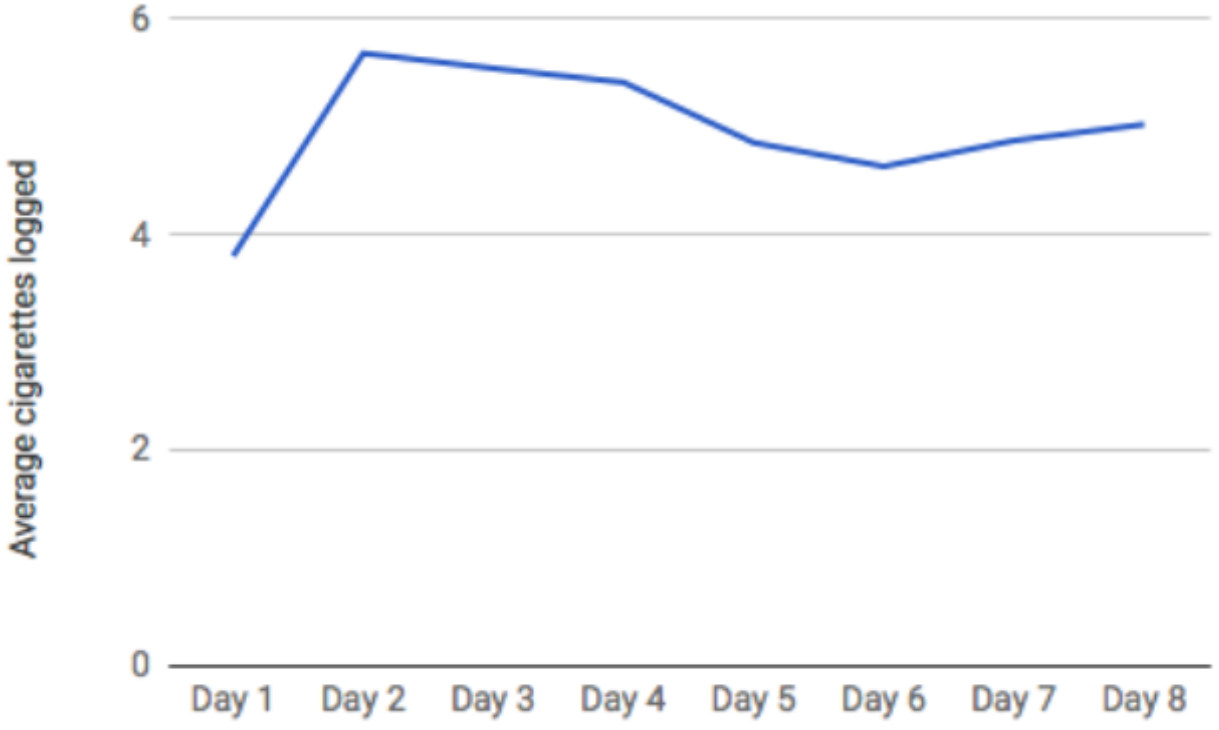
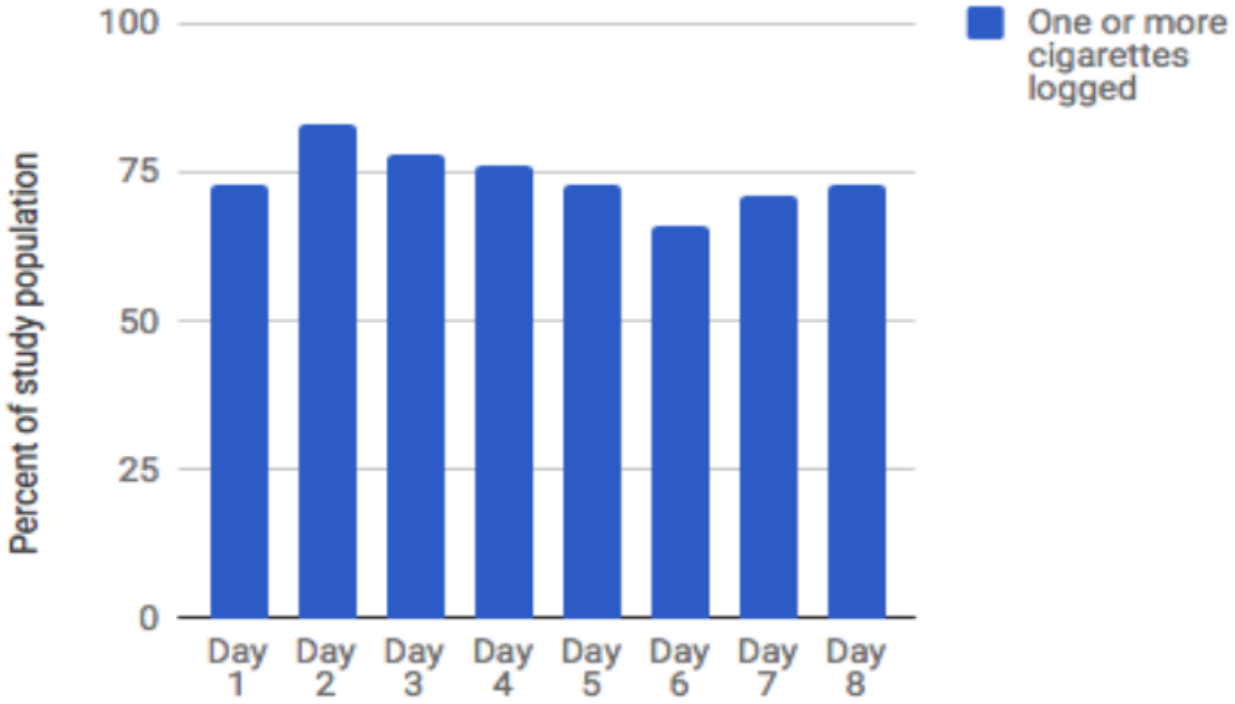


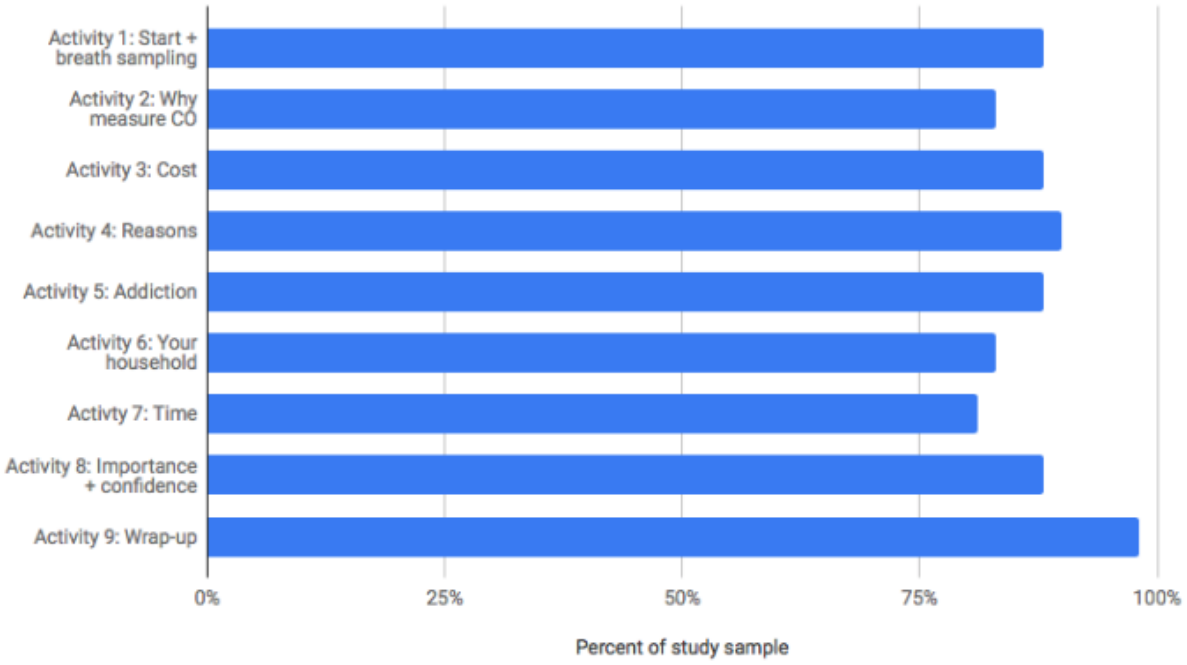
Figure 7. Percentage of study population logging at least one cigarette on each day.



Learning: Completion of In-App Activities

Figure 8 presents percent completion for each of the 9 daily, in-app activities. Participants who had not yet completed activity 9 when they attended the on-site study exit interview were asked to do so before exiting the study. Thus, this activity shows the highest completion rate. For all remaining activities, participants completed them on their own, within the context of their daily lives. As shown, completion rates were high, with each activity having a completion rate of 80% or higher (33-40 of 41).

Figure 8. Percent completion of in-app activities. CO: carbon monoxide.



Coaching Engagement

Table 2 presents data on engagement with coaching. As shown, each coach touchpoint had a participant response rate of 73% or greater (30-39 of 41). Coach-initiated touchpoints were conducted via a digital text messaging interface. The 2 columns on the right in Table 2 present mean number of messages per touchpoint from coaches and participants, respectively.

Table 2. Engagement with human coaching via digital text messaging interface.

Touchpoint	Participants responded, n (%)	Coach-sent messages, mean (SD)	Participant-sent messages, mean (SD)
1	39 (95)	9.8 (4.4)	6.7 (4.1)
2	31 (76)	8.6 (5.2)	6.0 (5.5)
3	32 (80)	7.8 (7.0)	4.7 (5.6)
4	30 (73)	6.7 (6.3)	3.9 (5.1)

Change in Attitudes Toward Quitting

Table 3 presents results from matched-pair analyses on changes in attitudes toward quitting smoking from baseline (T1) to end of study (T2). As shown, there were significant, positive changes in attitudes toward quitting smoking such that at study exit, participants expected lower difficulty in quitting, had greater confidence in success, and were more ready to quit.

Self-Reported Change in Smoking Behavior

More than three-fourths of the study population (78%, 32/41) indicated that their smoking behavior had changed from the beginning of the study to the end. Of participants who reported changing their smoking behavior, nearly all (31/32) indicated that they had decreased the amount that they smoked. The most common means by which participants decreased smoking included smoking fewer cigarettes per day (87%, 27/31) and increasing time between cigarettes (39%, 12/31). When asked what the change in smoking behavior was due to, the most common responses were tracking my cigarettes (45%, 14/31), tracking my CO levels (35%, 11/31), and being prompted to submit hourly breath samples (32%, 10/31).

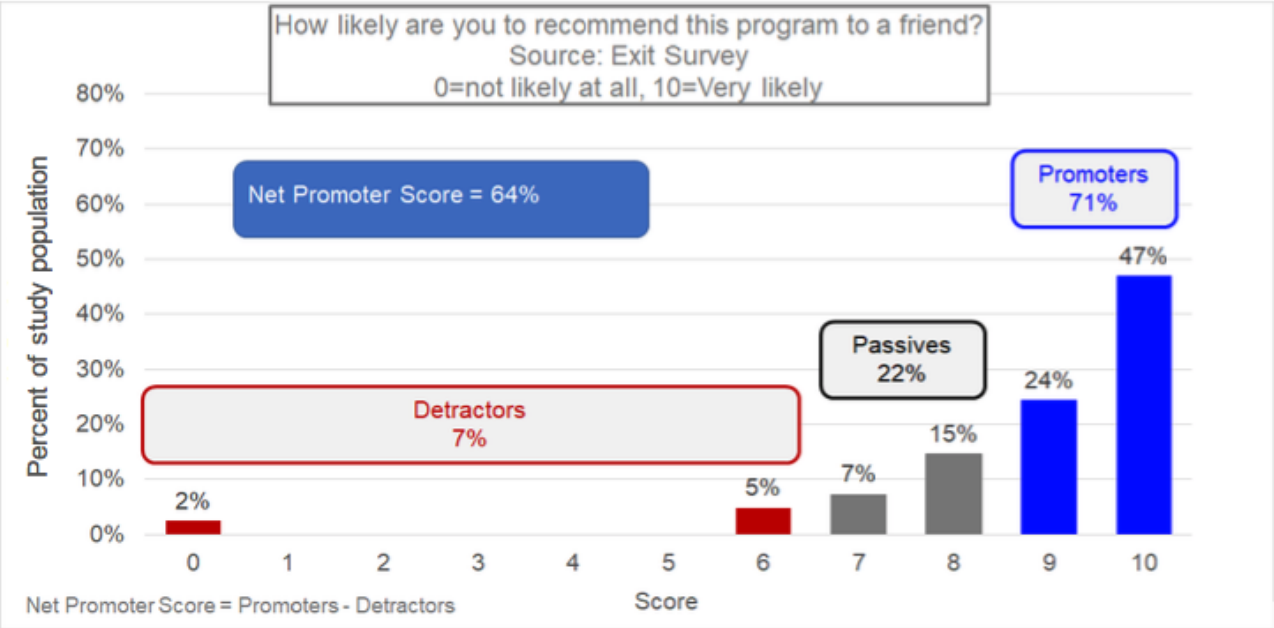
Satisfaction With Pivot

Overall, participants indicated that they thought the Pivot program was very informative (mean=8.4, SD=1.7) and that they were quite satisfied with the program overall (mean=8.5, SD=1.8). Figure 9 presents the distribution of responses to the question *How likely are you to recommend this program to a friend*, which was used to calculate NPS. As shown, the NPS score for the Explore phase of Pivot was 64. This is considered *excellent* using global NPS standards [24]. Most health apps do not report NPS publicly. As a point of reference from other health-relevant vendors, BlueStar (a product of Welldoc) reports an NPS of 70 [25], WebMD reports an NPS of 60 [26], and in a recent survey on NPS for primary care, patients indicated an NPS of 30 [27]. Finally, regarding willingness to use the breath sensor as part of a cessation program, on average, participants indicated they were willing to provide 10.3 breath samples per day. The interquartile range (25%-75%) was 6.5 to 12.0 samples per day.

Table 3. Changes in attitudes toward quitting smoking. For paired change analyses, positive numbers are favorable.

Attitude toward quitting	T1 (baseline), mean (SD)	T2 (end of study), mean (SD)	Paired change (SD)	*P* value
Difficulty	3.7 (3.2)	5.6 (2.7)	1.9 (3.4)	.001
Success	4.5 (2.7)	6.5 (2.6)	2.1 (3.0)	<.001
Readiness	6.1 (3.0)	7.4 (2.5)	1.2 (2.6)	.005

Figure 9. Net promoter score (NPS).



Discussion

Principal Findings

This was a feasibility study of the first phase of the Pivot program, Explore. Digital solutions offer one mechanism by which evidence-based approaches to smoking cessation can be disseminated at scale. Consistent with recommendations for features likely to increase program effectiveness, Pivot has been designed to offer the following: social context and support (coaching), multiple means of contact with the intervention (sensing, logging, learning, and coaching), tailoring (sensing and coaching), and self-management (sensing and logging) [28]. During the 9-day study, participants used all 4 program features: sensing via the personal mobile CO breath sensor, logging cigarettes, learning through completion of in-app activities, and coaching via a digital text messaging interface. They completed measures of attitudes toward quitting smoking at baseline and at the end of the study, and at the end of the study, answered questions about changes in smoking behavior and satisfaction with the program.

Program Engagement: Sensing, Logging, Learning, and Coaching

Across all program components, engagement was quite high. More than 80% of participants used the CO breath sensor at least once on each day of the study. At least two-thirds of participants used the in-app cigarette logging feature daily. Each of the 9 in-app activities had completion rates of 80% or higher, and responsiveness to each of the 4 coach-initiated touchpoints was 73% or higher.

This feasibility study provides initial support that participants are willing and able to engage with all components of Pivot's comprehensive evidence-based solution. This finding is encouraging, given the consistent dose-response relationship found in smoking cessation interventions, in which greater exposure to the intervention and in this case, multiple modes by which to have such exposure, is likely to improve cessation outcomes [5,28]. An important question to examine in future work is what constitutes *effective engagement* in Pivot (ie, the degree of engagement likely to yield the intended behavior change) [29]. Part of this will involve identifying the ideal amount and combination of engagement with each program component to yield clinical outcomes such as quit attempts and sustained cessation. It is also likely that different user profiles will emerge such that some components (or combinations of components) are particularly beneficial for different users based on user preferences, smoking history, readiness to change, and other characteristics. As we collect more data on user behavior and engagement with Explore and the remaining phases of Pivot, we will be able to examine different usage patterns. Different usage patterns may be related to baseline user characteristics such as smoking history and attitudes toward quitting, as well as subsequent quitting behavior including setting a quit date, building a quit plan, making quit attempts, and sustaining quits over time. This is a longer-term endeavor that will require much larger sample sizes and leveraging data collected surreptitiously through the app as well as user-provided data in the form of baseline and in-app surveys.

Changes in Attitudes and Self-Reported Behavior

The study also demonstrated positive, statistically significant improvements in three indicators of attitudes toward quitting smoking: increased readiness to quit, increased anticipation of success, and reduced perceptions of difficulty quitting. These attitude shifts are meaningful, considering the role of motivational factors (motivation to quit, confidence in quitting) in predicting quit attempts [30]. Well over half of participants demonstrated improvements in readiness (58%, 23/41), anticipation of success (73%, 29/41), and reduced perceptions of difficulty (65%, 26/41). In previous research, only 15% to 34% of participants demonstrated short-term changes in attitudes toward quitting over time frames ranging from 8 to 30 days [31-34]. It is possible that participants adopted other

strategies—such as the use of nicotine replacement therapy—to support changes in attitudes and behavior. However, nothing in the Explore phase of Pivot directly focuses on cigarette reduction, strategies for dealing with cravings or withdrawal, or the use of medications to support reducing or quitting smoking.

In addition, more than three-quarters of participants indicated that they had reduced the number of cigarettes they smoked during this 9-day experience. These findings are particularly promising, given that the focus of the first phase of Pivot is not intended to help users reduce or quit smoking. Rather, the purpose of the first 9 days of the program is to provide participants with an opportunity to learn about how smoking fits into their life and about how smoking and CO levels are linked. The fact that the experience of self-exploration and learning was associated with positive changes in how people feel about quitting smoking and with self-reported changes in smoking behavior suggests potential for the full Pivot program to lead to smoking cessation. In addition, because Pivot has been designed with this 9-day on-ramp that does not start with setting a quit date or working on a quit plan, Pivot may be uniquely positioned to appeal to smokers who otherwise would not participate in a cessation program.

Program Satisfaction and Acceptability

Satisfaction with and acceptability of Pivot was very high across multiple indicators including program informativeness, satisfaction, and NPS. As a metric of user loyalty, NPS is particularly important because Pivot is likely to be more effective with achieving clinical outcomes (ie, quit attempts and sustained cessation) for people who progress through later phases of the program. The fact that users indicated a high degree of loyalty for the first phase of Pivot shows promise for the potential longer-term *stickiness* of the program [24].

Limitations

This was a small feasibility study that examined only the first phase of a 6-phase smoking cessation program. Thus, we did not assess cessation outcomes. Participants were recruited from the San Francisco Bay Area, where policy, taxation, and social norms around smoking are more stringent than in other parts of the United States. In addition, given the influence of the technology sector in this area, participants were likely more tech-savvy than maybe expected elsewhere. Participants received a stipend (US $100) in exchange for their participation in the study. It is possible that this could have biased the findings. However, (1) the stipend was paid to them on completion of the final study assessment, regardless of their level of engagement in Pivot during the 9 days of the study, and (2) participants were told that the stipend was intended as a thank you for their time as a study participant and to cover transportation time and expenses for 2 trips to the study site. Thus, the stipend was not contingent on their levels of engagement with Pivot. Finally, although engagement across all components of the program was quite high, it is worth noting that utilization of the self-monitoring features (ie, sensing and logging) was somewhat lower than anticipated. With regard to the use of the CO breath sensor, participants were instructed to use the sensor hourly while awake; however, they used the sensor an average of 5.9 to 8.1 times per day. Use of the cigarette logging feature was also somewhat lower than anticipated. Participants reported smoking an average of 12.2 cigarettes per day at baseline and recorded 4.6 to 5.7 cigarettes daily using the in-app logging feature. Although possible, it seems unlikely that participants reduced smoking that drastically starting with the first full day of the study. Thus, we were unable to use cigarette logging as an indicator of smoking behavior and instead relied on participant's self-reported changes in smoking. This is similar to what others have found in a mobile app that includes cigarette-logging features, which have been used to support participants in cultivating awareness of behavioral patterns of smoking, rather than as an objective measure of cigarettes smoked [35]. A question on cigarettes smoked per day or the validated Fagerstrom test would have offered a more robust indication of behavioral shift. Despite the somewhat-lower-than-anticipated engagement with the self-monitoring features of Pivot, it is worth noting that using the breath sensor and tracking cigarettes were two of the most commonly noted features that participants felt contributed to their self-reported changes in smoking behavior.

Conclusions

Despite these limitations, the findings from this study provide support for the feasibility of the initial phase of Pivot: Explore. Pivot brings together a unique, comprehensive combination of technologies including the FDA–cleared CO breath sensor, evidence-based content presented through an engaging mobile app, and dedicated human coaching delivered via text messaging interface. Engagement across all program components was high as was program satisfaction and acceptability. In addition, the statistically significant, positive changes in attitudes toward smoking and self-reported changes in smoking behavior are promising, particularly because the first phase of Pivot does not directly address or promote smoking cessation. To confirm the promising results of these initial findings, additional research is underway to examine engagement with and progression through the full Pivot journey and evaluate program effectiveness for quit attempts and short- and long-term cessation outcomes. In addition, research is underway to better understand optimal program engagement and begin to identify and tailor on user profiles.

Acknowledgments

The authors would like to thank their scientific advisors Drs Judith (Jodi) Prochaska, Danielle E Ramo, and Sharon M Hall for their contributions to the development of Pivot. They also thank all their coaches for the clinical expertise and compassion they bring to these and all of their participants.

Authors' Contributions

JDM, DSU, CAF, and HP designed the study. JDM and CAF recruited participants and ran the study. CAF and DBG conducted data analyses. HP prepared the original draft of the manuscript. CAF, JDM, DBG, and DSU reviewed and provided comments on the manuscript before submission.

Conflicts of Interest

HP, CAF, JDM, DBG, and DSU are employees of Carrot Inc, the developer of the app and devices used in this study. DSU is the President and CEO of Carrot Inc and an investor in the company.

References

1. Centers for Disease Control and Prevention. 2018. Current cigarette smoking among adults in the United States URL: https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/ [WebCite Cache ID 70XnXzijw]
2. Centers for Disease Control and Prevention. Quitting smoking among adults--United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
3. National Center for Chronic Disease Prevention and Health Promotion. Centers for Disease Control and Prevention. 2018. The Health Consequences of Smoking-50 Years of Progress: A Report of the Surgeon General URL: http://www.cdc.gov/tobacco/quit_smoking/cessation/coverage/pdfs/coverage_tobacco_508_new.pdf [WebCite Cache ID 70XoMKNmL]
4. Centers for Disease Control and Prevention. 2014. Coverage for tobacco use cessation treatments URL: https://www.cdc.gov/tobacco/quit_smoking/cessation/pdfs/coverage.pdf [accessed 2018-06-29] [WebCite Cache ID 70XpMBdvJ]
5. Fiore MC, Jaén CR, Baker TB. Treating Tobacco Use and Dependence: 2008 Update--Clinical Practice Guidelines. Rockville, MD: US Department of Health and Human Services; 2008.
6. Schwartz RP, Gryczynski J, Mitchell SG, Gonzales A, Moseley A, Peterson TR, et al. Computerized versus in-person brief intervention for drug misuse: a randomized clinical trial. Addiction 2014 Jul;109(7):1091-1098 [FREE Full text] [doi: 10.1111/add.12502] [Medline: 24520906]
7. BinDhim NF, McGeechan K, Trevena L. Who uses smoking cessation apps? A feasibility study across three countries via smartphones. JMIR Mhealth Uhealth 2014 Feb 06;2(1):e4 [FREE Full text] [doi: 10.2196/mhealth.2841] [Medline: 25098439]
8. Kruse GR, Kelley JH, Linder JA, Park ER, Rigotti NA. Implementation of an electronic health record-based care management system to improve tobacco treatment. J Gen Intern Med 2012 Dec;27(12):1690-1696 [FREE Full text] [doi: 10.1007/s11606-012-2174-6] [Medline: 22865018]
9. Task Force on Community Preventive Services. Recommendations regarding interventions to reduce tobacco use and exposure to environmental tobacco smoke. Am J Prev Med 2001 Feb;20(2 Suppl):10-15. [Medline: 11173214]
10. Lichtenstein E, Zhu S, Tedeschi GJ. Smoking cessation quitlines: an underrecognized intervention success story. Am Psychol 2010;65(4):252-261 [FREE Full text] [doi: 10.1037/a0018598] [Medline: 20455619]
11. Schauer GL, Malarcher AM, Zhang L, Engstrom MC, Zhu S. Prevalence and correlates of quitline awareness and utilization in the United States: an update from the 2009-2010 National Adult Tobacco Survey. Nicotine Tob Res 2014 May;16(5):544-553. [doi: 10.1093/ntr/ntt181] [Medline: 24253378]
12. Pew Research Center. 2018. Mobile fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 70XprpjHp]
13. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
14. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Dec;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]
15. Beard E, West R. Pilot study of the use of personal carbon monoxide monitoring to achieve radical smoking reduction. J Smok Cessat 2012 Jul 13;7(01):12-17 [FREE Full text] [doi: 10.1017/jsc.2012.1]
16. Choi W, Kim C, Lee OG. Effects of brief smoking cessation education with expiratory carbon monoxide measurement on level of motivation to quit smoking. Korean J Fam Med 2013 May;34(3):190-198 [FREE Full text] [doi: 10.4082/kjfm.2013.34.3.190] [Medline: 23730486]
17. Risser NL, Belcher DW. Adding spirometry, carbon monoxide, and pulmonary symptom results to smoking cessation counseling: a randomized trial. J Gen Intern Med 1990;5(1):16-22. [Medline: 2405112]
18. Jamrozik K, Vessey M, Fowler G, Wald N, Parker G, Van Vunakis H. Controlled trial of three different antismoking interventions in general practice. Br Med J (Clin Res Ed) 1984 May 19;288(6429):1499-1503 [FREE Full text] [Medline: 6426618]
19. Fagerström KO. Assessment of the smoker who wants to quit. Monaldi Arch Chest Dis 2001 Apr;56(2):124-127. [Medline: 11499300]
20. Bittoun R. Carbon monoxide meter: the essential clinical tool—the 'Stethoscope'—of smoking cessation. J Smok Cessat 2012 Feb 21;3(02):69-70 [FREE Full text] [doi: 10.1375/jsc.3.2.69]

21. Lai DT, Cahill K, Qin Y, Tang JL. Motivational interviewing for smoking cessation. Cochrane Database Syst Rev 2010 Jan 20(1):CD006936. [doi: 10.1002/14651858.CD006936.pub2] [Medline: 20091612]
22. Valet V. Forbes. 2015. More Than Two-Thirds Of US Employers Currently Offer Wellness Programs, Study Says URL: https://www.forbes.com/sites/vickyvalet/2015/07/08/more-than-two-thirds-of-u-s-employers-currently-offer-wellness-programs-study-says/#168c66a2231d [accessed 2018-08-31] [WebCite Cache ID 72BoGd0jqampsaDampust1536195418227000ampusgAFQjCNELf4OZmZLq9dXCZ7o8TpuUXj2XA]
23. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010 Jul 16;10:67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
24. Schneider D, Berent M, Thomas R, Krosnick J. Van-haaften. 2018. Measuring customer satisfaction and loyalty: improving the Net-Promoter Score, 2008 URL: http://www.van-haaften.nl/images/documents/pdf/Measuring%20customer%20satisfaction%20and%20loyalty.pdf [accessed 2018-06-29] [WebCite Cache ID 70XqO5Yvc]
25. Welldoc. Results that Redefine Chronic URL: https://www.welldoc.com/ [accessed 2018-08-31] [WebCite Cache ID 72BtGUxLO]
26. Customer.guru. WEBMD Health corp net promoter score 2018 benchmarks URL: https://customer.guru/net-promoter-score/webmd-health-corp [accessed 2018-08-31] [WebCite Cache ID 72BtOSQEh]
27. Advisory. Is it fair to compare physicians' Net Promoter Score to Apple's? URL: https://www.advisory.com/research/market-innovation-center/the-growth-channel/2015/09/pcp-consumer-loyalty-survey [accessed 2018-08-31] [WebCite Cache ID 72Bthz4JdampsaDampust1536195418228000ampusgAFQjCNEZmuzrCVw5gRjg1kjMBqYsjtA]
28. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
29. 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]
30. Vangeli E, Stapleton J, Smit ES, Borland R, West R. Predictors of attempts to stop smoking and their success in adult general population samples: a systematic review. Addiction 2011 Dec;106(12):2110-2121. [doi: 10.1111/j.1360-0443.2011.03565.x] [Medline: 21752135]
31. Donovan R, Jones S, Holman CD, Corti B. Assessing the reliability of a stage of change scale. Health Educ Res 1998 Jun;13(2):285-291. [Medline: 10181026]
32. Etter JF, Sutton S. Assessing 'stage of change' in current and former smokers. Addiction 2002 Sep;97(9):1171-1182. [doi: 10.1046/j.1360-0443.2002.00198.x] [Medline: 12199833]
33. Hughes JR, Keely JP, Fagerstrom KO, Callas PW. Intentions to quit smoking change over short periods of time. Addict Behav 2005 May;30(4):653-662. [doi: 10.1016/j.addbeh.2004.08.011] [Medline: 15833571]
34. Peters EN, Hughes JR. The day-to-day process of stopping or reducing smoking: a prospective study of self-changers. Nicotine Tob Res 2009 Sep;11(9):1083-1092 [FREE Full text] [doi: 10.1093/ntr/ntp105] [Medline: 19561132]
35. Iacoviello BM, Steinerman JR, Klein DB, Silver TL, Berger AG, Luo SX, et al. Clickotine, a personalized smartphone app for smoking cessation: initial evaluation. JMIR Mhealth Uhealth 2017 Apr 25;5(4):e56 [FREE Full text] [doi: 10.2196/mhealth.7226] [Medline: 28442453]

Abbreviations

CO: carbon monoxide
FDA: Food and Drug Administration
NPS: net promoter score

Edited by G Eysenbach; submitted 26.07.18; peer-reviewed by D Christofferson, C Rey-Reñones; comments to author 16.08.18; revised version received 17.09.18; accepted 22.09.18; published 18.12.18.

Please cite as:
Patrick H, Fujii CA, Glaser DB, Utley DS, Marler JD
A Comprehensive Digital Program for Smoking Cessation: Assessing Feasibility in a Single-Group Cohort Study
JMIR Mhealth Uhealth 2018;6(12):e11708
URL: https://mhealth.jmir.org/2018/12/e11708/
doi:10.2196/11708
PMID:30563807

Original Paper

Coping Strategies and Social Support in a Mobile Phone Chat App Designed to Support Smoking Cessation: Qualitative Analysis

Esther Granado-Font[1,2], SRA; Carme Ferré-Grau[2], PhD; Cristina Rey-Reñones[2,3], PhD; Mariona Pons-Vigués[4,5,6], PhD; Enriqueta Pujol Ribera[4,5,6], PhD; Anna Berenguera[6], PhD; Maria Luisa Barrera-Uriarte[7,8], SRA; Josep Basora[8], MD; Araceli Valverde-Trillo[9], MD; Jordi Duch[10], PhD; Gemma Flores-Mateo[8,11], MD

[1]Centre d'Atenció Primària Horts de Miró (Reus - 4), Gerència d'Àmbit d'Atenció Primària Camp de Tarragona, Institut Català de la Salut, Tarragona, Spain

[2]Facultat d'Infermeria, Departament d'Infermeria, Universitat Rovira i Virgili, Tarragona, Spain

[3]Institut Català de la Salut, Unitat de Suport a la Recerca Tarragona-Reus, Institut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Reus, Spain

[4]Institut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Barcelona, Spain

[5]Universitat de Girona, Girona, Spain

[6]Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), Spain

[7]Centre d'Atenció Primària La Granja (Tarragona-2), Gerència d'Àmbit d'Atenció Primària Camp de Tarragona, Institut Català de la Salut, Torreforta (Tarragona), Spain

[8]Unitat de Suport a la Recerca Tarragona-Reus, Institut Universitari en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Reus, Spain

[9]Departament de Salut, Barcelona, Spain

[10]Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain

[11]Unitat d'Anàlisi i Qualitat, Xarxa Sanitària i Social Santa Tecla, Tarragona, Spain

Corresponding Author:
Cristina Rey-Reñones, PhD
Institut Català de la Salut
Unitat de Suport a la Recerca Tarragona-Reus
Institut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol)
Camí de Riudoms, 53-55
Reus, 43202
Spain
Phone: 34 977778515
Fax: 34 977778518
Email: crey.tgn.ics@gencat.cat

Abstract

Background: Smoking is one of the most significant factors contributing to low life expectancy, health inequalities, and illness at the worldwide scale. Smoking cessation attempts benefit from social support. Mobile phones have changed the way we communicate through the use of freely available message-oriented apps. Mobile app–based interventions for smoking cessation programs can provide interactive, supportive, and individually tailored interventions.

Objective: This study aimed to identify emotions, coping strategies, beliefs, values, and cognitive evaluations of smokers who are in the process of quitting, and to analyze online social support provided through the analysis of messages posted to a chat function integrated into a mobile app.

Methods: In this descriptive qualitative study, informants were smokers who participated in the chat of Tobbstop. The technique to generate information was documentary through messages collected from September 2014 through June 2016, specifically designed to support a smoking cessation intervention. A thematic content analysis of the messages applied 2 conceptual models: the Lazarus and Folkman model to assess participant's experiences and perceptions and the Cutrona model to evaluate online social support.

Results: During the study period, 11,788 text messages were posted to the chat by 101 users. The most frequent messages offered information and emotional support, and all the basic emotions were reported in the chat. The 3 most frequent coping strategies identified were physical activity, different types of treatment such as nicotine replacement, and humor. Beliefs about

quitting smoking included the inevitability of weight gain and the notion that not using any type of medications is better for smoking cessation. Health and family were the values more frequently described, followed by freedom. A smoke-free environment was perceived as important to successful smoking cessation. The social support group that was developed with the app offered mainly emotional and informational support.

Conclusions: Our analysis suggests that a chat integrated into a mobile app focused on supporting smoking cessation provides a useful tool for smokers who are in the process of quitting, by offering social support and a space to share concerns, information, or strategies.

(JMIR Mhealth Uhealth 2018;6(12):e11071) doi:10.2196/11071

KEYWORDS

primary health care; qualitative research; mobile apps; smoking cessation; social support; psychological adaptation

Introduction

Background

Smoking is one of the most significant factors contributing to low life expectancy, health inequalities, and illness at the worldwide scale [1]. Every year, tobacco kills approximately 6 million people and causes economic losses in the order of half a billion dollars. However, the deaths caused by smoking are the most preventable and, as a World Health Organization report points out, the impact of the tobacco epidemic can be reduced by using low-cost, high-efficiency means [1]. According to data from the 2014 European Health Survey, conducted in Spain by the National Institute of Statistics, 30.4% of men and 20.5 % of women are smokers, compared with the European mean of 21.9% and 15.1%, respectively [2].

Online Support for Smoking Cessation

The rise of internet use and mobile phones has introduced 2 key features in the way we communicate with each other: communications are now (1) ubiquitous, that is, you can talk with almost anyone anytime (24 hours a day) anywhere and (2) nearly instantaneous, as messages can be received and answered within seconds or minutes. Moreover, individuals can be connected at minimal cost, eliminating barriers to in-person participation in group programs such as childcare, disability, and employment responsibilities [3].

The current guidelines recommend that all smoking cessation programs incorporate some type of social support [4]. This may include social networks and mobile communication–based systems that provide a platform where those trying to quit smoking can share concerns and offer emotional support, useful advice, personal stories, and reinforcement during all the smoking cessation process [5]. Online support groups also offer a degree of anonymity that would not be possible in face-to-face communication, which may encourage individuals to openly discuss their experiences without fear of a negative reaction [6].

However, little is known about the efficiency and the importance of online support in smoking cessation programs. To our knowledge, previous studies published about this topic found that the support of social networks may be beneficial immediately when smokers want to quit and also during the first weeks of a smoking cessation program [3,7].

Conceptual Models

Lazarus and Folkman model defines the concept of stress by referring to the interrelationships that occur between a person and the context in which that individual finds himself or herself. The Lazarus and Folkman model may be transferred to the smoking cessation field to study psychological factors. Their model suggests that anxiety levels depend on the ability to handle external demands and internal evaluations that exceed the resources of the individual and on the strategies used to cope with them. This framework is appropriate for this study because smoking cessation is considered an important stressful factor. Although most smokers aged 18 years or older expressed a desire to quit and 52% had attempted to do so, only 6% of them had successfully quit at 12 months. Previous studies have found stress-induced craving response to be particularly important in smokers with high levels of nicotine dependence, who may be at greatest risk for cessation failure [8,9]. Stress-coping progams increase success in quitting smoking [8,9].

The Lazarus and Folkman model addresses 6 categories [10]: (1) emotions, for which our study applied the Ekman classification of primary emotions (joy, sadness, anger, disgust, fear, and surprise) [8]; (2) coping strategies, both task-oriented and emotion-oriented; (3) beliefs, defined as preexisting notions of reality, whether individually created or culturally shared, and in this case, referring to the smoking cessation process; (4) values, encompassing the objectives that express what is important to the individual and will help him or her to quit smoking; (5) cognitive evaluation, a process that determines the consequences a particular event will generate in an individual; and (6) social support, a coping resource whereby someone provides emotional, informative, and/or tangible support.

The discussion of emotions is a key element of online support groups. Cutrona and Suhr developed a coding scheme to classify social support behaviors as emotional, informational, self-esteem, social network, and tangible support [9]. They identified all 5 types of social support in online posts, with informational and emotional support most frequently observed. People who decide to quit smoking may benefit from having developed coping strategies to overcome the habit.

Tobbstop Trial

The Tobbstop trial was a multicenter randomized clinical trial (Registration: clinicaltrial.gov NCT01734421) carried out in Tarragona, Reus, and surrounding areas in Catalonia (Spain)

that aimed to assess the efficacy of a mobile phone app for smoking cessation. Smokers were recruited from primary health care centers and were randomized into 2 groups: (1) an intervention group that included access to the Tobbstop mobile app and the usual counseling about smoking cessation provided in primary health care consultations [11] and (2) a control group that received only the usual smoking cessation counseling.

This study analyzed one of the components of the Tobbstop app, a private chat that allowed study participants to communicate with each other [12]. The objectives of this study were to identify emotions, coping strategies, beliefs, and values, together with cognitive evaluation of smokers during the process of quitting, and to analyze online social support provided through messages posted to this chat.

Methods

Design

Descriptive qualitative study to identify the emotions, motivations, and perceived benefits that could be observed in daily experiences within the process of change experienced by people who used this chat function during the action phase of the change process.

Participants

Of the 309 participants randomly selected for the intervention group, 102 participated writing comments in the chat, constituting our study population. The sample was opportunistic [13]. Inclusion criteria were being adults (older than 18 years) with a motivation ≥6 points on the Richmond test [11], in the action phase according to Prochaska and DiClemente model of change [14], and who had an iOS or Android-based mobile phone.

The Prochaska and DiClemente model describes stages related to addictive behaviors in individuals trying to abandon substance use. The stages according to this model are precontemplation (denial a problem exists), contemplation (self-awareness of problem begins), preparation stage (individual starts making concrete plans to abandon substance use), action stage (reduction and cessation of smoking), and finally, a maintenance stage.

Description of the Mobile App

The Tobbstop app was designed to support participants during the first 3 months of the smoking cessation progress, with 3 main goals in mind: (1) to help individuals record their progress in the smoking cessation program; (2) to increase the user's knowledge about the problems related to smoking and the health benefits associated with smoking cessation; and (3) to provide distraction for moments of craving.

The Tobbstop app included 4 components: (1) a library with information about tobacco; (2) a private chat for study participants where they could ask for help, share concerns, or offer help to others; (3) a set of minigames designed specifically to entertain and educate participants; and (4) a progress registry to show the evolution of the participant's health throughout the treatment process. The app also included a panic button and consultation with an expert.

Technique to Generate Data

The technique to generate information was documentary through written text messages. During the Tobbstop study period (September 2014 to June 2016), 11,788 text messages were written in Catalan and Spanish by participants. These were downloaded into an Excel table for analysis, replacing personal information about the participants with identification codes that protected anonymity.

Analysis

A thematic content analysis of the messages posted in the chat was performed by 2 members of the research team (EGF and GFM) as follows: (1) an initial reading of all messages; (2) identification of relevant topics and text messages; (3) fragmentation of the texts into units of meaning; (4) codification of texts by topics; (5) creation of categories based on the Lazarus and Folkman and the Cutrona model, grouping the codes; and (6) interpretation of the meanings of each category. Analysis was conducted with the support of the ATLAS.ti 7 program.

Criteria of Rigor and Quality

To ensure the rigor and quality of the study, the following criteria of rigor suggested by Calderón were followed: epistemological and methodological adequacy, relevance, validity, and reflexivity [15]. The context, the characteristics of the participants, and the research process were described. The messages obtained were analyzed, and a period of reflection was carried out by 2 members of the research team.

Ethical Aspects

The study entitled "Efficacy of an application for mobile devices in smoking cessation in young people (Smart_Smoke): a cluster-randomized clinical trial" was approved by the ethics committee of Instituto de Investigación en Atención Primaria (IDIAP) Jordi Gol (P12 / 041). The app used was called Tobbstop.

Participants voluntarily agreed to participate and provided their signed informed consent. The research team coded the stored messages with an identification number to guarantee confidentiality and protection of the participants' identity. No names were used in the reported quotations.

Results

Table 1 shows the characteristics of the participants included. The results are structured in 2 blocks according to the Lazarus and Folkaman model, and Cutrona model categories (Figure 1).

Lazarus and Folkman Model

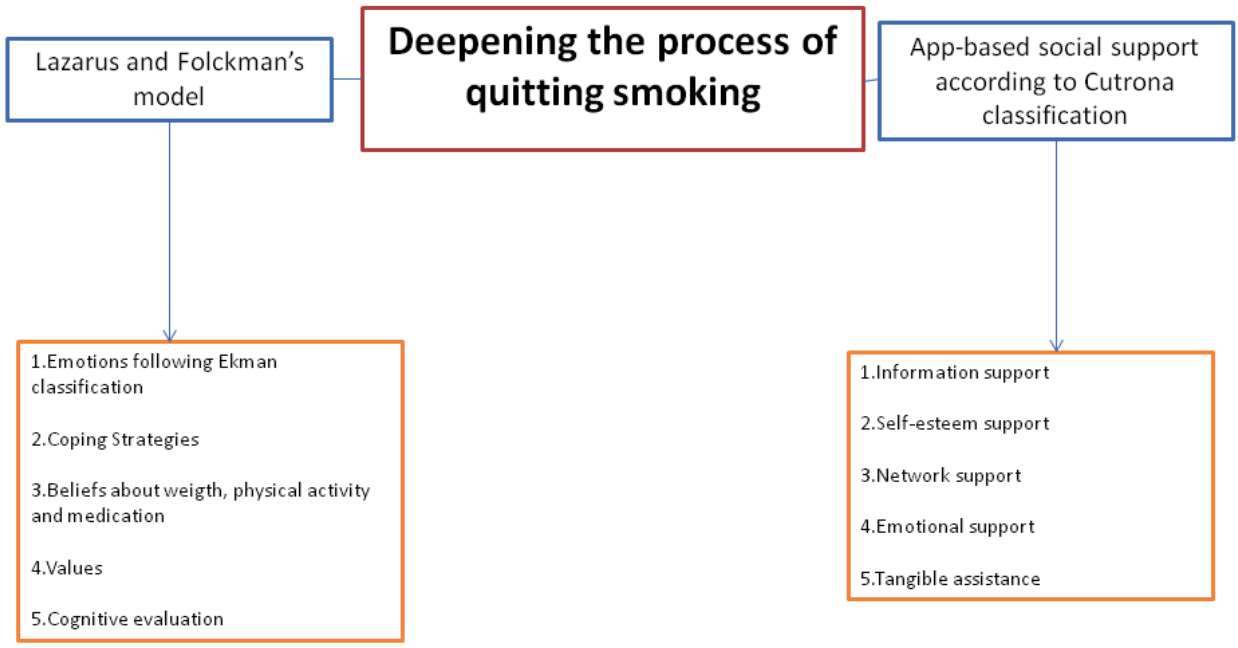
On the basis of Lazarus and Folkman model, analysis revealed the following 5 main categories:

Table 1. Sociodemographic characteristics of users of the Tobbstop chat (N=102).

Chat participants	Statistics
Sex, n (%)	
Male	59 (57.8)
Age (years), mean (SD)	45.3 (8.9)
Civil status, n (%)	
Single	21 (20.8)
Married	61 (59.2)
Widower	2 (2.5)
Divorced	18 (17.5)
Educational level, n (%)	
No schooling	0 (0.0)
Primary	26 (25.7)
Secondary	57 (56.4)
University or higher	17 (16.8)
Age started smoking (years), mean (SD)	16.1 (3.1)
Number of quit attempts, median (IQR[a])	2 (1-3)
Maximum months of smoking abstinence, median (IQR)	2 (0.7-10.5)

[a]IQR: interquartile range.

Figure 1. Deepening the process of quitting smoking.



Emotions

Participants express positive and negative emotions related to smoking cessation process following Ekman classification [8]: fear, surprise, disgust, sadness, joy, and anger. At the beginning, they send messages of sadness and as the time passes, the messages are more positive. The emotion that appeared most frequently at the beginning was fear such as a fear of facing certain social events at which they usually would have smoked. Once these events have passed without falling back into the habit, they showed the joy they felt at remaining abstinent:

> *I have a dinner tonight and I'm really scared about it.* [ID 548, woman, 56 years]
>
> *Well, I passed an important test, a calçotada party with friends.* [ID 399, woman, 58 years]

The emotion of surprise also appeared when a participant realized he or she had not thought about tobacco for a long time:

> *...which means that I did not have time to think about smoking, and even I was surprised.* [ID 422, woman, 57 years]
>
> *I think I'm the newbie here, but it is truly surprising, the changes you notice, from smell to taste...and I'm only on my fifth day!* [ID 192, man, 41 years]

Some people resort to the emotion of disgust to avoid smoking again:

> *I pressed the panic button because I would like one cigarette. I'm not so well today.* [ID 452, woman, 54 years]

As the following dialog shows, in the early stages of the cessation process, sadness emerged as a powerful emotion and there was a sort of a duel about quitting smoking, a habit that had been with them for a long time. This mourning for what participants got out of smoking is also observed in these statements acknowledging that the pleasure of smoking was because of addiction:

> *I wonder if other people have also felt sad, thinking how happy they were to smoke and now not smoking...I enjoyed the 'misbehavior' of smoking...* [ID 647, woman, 41 years]
>
> *Yes...I didn't even want to get out of bed because I thought, What will I do if I don't smoke???* [ID 548, woman, 54 years]
>
> *And did you get over it?* [ID 647, woman, 41 years]
>
> *Oh, sure. But it will take you a few more days yet.* [ID 548, woman, 54 years]

Once the first few weeks have passed, participants reported an immense joy. Many were counting the days that they had gone without smoking and expressed pride in their achievement:

> *Happy Wednesday. Today makes 120 days I have not smoked. I am very happy and proud of what I'm doing. Don't give up, everybody!!!* [ID 548, woman, 54 years]

The Tobbstop app asks about emotions (anger, sadness, and bad or happy mood) every day when individuals start the app. Emotions are dynamic and responsive, and some emotions can be replaced by others throughout the smoking cessation process. The negative emotions appeared in the first days when participants started to quit smoking:

> *Every time I open the application and it asks me how am I...the first days I answered I am angry, sad, badmood...and now I have been saying for some time that I am happy.* [ID 543, woman, 54 years]

Coping Strategies

External and internal demands assessed as excessive or overwhelming might be confronted with different coping strategies. The most frequent coping strategy identified to decrease psychological stress, anxiety, and fear of relapse was physical activity. Other strategies used to decrease psychological distress and avoid thinking about tobacco consumption included listening to music, cleaning, reading, cooking, or playing:

> *Stationary bicycle...to not think about tobacco.* [ID 279, woman, 40 years]
>
> *All I do is clean and listen to music to not think about it.* [ID 429, woman, 40 years]
>
> *I fix supper and spend time on that.* [ID 406, man, 52 years]
>
> *When I get overwhelmed I look for a game and I get over it.* [ID 363, woman, 54 years]

Eating certain foods, drinking liquids, and eating candies as a distraction were also used as a way to reduce stress:

> *I'm sure that sunflower seed salesmen are happy about my not smoking!!!! I've already eaten 2 packages today!* [ID 192, man, 41 years]
>
> *I don't know how to get over that need, I'm trying right now to think about other things and I am at work but I would really like to smoke. I am going to drink water or a Coca-Cola I might find around the office.* [ID 259, woman, 42 years]

In order to decrease symptoms of nicotine abstinence, the participants used different types of treatments such as Varenicline, a nicotine substitute. The participants explained the difficulties they experienced with nicotine substitutes to calm the anxiety produced by not having nicotine, especially patches in the case of those receiving no treatment or a different treatment:

> *I take the pills but now the doctor called to give me the patches because I had a problem with Champix but the truth is that they work.* [ID 399, woman, 58 years]
>
> *For now, the patch gives me the nicotine I need...I only miss having a cigarette between my fingers...* [ID 266, woman, 51 years]

The following dialog shows that humor is another strategy that was frequently used to reduce stress and anxiety:

> *So, how's the car repair going?* [ID 647, woman, 41 years]
>
> *No defects. Hahaha.* [ID 485, man, 49 years]
>
> *And the O2 buffer?* [ID 647, woman, 41 years]
>
> *Bad joke, no? "carboximetry at 2."* [ID 485, man, 49 years]
>
> *Tell the nurse to send the carbon dioxide meter out for repairs. It must not be working right...* [ID 422, woman, 57 years]

At follow-up visits, participants tested their carboximetric level; seeing a score of 0 became an element of self-reinforcement. In addition, they thought that a low (or 0) score meant they had clear lungs:

> *Today at the exhalation test I almost jumped out of the chair, I was so happy. I never thought I'd react like that!* [ID 422, woman, 57 years]

> *I went to see the nurse and I got a "2" on lung toxicity.* [ID 491, woman, 44 years]

Beliefs About Weight, Physical Activity, and Medication

Beliefs are cognitive configurations individually created or culturally shared preexisting concepts of reality that act as a perceptual lens. Despite being counseled to follow a healthy diet and drink a lot of liquids, the chat participants showed a belief that it would be impossible to avoid gaining weight during the smoking cessation process. If the timing coincided with menopause and other aspects of aging, they believed there would be greater weight gain:

> *Did you gain weight too?* [ID 694, woman, 50 years]
>
> *Five kilos (10 pounds) in three months!! But it was worth it. Later on you can lose them, but slowly.* [ID 577, woman, 43 years]
>
> *Age is helping me [gain weight] too...* [ID 548, woman, 54 years]
>
> *That could be.* [ID 577, woman, 43 years]
>
> *Well, with menopause besides, I can't tell you what all else is going on with me, hahaha.* [ID 422, woman, 57 years]

Nonetheless, they believed it was worse to smoke than to gain weight and that they would not be able to lose the weight while trying to quit smoking. Instead, they proposed it as a challenge for the near future:

> *The weight doesn't worry me too much if we don't put on too much! Anyway, better 10 extra pounds than smoking again, no?* [ID 192, man, 41 years]
>
> *Yeah, we'll get rid of the kilos and we will also be rid of the addiction to tobacco.* [ID 399, woman, 58 years]

Participants believe that physical activity is positive and useful to decrease craving symptoms. Most participants explained that they were exercising regularly:

> *The truth is that doing sports helps a lot to overcome this vice.* [ID 192, man, 41 years]
>
> *Doing sports is the best !!* [ID 164, woman, 32 years]

Although many participants were quitting smoking with the help of medication, there was still a belief that it was best to do it without any pharmacological help:

> *Anyway, if you can leave without anything, that's the best. It's just a head issue. Be strong. And say no. In a week, the cold turkey effect just no longer exists.* [ID 164, woman, 32 years]
>
> *I don't know anyone who has used medication to quit smoking. My friends have quit with nothing.* [ID 429, woman, 40 years]

Participants believe that being in a tobacco-free environment would help them in their quit process and were concerned when they faced situations where they normally would have smoked and where they knew they would meet other smokers:

> *I'm lucky that my circle of friends does not smoke, almost nobody.* [ID 499, woman, 47 years]
>
> *My son and daughter-in-law also quit a few years ago. My daughter sees it as more difficult for her partner, who also smokes.* [ID 422, woman, 57 years]

Values

Values are expressions of what is important for that person. The most important value expressed by participants was *health*, which was stated as the main motivation to quit smoking:

> *I want to quit for my health and for my wife and my son. If I get sick they will have a very hard time.* [ID 406, man, 52 years]
>
> *I'm quitting because I do not want the doctor to tell me one day, either quit or you will die.* [ID 470, woman, 57 years]
>
> *Not even 24h yet but I am happy because I need to quit smoking. To health!!!* [ID 548, woman, 54 years]

They also highlighted the benefits of smoking cessation for their health. They described benefits they perceived in their body, how food smelled and tasted better, how they did not run out of breath while exercising, and how the various follow-up tests reinforced their decision to stop smoking:

> *but it really is surprising, the changes you notice, from the sense of smell to taste...and I'm only on my fifth day!* [ID 192, man, 41 years]
>
> *I can breathe better, I can smell better AND I smell better. Food tastes better, and I've saved 280 euros.* [ID 164, woman, 32 years]

Family was another important value and one of the main motivators to start the process of smoking cessation. Concerns included having a negative influence on their children, grandchildren, or other relatives, or impairing the health of family members, especially children, with second-hand tobacco smoke:

> *My 4-year-old granddaughter, whose parents do not smoke, saw me and her aunt smoking, and told my daughter-in-law that "when I grow up I'm going to smoke like yaya and auntie." You should have seen my face, and I told her that I would not smoke anymore because it makes 'owies'.* [ID 422, woman, 57 years]
>
> *I have 2 little children who have a lot of bronchitis and the smoke is really bad for them...and then for my health...I'm 33 years old and I have a lot of breathing problems.* [ID 843, man, 35 years]

In other cases, the family had asked them to stop smoking, but the participants were not always receptive to these messages when they were in the early precontemplative stage. They still did not see that smoking would cause any harm and it was not until the contemplative and preparatory stage when they became aware of all the messages they had gotten from relatives, health professionals, and friends:

> *My two children asked me to quit and I am doing it for them.* [ID 280, man, 55 years]
>
> *My husband is anti-smoking and has been telling me to quit for 9 years and I paid no attention; then one*

day, looking at Facebook, I saw [the TOBBSTOP study] and I decided. [ID 429, woman, 40 years]

Another important value was *freedom* when participants tired of being dependent on tobacco and became conscious of their nicotine addiction. Once they had started the cessation process and were aware they had regained their freedom, this awareness became an important motivator to succeed. They were also critical of people who continued to smoke:

First of all, I didn't smoke because I liked it but because I was addicted. Tobacco tastes bad. [ID 485, man, 49 years]

I'm tired of being dependent on tobacco. [ID 577, woman, 30 years]

I won't go back to it because even when I smoked the monkey stayed on my back. Many times, right after I smoked, I thought, "Really? Again? and I would have smoked another one..." [ID 647, woman, 41 years]

Look, yesterday I was inLes Gavarres and for the first time I was inside, not out on the terrace, and I looked at the smokers who were out there. The image was grotesque to look at. It was like they were being controlled. Think about that. [ID 337, man, 41 years]

Another value was to *help others to quit smoking*, especially friends and family:

I have convinced three people in my circle to quit...I'm on a crusade against tobaccoooo! [ID 192, man, 41 years]

Money was a great motivator to begin a cessation attempt although it was combined with other elements. Saving money was an important value in remaining abstinent. The participants talked about what they wanted to buy with the money saved:

I don't want to smoke -- for my health, the money, the smell on my clothing... [ID 541, woman, 56 years]

With what we spent on tobacco, my partner and I could take a cruise you won't believe... [ID 364, woman, 49 years]

Cognitive Evaluation

Cognitive evaluation is the process that determines the consequences of smoking cessation in the individual. When participants smoked a cigarette or just took a drag, they were less active in the chat group because they felt sad, guilty, and ashamed although the group encouraged them to continue trying:

I keep reading you...but I have not been able to quit. [ID 259, woman, 42 years]

I feel bad because I wear the patch and don't smoke cigarettes (it bothers me) but I use the electronic cigarette. With non-nicotine liquid and I don't inhale. I don't inhale the smoke but I use it and feel guilty about it. [ID 399, woman, 58 years]

Although they were aware that the first days are the most difficult, and are when it is easiest to have a relapse, in some cases, they minimized the risk they had overcome during the first weeks:

No prob, man. Once you get through the first week and say a few times, "No, I don't smoke anymore," that's it. [ID 192, man, 41 years]

So, yeah, it's true that it's hard at first but after that it's not. [ID 422, woman, 57 years]

Cutrona Model

According to the Cutrona model, online social support was classified into 5 subcategories:

Information Support

Chat was perceived by participants as a strength, as it provides cognitive support by sharing advice and practical information with others. Many of the messages offered suggestions about not gaining weight during the cessation attempt:

My advice is to be careful; enjoy the food, which will taste better than ever...but do not forget that you can go from gaining 4 kilos to 10 without even noticing. [ID 164, woman, 32 years]

Try with natural juices and sport. Cheer up!! These are the first few days. [ID 429, woman, 40 years]

In many messages, participants recommended physical exercise as a method to control anxiety and described the different activities they performed:

The trick is to make up your mind that you really want to stop and do some sport. Try a "fun run" event and I'm sure you will get hooked on it. [ID 192, man, 41 years]

Sport or physical activities works well !!! [ID 162, woman, 32 years]

Among the advice given to help overcome the withdrawal syndrome was natural remedies (eg, herbal teas and tryptophan):

It is important to stand firm and not smoke or take even one puff. Lime-blossom or valerian tea can help you. [ID 623, woman, 37 years]

Often the participants had made previous attempts to quit smoking, and they shared these experiences with the group, including the reasons that led them to relapse. They warned the others not to smoke even a single cigarette because that was what led them to fail:

From all this I learned that if you stop smoking you should never smoke even one. [ID 363, woman, 54 years]

You're right. I did not smoke with the pregnancies and then people offered me one and I went back to smoking. [ID 422, woman, 57 years]

Some messages referred to the opinions and advice of experts, sometimes with verbatim phrases of what a doctor had told them:

As my doctor says, quitting smoking is learning and there is no learning without relapses. [ID 320, woman, 50 years]

Self-Esteem Support

The group offered compliments to those who were achieving their goals and considered them to be role models:

> *Thank you, you have become champions!* [ID 422, woman, 57 years]
>
> *Congratulations!!! A good example to follow.* [ID 477, woman, 30 years]

Network Support

Recurrent messages were found to provide group support to overcome the worst moments, especially in the first few days. In the group, people found others who were in the same situation and understood what was happening to them at that time:

> *It's my second day. I'm having a nervous breakdown!!!!* [ID 354, woman, 43 years]
>
> *Hang in there! It's my FIRST day and I don't have to tell you anything you don't already know.* [ID 348, woman, 42 years]
>
> *Together we'll make it through!* [ID 363, woman, 54 years]

The chat also allowed those who were just beginning the program to ask questions to, instead of those who had been in the process longer:

> *[NAME], I have a question. After 103 days, do you still think about smoking?* [ID 355, man, 42 years]

Emotional Support

Participants sought support from the group when they felt a need to smoke, and received messages of support and motivation to help them get past the *craving* episodes [16]:

> *I need a cigarrooooo.* [ID 375, man, 36 years]
>
> *Don't smoke, you are stronger than that.* [ID 361, woman, 34 years]

They often needed to validate their emotions with the group, especially when they had not smoked for a number of days. It became important to count the days without smoking and to seek congratulations from the group; this positive feedback rewarded them emotionally:

> *83 days without smoke.* [ID 485, man, 49 years]
>
> *Congratulations!* [ID 548, woman, 54 years]

The group also offered encouragement when participants relapsed and had a cigarette:

> *Don't worry, try again.* [ID 361, woman, 43 years]
>
> *Don't believe that more than one hasn't had a fall and still do; they are not all so strong.* [ID 422, woman, 57 years]

Several participants mentioned eating snacks to quell anxiety. In these cases, the group downplayed the weight gain, considering smoking to be worse than gaining a few extra pounds:

> *Relax, the extra pounds go away but your lungs and your body in general will thank you...* [ID 270, man, 35 years]

Several messages show virtual affection:

> *You're welcome. When you get the urge to smoke, think "maybe later" and that's how you get past it. [Sending you] a kiss.* [ID 422, woman, 57 years]
>
> *I am so sorry...there are situations that require your energy...When you start again, you will achieve it, and will do better with experience! Asuper-hug!* [ID 477, woman, 30 years]
>
> *Thank you. Everybody in this group is super-cool!!!* [ID 548, woman, 54 years]

Tangible Assistance

First of all, the group decided to make closer contact and a *whatsapp* group was proposed:

> *We could do a whatsapp group.* [ID 299, woman, 49 years]
>
> *Yes, that would be cool.* [ID 192, man, 41 years]
>
> *Great, so who's going to do it?* [ID 299, woman, 49 years]
>
> *If you want, I'll set it up.* [ID 192, man, 41 years]

As new participants were being integrated into the chat, they were invited to join the group:

> *Some colleagues formed a whatsapp group a few days ago to help us more personally in case someone needs it. It is a complement to the [study] app. Anybody who wants to join will be welcome.* [ID 192, man, 41 years]

The connection between participants that was made in the group was so strong that the need arose to get to know each other outside of the study:

> *It would be good to meet someday, and not just those fromTarragona – everybody whowants to and can!* [ID 548, woman, 54 years]

In total, 10 people arranged a day to meet. As a separate whatsapp group was established, to which the research team did not have access, we do not know exactly how many people got together. We do know that it was satisfactory because they talked about organizing a second one for the people who could not attend:

> *A great get-together!!!! At the end of summer, another one, ehhhh?* [ID 647, woman, 41 years]

Discussion

Principal Findings

This study found that a chat integrated into a mobile app was a useful tool for offering social support and sharing emotions, information, or coping strategies to smokers in the process of quitting the habit. To our knowledge, this was the first qualitative descriptive analysis of a chat included in an app aimed at people in the *action* stage of change who were trying to quit smoking.

The analysis of the chat messages showed that it was an active forum used by participants to exchange information, concerns, and social support. Some of the emotions described by Eckman

appeared in the chat [8]. In the first phases, users show sadness and fear a relapse. These withdrawal symptoms peak within the first weeks and last for about 2 to 4 weeks [17]. As they progressed in the process, participants moved to more positive emotions such as joy and even euphoria. In addition, in the more advanced phases the participants minimized the risk of relapse as considered themselves to be past that phase.

Values and beliefs are essential to initiate the smoking cessation process as well as to maintain abstinence. As in previous studies [18-21], we found that health and family, including concerns about a family member's health or illness, or not willing to be a bad example for children, were primary reasons for quitting smoking. Moreover, the health benefits of smoking cessation are an important motivation to maintain abstinence as well as passing health checkups revisions, mainly to obtain a 0 in the co-oximetry.

Although cigarette taxes have shown dramatic increases in Spain, following the European Union legislation (Council Directive CD 2011/64/EU), participants infrequently reported money as a reason for quitting smoking. When they did report it, money appeared as a motivator in combination with other elements such as health. These results differ from a previous study in France in which money was the most reported reason for quitting smoking [20]; France suffered a dramatic increase of cigarette taxes between 2003 and 2004 [20]. The difference in effect between France and Spain in tax increases may be mediated through the height of the increase that in France was acute and large, not stepwise as Spain. However, our results concurred with a study performed in Spain in which money was not a main reason for quitting smoking [21]. Moreover, a 2006 study in Spain found that the introduction of a tax on manufactured cigarettes did not affect smoking prevalence in men and had a weak effect in women [22].

Most information provided in the chat was related to avoiding weight gain. Although weight varies greatly after quitting cigarettes, a published meta-analysis found that about 16% of quitters lost weight and 13% gained more than 10 kg [23]. The participants in our study believed it to be impossible to avoid weight gain. Smokers, and particularly women, have a high level of weight concerns that influence the likelihood of initiating a smoking cessation process [24]. However, the users of our chat believed that smoking is worse than increased weight.

Within the chat group, we observed that people who had a relapse were embarrassed and, although some sought the help of the group, some participants might not have asked for help because they felt guilty about deceiving themselves and above all for deceiving the group at the same time.

Within its social support, the group also offered emotional and informational support. We found similar results in other research, such as the study by Coulson et al, which indicates that group members offer informational and emotional support [25]; Ko et al [26] suggested that self-disclosure in blogs or Facebook is beneficial to users in obtaining social support and establishing or maintaining friendships [27]. However, in our study, various members of the group felt a need to meet each other and organized a time to get together. This could be because *the group* acquired such importance that its components wanted to connect in person.

Clinical Implications

The Tobbstop app was designed to accompany the process of quitting for the first 90 days, the most critical days for a possible relapse. Participants who succeeded in abstaining from smoking used the chat to help newcomers providing advice, information, and emotional support. However, previous studies found that more than half of the messages from the support group were posted during the first months of the smoking cessation process indicating that people require more support in the first steps of quitting [3].

We found an important online social support community that complemented the information and support provided in primary health care consultations and other resources (expert patient of tobacco cessation, group activities, and community activities) in the first phases of smoking cessation programs. Moreover, online support groups have the potential to provide a unique opportunity for health professionals to learn about the experiences and views of individuals.

Online social support from an established group during the change process has several benefits. Participants are not restricted by the temporal, geographical, and spatial limitations typically associated with face-to-face groups; individuals can send and receive messages at any time of the day or night. In addition, online support groups may bring together a more varied range of individuals to offer diverse perspectives, experiences, opinions, and sources of information.

The emotional support obtained from the app may help some people deal with relapses. Little is known about how online discussions transform into real-life behavioral changes [28]. Efficacy is a concern because a recent review concluded that no robust evidence exists of the effectiveness of online peer-to-peer support groups [28]. An important next step is to assess the efficacy of online app forums by conducting randomized controlled trials.

New technologies and, more specifically, chat as a channel of communication may be able to help us to create groups of people who are engaged in the same process such as smoking cessation. The chat group can provide support and help 24 hours a day.

Limitations and Strengths

We neither know the reasons why some participants did not use the chat nor what their comments might have been; it is possible that some users only read posts and did not contribute to them. For those users, it would be useful to determine which channel of communications would work best. A descriptive analysis has been done. It would be interesting to conduct a more in-depth and interpretative analysis according to sex, age group, studies, and other characteristics considered. According to the study protocol, participants who relapsed to tobacco consumption were removed from the Tobbstop app [12]. Those who relapsed were dismissed from the study and could not use the app, so we lack information to determine their emotions and feelings before the relapse, a process contemplated within Prochaska and DiClemente stages of change [14].

Among the strengths of the study was the interaction between participants who were in different phases of the process. Some people were just starting and others had already gone 180 days without smoking. A person who has already passed through a given stage will show empathy, respect, and confidence in others' abilities and reinforce the social support. In addition, the chat showed a diverse and pluralistic discourse.

Conclusions

The results of this study suggest that a chat integrated into a mobile app can be a useful tool for smokers who are in the process of quitting. In our study, the app offered social support and a space where participants shared concerns, information, and strategies This type of online social support could complement the information and support provided in primary health care consultations and other resources in smoking cessation programs.

Acknowledgments

The authors would like to thank all who participated in the Tobbstop study, including all the health professionals, and the Catalan Institute of Health for publicizing our study and the Instituto de Investigación en Atención Primaria Jordi Gol for all of the support we received. The project received a research grant from the Carlos III Institute of Health, Ministry of Economy and Competitiveness (Spain), awarded by the 2012 call under Health Strategy Action 2013-2016 within the National Research Program oriented to Societal Challenges of the Technical, Scientific, and Innovation Research National Plan 2013-2016, with reference PI12/01977, cofunded by the European Regional Development Fund (European Union).

JDG has been funded by the MINECO project FIS2016-78904-C3-1-P.

Authors' Contributions

All authors participated in developing the study design, in data interpretation, or in writing and revision of the paper. All authors approved the final version of the paper for publication.

The Tobbstop Study Group is comprised of: Susana Abajo Llama, Manoli Aliseda Tienza, Elena Alvarez Soler, Mª Teresa Antich Salvado, Vicens Ramon Arnau Adan, Teresa Aviño Llopis, Iolanda Barba Pérez, Isabel Barco Carrillo, Anna Bibiloni Sola, Josep Blanch Aubia, Joan Boj Casajuana, Marinela Boldu Ortega, Dolors Borras Vicente, Olga Briones Carcedo, Enric Buera Sales, Olivia Bujaldon Castillo, Paloma Camos Guijosa, Ruth Caro Garduaño, Sonia Cortiles Bosch, Mª Teresa Del Amo Martinez, Eva Mª Dominguez Ibañez, Roxana Elena Catalin, Guillermo Elizalde Del Rio, Pilar Escoda Fuentes, Miquel Raiget, Iolanda Farre Torra, Nuria Ferrando Ruana, Carme Fonollosa Queralt, Cruzma Fuentes Bellido, Carme Giner Aguiló, Mª Purificación Gomez Martinez, Javier Goni Garin, Maria Gorgues Espasa, Montserrat Grivé Isern, Carmen Hernandez, Carolina Inza Henry, Aratxa Jara Ruiz de Gaona, Carmen Jimenez Civit, Montserrat Juncosa Cabre, Julia Ledo Garcia, Miriam Leiva Caballero, Montserrat Llauradó Vernet, Daniela Lleixa Lleixa, Antonio Lopez Barea, Carol Lopez Guerrero, Sonia Loran Valcarcel, Ana Maria Lorente Zozaya, Maria Jesus Lorenzo Posada, Emilia Martí Suau, Immaculada Martí Sendra, Paco Martin Lujan, Trinidad Martinez Perez, Dolors Matamoros Belsue, Oscar Molina Pla, Francisca Monterroso Segura, Mª Dolores Moran Martinez, Maria Moreno, Gemma Muñoz Alvarez, Manuel Murillo Molina, Mª Angels Naranjo, Elia Nevot Bueno, Angels Odena Estrada, Immaculada Palou Val, Joan Pardo Andujar, Roser Pedret Llaveria, Pilar Perez Acin, Armand Pi Coll, Concepción Rambla Vidal, Anna Riart Miquel, Marga Robert Vilà, Ramon Rodriguez Perez, Margarida Royo Cavalle, Jose Miguel Royo, Laura Ruiz de Porras Rosselló, Ramon Sagarra Álamo, Luis Juan San José Mallen, Maria Jesus Sancho Vilellas, Nuria Sarrà Manetes, Eulalia Sigró Civit, Rocío Solis Narvaez, Alberto Suelves Ascaso, Mercè Timon Torres, Montserrat Tomàs Martí, Eva Torres Maillo, Francisca Vallespí Borrull and Emma Vidal Rojas.

Conflicts of Interest

None declared.

References

1. Organización Mundial de la Salud. Hacer cumplir las prohibiciones sobre publicidad, promoción y patrocinio del tabaco. Informe OMS sobre la epidemia mundial de tabaquismo URL: http://www.who.int/tobacco/global_report/2013/es_summary.pdf [accessed 2018-11-08] [WebCite Cache ID 73m2ZgYq2]
2. Ministerio de Sanidad, Servicios Sociales e Igualdad. Evolución de los indicadores del estado de salud en España y su magnitud en el contexto de la Unión Europea. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad; 2017. Indicadores de salud 2017 URL: https://www.mscbs.gob.es/estadEstudios/estadisticas/inforRecopilaciones/docs/Indicadores2017.pdf [accessed 2018-11-08] [WebCite Cache ID 73m5UtEvo]
3. Selby P, van Mierlo T, Voci SC, Parent D, Cunningham JA. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. J Med Internet Res 2010;12(3):e34 [FREE Full text] [doi: 10.2196/jmir.1340] [Medline: 20719739]

4. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. Am J Prev Med 2008 Aug;35(2):158-176 [FREE Full text] [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
5. Ramo DE, Liu H, Prochaska JJ. A mixed-methods study of young adults' receptivity to using Facebook for smoking cessation: if you build it, will they come? Am J Health Promot 2015 Mar;29(4):e126-e135. [doi: 10.4278/ajhp.130326-QUAL-128] [Medline: 24575728]
6. Coulson NS, Knibb RC. Coping with food allergy: exploring the role of the online support group. Cyberpsychol Behav 2007 Feb;10(1):145-148 [FREE Full text] [doi: 10.1089/cpb.2006.9978] [Medline: 17305463]
7. Burri M, Baujard V, Etter J. A qualitative analysis of an internet discussion forum for recent ex-smokers. Nicotine Tob Res 2006 Dec;8(Suppl 1):S13-S19. [Medline: 17491166]
8. Ekman P. An argument for basic emotions. Cogn Emot 1992;6:169-200. [doi: 10.1080/02699939208411068]
9. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Commun Res 1992 Apr;19(2):154-174. [doi: 10.1177/009365092019002002]
10. Lazarus R, Folkman S. Estrés y procesos cognitivos. Barcelona: Ediciones Martínez Roca; 1986:1-468.
11. Ballvé MJ, Bladé CJ, Borràs IM, Granollers MS, Morera JC, Serra AA, et al. Guies de pràctica clínica. Barcelona: Institut Català de la Salut; 2009. Detecció i tractament del consum de tabac URL: http://ics.gencat.cat/web/.content/documents/assistencia/gpc/guia_tabac.pdf [accessed 2018-11-08] [WebCite Cache ID 73m4CDPOJ]
12. Valdivieso-López E, Flores-Mateo G, Molina-Gómez J, Rey-Reñones C, Barrera UM, Duch J, et al. Efficacy of a mobile application for smoking cessation in young people: study protocol for a clustered, randomized trial. BMC Public Health 2013;13:704 [FREE Full text] [doi: 10.1186/1471-2458-13-704] [Medline: 23915067]
13. Berenguera A, Fernández DS, Pons M, Pujol ER, Saura S. Saludcomunitaria. 2014. Escuchar, observar y comprender URL: https://saludcomunitaria.files.wordpress.com/2014/12/escucharobservarcomprender.pdf [accessed 2018-11-08] [WebCite Cache ID 73m4Ozyy3]
14. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 6863699]
15. Calderon C. Criterios de calidad en la investigación cualitativa en salud (ICS): apuntes para un debate necesario [in Spanish]. Revista Española de Salud Pública 2002;76(5):473-482 [FREE Full text]
16. Muñoz M, Sanjuan R, Fernández-Santaella MC, Vila JM. Aspectos neuropsicológicos del craving por la nicotina. Adicciones 2011;23(2):123 [FREE Full text]
17. Hughes JR. Effects of abstinence from tobacco: valid symptoms and time course. Nicotine Tob Res 2007 Mar;9(3):315-327. [doi: 10.1080/14622200701188919] [Medline: 17365764]
18. McCaul KD, Hockemeyer JR, Johnson RJ, Zetocha K, Quinlan K, Glasgow RE. Motivation to quit using cigarettes: a review. Addict Behav 2006 Jan;31(1):42-56. [doi: 10.1016/j.addbeh.2005.04.004] [Medline: 15916861]
19. Sieminska A, Buczkowski K, Jassem E, Lewandowska K, Ucinska R, Chelminska M. Patterns of motivations and ways of quitting smoking among Polish smokers: a questionnaire study. BMC Public Health 2008 Aug 04;8:274 [FREE Full text] [doi: 10.1186/1471-2458-8-274] [Medline: 18680600]
20. Aubin H, Peiffer G, Stoebner-Delbarre A, Vicaut E, Jeanpetit Y, Solesse A, et al. The French Observational Cohort of Usual Smokers (FOCUS) cohort: French smokers perceptions and attitudes towards smoking cessation. BMC Public Health 2010 Feb 26;10:100 [FREE Full text] [doi: 10.1186/1471-2458-10-100] [Medline: 20184784]
21. Ramon TJ, Bruguera CE, Fernández PC, Sanz DB, Ramírez VE. [Reasons for smoking cessation in Spain by gender and age]. Gac Sanit 2009;23(6):539.e1-539.e6 [FREE Full text] [doi: 10.1016/j.gaceta.2009.07.004] [Medline: 19880217]
22. López-Nicolás A, Badillo-Amador L, Cobacho-Tornel MB. Will the European Union's new tobacco tax legislation lead to reductions in smoking prevalence? Evidence from a quasi-experiment in Spain. Nicotine Tob Res 2013 Dec;15(12):1963-1970. [doi: 10.1093/ntr/ntt038] [Medline: 23558971]
23. Aubin H, Farley A, Lycett D, Lahmek P, Aveyard P. Weight gain in smokers after quitting cigarettes: meta-analysis. Br Med J 2012 Jul 10;345:e4439 [FREE Full text] [Medline: 22782848]
24. Aubin H, Berlin I, Smadja E, West R. Factors associated with higher body mass index, weight concern, and weight gain in a multinational cohort study of smokers intending to quit. Int J Environ Res Public Health 2009 Dec;6(3):943-957 [FREE Full text] [doi: 10.3390/ijerph6030943] [Medline: 19440425]
25. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007 Oct;68(2):173-178. [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]
26. Ko H, Wang L, Xu Y. What kinds of social support does the audience offer to a-list journal-type bloggers? In: PACIS 2011 Proceedings. 2011 Jul 09 Presented at: Pacific Asia Conference on Information Systems (PACIS); 7 July, 2011; Brisbane URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.232.1780&rep=rep1&type=pdf
27. Kummervold PE, Gammon D, Bergvik S, Johnsen JK, Hasvold T, Rosenvinge JH. Social support in a wired world: use of online mental health forums in Norway. Nord J Psychiatry 2002;56(1):59-65. [doi: 10.1080/08039480252803945] [Medline: 11869468]

28. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. Br Med J 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]

Abbreviations

IDIAP: Investigación en Atención Primaria
IQR: interquartile range

Edited by G Eysenbach; submitted 24.05.18; peer-reviewed by C Tome-Pires, MA Styn, N Chavannes; comments to author 26.07.18; revised version received 20.09.18; accepted 23.09.18; published 20.12.18.

Please cite as:
Granado-Font E, Ferré-Grau C, Rey-Reñones C, Pons-Vigués M, Pujol Ribera E, Berenguera A, Barrera-Uriarte ML, Basora J, Valverde-Trillo A, Duch J, Flores-Mateo G
Coping Strategies and Social Support in a Mobile Phone Chat App Designed to Support Smoking Cessation: Qualitative Analysis
JMIR Mhealth Uhealth 2018;6(12):e11071
URL: http://mhealth.jmir.org/2018/12/e11071/
doi:10.2196/11071
PMID:30573445

Original Paper

Parents' Perspectives on the Theoretical Domains Framework Elements Needed in a Pediatric Health Behavior App: A Crowdsourced Social Validity Study

Christopher C Cushing[1,2], PhD; David A Fedele[3], PhD; Erin E Brannon[4], PhD; Tiffany Kichline[1], MS

[1]Clinical Child Psychology Program, University of Kansas, Lawrence, KS, United States
[2]Schiefelbusch Institute for Life Span Studies, University of Kansas, Lawrence, KS, United States
[3]Department of Clinical & Health Psychology, University of Florida, Gainesville, FL, United States
[4]Dayton Children's Hospital, Dayton, OH, United States

Corresponding Author:
Christopher C Cushing, PhD
Clinical Child Psychology Program
University of Kansas
Dole Human Development Center
1000 Sunnyside Avenue, Room 2011
Lawrence, KS, 66049
United States
Phone: 1 7858640713
Email: christopher.cushing@ku.edu

Abstract

Background: Most pediatric studies do not include parent stakeholders in the design of the intervention itself and many pediatric mobile health (mHealth) interventions are not meaningfully disseminated after the trial period ends. Consequently, the consumer desire for mobile apps targeting pediatric health behavior is likely to be met by commercial products that are not based in theory or evidence and may not take stakeholder preferences into account.

Objective: The aim was to assess parent preference for mobile app features that map onto specific Theoretical Domains Framework (TDF) elements.

Methods: This study was a crowdsourced social validity study of 183 parents who were asked to rate their preferences for mobile app features that correspond to elements of the TDF. The TDF organizes a large number of theoretical models and constructs into three components: (1) capability, (2) motivation, and (3) opportunity. Parents of children were recruited through Amazon Mechanical Turk.

Results: The majority of participants were Caucasian and mean age was 36.9 (SD 8.0) years. Results revealed broad acceptability of communication, motivation, and opportunity domains. However, the degree to which each domain was valued varied within behavioral category. Parents demonstrated a preference for increasing procedural knowledge for physical activity and diet behaviors over sleep ($F_{2,545}$=5.18, P=.006). Similarly, parents valued self-monitoring as more important for physical activity than sleep ($F_{2,546}$=4.04, P=.02). When asked about the value of features to help children develop skills, parents preferred those features for dietary behavior over sleep ($F_{2,546}$=3.57, P=.03). Parents perceived that goal-setting features would be most useful for physical activity over sleep and diet ($F_{2,545}$=5.30, P=.005). Incentive features within the app were seen as most useful for physical activity over sleep ($F_{2,546}$=4.34, P=.01).

Conclusions: This study presents a low-cost strategy for involving a large number of stakeholders in the discussion of how health behavior theory should be applied in a mHealth intervention. Our approach is innovative in that it took a scientific framework (ie, TDF) and made it digestible to parents so that they could then provide their opinions about features that might appear in a future app. Our survey items discriminated between various health behaviors allowing stakeholders to communicate the different health behaviors that they would like a TDF feature to change. Moreover, we were able to develop a set of consumer opinions about features that were directly linked to elements of the TDF.

(JMIR Mhealth Uhealth 2018;6(12):e192) doi:10.2196/mhealth.9808

KEYWORDS

mHealth; adolescent; children; parent; stakeholder; consumer preference

Introduction

Mobile health (mHealth) has the potential to extend the reach and dose of clinical care and expand the kinds of intervention that can be provided to youth (ie, those younger than 18 years) at the point of care [1]. mHealth interventions can be effective at improving youth health outcomes, especially when a parent is an active recipient of the intervention [2]. This finding is consistent with evidence from face-to-face health promotion and prevention interventions that indicate that multisystemic interventions are more effective at improving youth health than those that solely target the individual [3].

There is conceptual agreement that the first phase of digital intervention development should be both grounded in theory and include stakeholders in the process [4]. However, parent preferences and opinions about intervention components are commonly missing from mHealth interventions because stakeholders are not typically included in the development phase of the applications [2]. Moreover, widely available digital health products are not meaningfully based on theory or informed by published evidence [5-7]. The fact that there is a large and vibrant marketplace of digital health tools, even if not evidence-based, suggests that currently available digital health solutions do tap into at least some perceived stakeholder need. If true, this conclusion suggests that research teams should look early and frequently for stakeholder opinions about the design of digital health interventions. One way to ensure that interventions are both based in theory and involve stakeholders is to use a theoretical framework to inform feedback from the stakeholders themselves.

Stakeholder engagement in intervention design can be conceptualized as part of a larger social validity framework [8]. Social validity refers to the degree to which a given intervention is (1) important to society; (2) involves treatment that is acceptable to consumers in terms of cost, ethical considerations, and practicality; and (3) results in treatment outcomes that are acceptable to consumers [9]. Crowdsourcing platforms and the advancement of population screening approaches present an externally valid opportunity to interact with stakeholders [10]. This study used social validity methodology to engage parent stakeholders in a crowdsourced study of stakeholder preferences for mobile app features that correspond to evidence-based behavior change strategies.

To empirically evaluate stakeholder perceptions of intervention elements they might encounter in mobile apps, it is imperative that there is a shared operational definition of the behavioral strategies that investigators would use to change health behavior [11]. Moreover, there is relatively good consensus that interventions should be based on a guiding theory rather than disconnected techniques [1,12]. Incorporating a full-featured theoretical framework, such as the Theoretical Domains Framework (TDF; see Cane et al [13] for review), allows interventionists to make falsifiable predictions about the performance of an intervention. The TDF has been proposed as a system for defining active theoretical elements within larger studies [12,13]. It organizes a large number of theoretical models and constructs into 12 independent domains which can be further aggregated into three components from the behavior change wheel posited by Michie et al [14]: (1) capability, (2) motivation, and (3) opportunity. Capability refers to both one's physical ability to perform a behavior and also to psychological factors such as knowledge, cognitive capacity, self-regulation, and mental skill. Motivation is subdivided into two elements: automatic and active. The automatic element of motivation involves noncognitive influencers of behavior such as reinforcement, emotion, and habit. The active element of motivation involves more cognitive decision-making processes. Opportunity represents external influences on behavior that facilitate or prompt a behavior via either environmental or social means.

Although theoretical frameworks are useful for researchers to design interventions, they can be particularly challenging to communicate to stakeholders due in part to the discipline-specific knowledge required to understand the definitions and predictions of these frameworks. Within pediatric mHealth interventions, there is a potential that developing a shared language could help parent stakeholders involved in intervention to fully communicate their preference for which behavior change techniques should be included in the final intervention.

Therefore, this study aimed to:

- Distill elements of the TDF into features that might be included in a mobile app (make translatable for parents);
- Quantify parent stakeholder preferences for elements of the TDF in mHealth interventions by asking them to consider specific features that would be consistent with each TDF element; and
- Quantify how parents are currently using mHealth to manage their children's health.

Methods

Recruitment

Participants were parents of children recruited via the Amazon MTurk platform between July and September 2016. We posted an advertisement using a requester account on MTurk and advertised our survey as a project soliciting "Opinions about smartphone apps." The advertisement noted that we estimated the Human Intelligence Task (HIT) to take 1 hour to complete and that valid responses would receive US $1.66 in compensation. All recruitment procedures and materials were approved by the Institutional Review Board at the University of Kansas (STUDY00003083).

Study Eligibility

Mechanical Turk (MTurk) is an online marketplace where workers can receive compensation for completing a range of HITs such as completing questionnaires. We required that

participants have a Master certification on the MTurk platform (ie, complete at least 1000 tasks and maintain at least 99% approval). This was done to ensure a participant pool who was likely to provide high-quality data and reduce data screening burden. Participants were also required to be the parent of at least one child younger than 18 years. To assess for this eligibility criterion, we followed procedures similar to Schleider and Weisz [10]. Specifically, we included a six-question screener that asked interested participants to self-report their sleep, physical activity, and dietary behaviors, as well as whether they had a child younger than the age of 18. All items other than child age were distraction questions and were not used for eligibility. If the answer to the question about parental status was affirmative, participants were allowed to complete the rest of the survey.

Participants

There were 572 participants who attempted to take the survey. Participants were screened out of the study if they were not parents or did not answer all the nine attention items correctly. Question examples included "Who was the first president of the United States?," "Who invented the lightbulb?," and "What is the current year?" We analyzed a final dataset consisting of 183 valid responses (age: mean 36.9, SD 8.0 years; see Table 1 for more demographic information).

Study Survey

A common approach to gathering social validity information from stakeholders is to develop short prompts and a questionnaire that can be administered to large groups relatively quickly [15,16]. Participants were directed from their participant recruitment portals to a Qualtrics survey. The full survey is available from the corresponding author.

Cell Phone Usage

Fifteen items assessed the participants' usage of mHealth apps including if they have downloaded apps, how often they use the app, and the platform on which they have downloaded apps. Participants also indicated their willingness to purchase apps and were asked to estimate the amount of money they would spend on a mHealth app. Questions also assessed if participants' have downloaded an app for their child or adolescent and the focus of the app.

Domains of the Theoretical Domains Framework

For the purposes of this study we chose to focus on health behaviors of diet, physical activity, and sleep as these were of primary intellectual interest to us. We selected elements of the TDF that were most directly applicable to mobile phone apps in order to characterize parents' perception of importance for inclusion in mHealth apps to target diet, physical activity, and sleep. The TDF domain questions were developed in an iterative process by three of the study authors (CC, DF, EB). Each author is a doctoral-level pediatric psychologist with expertise in behavior change. The TDF domain definitions were ranked by each author to evaluate their applicability to mHealth. After the group reached consensus, EB developed examples of mHealth features that would map on to each TDF domain. Examples were refined until consensus agreement was reached among CC, DF, and EB.

Definitions of the TDF domains were provided in the question stem along with examples for how the TDF element would be incorporated into a hypothetical app. Parents rated the helpfulness of that domain for inclusion in an app on a 5-point scale ranging from 1=not helpful to 5=really helpful. Example items included "Apps could also provide a way to assist children and adolescents in scheduling time to practice the behavior. The app could send an alarm for when the child needs to engage in physical activity or send a reminder to help mom and dad cook a nutritious meal, or lastly to lock the phone to encourage not using electronics at night" (example of practice) and "Apps may also be useful in giving the positive benefits of changing these healthy behaviors. For instance, sleep is related to better concentration, diet is related to more energy, and physical activity could lead to weight loss and better sleep" (example of optimism; see Multimedia Appendix 1 for a list of sample survey questions).

All parents were prompted to answer questions on the TDF domains regarding diet, physical activity, and sleep. Attention questions were also interspersed within the study questions as a manipulation check. Question examples include "Who was the first president of the United States?," "Who invented the lightbulb?," and "What is the current year?" Incorrect answers to these questions resulted in the participant's data being excluded from the study.

Data Analysis

All analyses were conducted using SPSS statistical software, version 23, for frequencies and descriptive statistics. To understand parents' views on the helpfulness of mHealth apps and the TDF in relation to different health behaviors, one-way between-subjects analyses of variances (ANOVA) were conducted with the construct as the independent variable and type of health behavior (ie, physical activity, diet, and sleep) as the dependent variable. Significant differences between health behaviors were followed up with post hoc comparisons using the Tukey honestly significant difference test to better understand the parents' ratings of the helpfulness of mHealth apps. Degrees of freedom were larger than the number of participants in the sample because three different ratings (ie, physical activity, diet, and sleep) were provided for each prompt by a single participant. Statistical tests were considered significant if type I error rates were less than 5%.

Table 1. Demographics of participants (N=183).

Variable	Participants
Age (years), mean (SD); range	36.9 (8.0); 23-68
Race, n (%)	
Caucasian/White	124 (67.8)
African American	16 (8.7)
Asian	28 (15.3)
Native American	4 (2.2)
Biracial	6 (3.3)
Other	5 (2.7)
Ethnicity, n (%)	
Hispanic or Latino	17 (9.3)
Not Hispanic or Latino	165 (90.2)
No response	1 (0.0)
Relationship status, n (%)	
Single	31 (16.9)
Married	126 (68.9)
Separated	4 (2.2)
Divorced	22 (12.0)
Annual family income (US$), n (%)	
0-19,999	19 (10.4)
20,000-39,999	44 (24.0)
40,000-59,000	38 (20.8)
60,000-79,999	38 (20.8)
80,000-99,999	25 (13.7)
≥100,000	19 (10.4)
Number of children, n (%)	
1	77 (42.1)
2	69 (37.7)
3	23 (12.6)
4-7	14 (7.7)
Current cell phone platform, n (%)	
Android	104 (56.8)
iPhone	75 (41.0)
Windows	3 (1.6)
Blackberry	1 (0.5)

Results

Parent mHealth App Use

A total of 105 of 183 parents (57.4%) reported that they had some type of mHealth app on their phone (this could be an app that was preloaded on the phone or downloaded intentionally). The majority of these parents (85/105, 81.0%) reported that they used their mHealth app every day or multiple times a week. However, when asked if they had downloaded a mHealth app, 156 parents (85.2%) reported they had not; additionally, only 29.6% (8/27) of participants who had downloaded a mHealth app reported that they downloaded the app on their child's cell phone.

Likelihood of Using mHealth Apps to Target Child Health Behaviors

When asked to rate how much of a problem physical activity was for their child, 70 of 183 (38.3%) parents reported it was little to a lot of a problem; 90 of 183 (49.2%) parents reported

diet was a little to a lot of a problem for their child. When asked to rate how much of a problem sleep was for their child, 83 of 183 (45.4%) parents rated sleep as little to a lot of a problem. When asked how likely they would be to use a mHealth app to target their child's physical activity, most parents (66.7%, 122/183) reported somewhat likely to very likely. A majority of parents (60.6%, 111/183) reported they were somewhat likely to very likely to use a mHealth app to target their child's diet. Parents also reported that they would be likely to use a mHealth app to target their child's sleep (51.9%, 95/183 of parents reported somewhat likely to very likely).

Theoretical Domains Framework and mHealth Behavior Change Techniques

Capability

The majority of parents rated the capability component as helpful for their child's physical activity (73.1%, 132/183 reported somewhat helpful to really helpful), diet (73.8%, 135/183 reported somewhat helpful to really helpful), and sleep (66.7%,122/183 reported somewhat helpful to really helpful). These results suggest that, overall, parents found the mHealth TDF domains related to capability to be helpful strategies for their child's health behaviors.

Within the knowledge domain, there were no significant differences in parents' ratings of the helpfulness of the knowledge and knowledge of task environment constructs. This suggests parents may view mHealth apps that provide knowledge and knowledge of task environment to be equally helpful for their child's physical activity, diet, and sleep behaviors. However, parents rated the helpfulness of procedural knowledge differently depending on the health behavior ($F_{2,545}$=5.18, P=.006). Post hoc comparisons revealed that parents rated features to enhance physical activity (mean 3.96, SD 1.08, P=.008; Cohen d=0.31) and diet knowledge (mean 3.91, SD 1.17, P=.03; Cohen d=0.26) as significantly more helpful than sleep features (mean 3.60, SD 1.21). There was no significant difference between diet and physical activity. Therefore, parents may view mHealth apps that provide training in how to perform the health behavior to be more helpful for their child's physical activity and diet than sleep.

In the skill domain, there was no significant difference in parents' ratings of helpfulness of the practice construct or skills construct. Parents may be more likely to view mHealth apps that promote repeated practice for the health behavior or provide education and training on a skill to be helpful for their child's overall health behaviors. There was a significant difference depending on health behavior when parents rated the helpfulness of skill development ($F_{2,546}$=3.57, P=.03). Post hoc analyses revealed diet (mean 4.03, SD 1.07, P=.02; Cohen d=0.26) significantly differed from sleep (mean 3.73, SD 1.21), whereby parents rated features designed to enhance dietary skill as more helpful than sleep skill features. There was no significant difference between physical activity and diet or physical activity and sleep. Therefore, parents may view mHealth apps aimed at skill development to be more helpful for their child's diet than sleep.

Parents' ratings of the helpfulness of the decision processes domain did not significantly differ suggesting parents may view mHealth apps that assist with making decisions related to health behaviors to be equally helpful for physical activity, diet, and sleep.

Within the behavioral regulation domain, parents' ratings did not significantly differ in the action planning and breaking habit constructs. These results suggest parents view mHealth apps that assist with creating a specific plan for healthy behaviors or provide information on how to break bad habits and form healthy habits to be useful for their child's overall health behaviors. However, there was a significant difference in parents' ratings of the self-monitoring construct ($F_{2,546}$=4.04, P=.02). Post hoc analyses revealed parents rated features designed to enhance physical activity self-monitoring (mean 4.11, SD 1.14, P=.01; Cohen d=0.29) as significantly more helpful than features for sleep self-monitoring (mean 3.78, SD 1.14). There was no significant difference between diet and physical activity or sleep. Therefore, parents may view mobile apps that promote self-monitoring as more helpful for their child's physical activity than sleep.

Motivation

Overall, parents found the use of mHealth apps within the motivation component to be helpful for physical activity (73.2%, 134/183 reported somewhat helpful to really helpful), diet (68.9%, 126/183 reported somewhat helpful to really helpful), and sleep (66.8%, 122/183 reported somewhat helpful to really helpful). Within the motivation component, there was no significant difference in parents' ratings of the helpfulness in the beliefs about consequences domain, intentions domain, and the optimism domain. These results suggest parents are likely to view mobile apps that remind their child of outcome expectancies and consequences related to health behaviors, understand if their child is ready to make a change, teach their child to think more positively about being healthy, and incorporate their child's identity in relation to health behaviors as equally helpful toward improving their child's physical activity, diet, and sleep behaviors.

In the goal domain, there was no significant difference in parents' ratings of goal setting; however, there was a significant difference in parents' ratings of distal and proximal goals ($F_{2,545}$=5.30, P=.005). Post hoc analyses revealed physical activity (mean 4.01, SD 0.96) was significantly different from diet (mean 3.68, SD 1.26, P=.02; Cohen d=0.29) and sleep (mean 3.67, SD 1.15, P=.01; Cohen d=0.32). Therefore, parents are likely to view mHealth apps focused on goal setting as equally helpful for their child's physical activity, diet, and sleep behaviors. However, parents may view mHealth apps that assist with distal and proximal goals to be most helpful for their child's physical activity.

Table 2. Descriptive mean scores[a] of helpfulness of the mHealth TDF domains.

Domain	Total (N=183), mean (SD)	Child (N=8), mean (SD)	Self (N=19), mean (SD)	None (N=156), mean (SD)
Capability				
Knowledge				
Knowledge	3.55 (1.22)	3.96 (0.96)	4.33 (0.93)	3.43 (1.22)
Procedural knowledge	3.82 (1.16)	3.92 (1.02)	4.39 (1.02)	3.75 (1.17)
Knowledge of task environment	3.65 (1.19)	4.17 (0.87)	4.26 (1.03)	3.54 (1.19)
Skill				
Skill development	3.90 (1.12)	4.21 (0.88)	4.47 (0.78)	3.81 (1.14)
Practice	3.91 (1.10)	4.17 (0.96)	4.33 (0.85)	3.84 (1.12)
Skills	3.85 (1.10)	4.25 (0.90)	4.37 (0.84)	3.76 (1.11)
Memory, attention, and decision processes				
Decision making	3.73 (1.17)	3.96 (1.30)	4.21 (1.10)	3.66 (1.16)
Behavioral regulation				
Self-monitor	3.94 (1.15)	3.96 (0.86)	4.25 (0.99)	3.90 (1.18)
Action planning	3.83 (1.16)	4.04 (0.91)	4.25 (0.85)	3.77 (1.19)
Break habit	3.61 (1.25)	4.16 (0.69)	4.04 (1.04)	3.55 (1.26)
Motivation				
Beliefs about consequences				
Outcome expectancies	3.72 (1.15)	4.04 (1.00)	4.28 (0.98)	3.63 (1.16)
Consequences	3.66 (1.21)	4.00 (1.14)	4.09 (1.11)	3.59 (1.22)
Goals				
Goal setting	3.87 (1.05)	3.92 (0.88)	4.39 (.077)	3.81 (1.07)
Distal and proximal goals	3.79 (1.14)	4.13 (0.80)	3.98 (1.10)	3.75 (1.15)
Intentions				
Transtheoretical	3.73 (1.17)	4.21 (0.83)	4.21 (1.10)	3.65 (1.18)
Optimism				
Optimism	3.88 (1.13)	4.21 (1.06)	4.33 (0.89)	3.81 (1.15)
Identity	3.19 (1.28)	3.33 (1.47)	3.84 (1.10)	3.11 (1.28)
Reinforcement				
Rewards	3.98 (1.14)	3.79 (1.14)	4.09 (1.11)	3.98 (1.15)
Reinforcement	3.89 (1.13)	4.17 (0.76)	4.23 (0.98)	3.84 (1.16)
Incentives	3.90 (1.17)	4.04 (1.12)	3.93 (1.21)	3.89 (1.16)
Opportunity				
Environmental context				
Resources	3.60 (1.23)	3.83 (1.17)	4.12 (1.12)	3.52 (1.23)
Barriers and facilitators	3.38 (1.24)	4.00 (0.66)	3.93 (1.12)	3.28 (1.25)
Social influences				
Social comparisons	3.47 (1.32)	3.71 (1.16)	3.84 (1.36)	3.41 (1.32)
Social support	3.54 (1.24)	3.71 (0.96)	3.84 (1.31)	3.49 (1.25)

[a]Scores were based on a five-point Likert scale (higher scores indicating more helpfulness). Columns are broken down based on participant experience. Self: indicates the parent downloaded a health app; child: indicates that the health app is on the child's phone; none: indicates that the parent believes that neither they nor the child have downloaded a health app.

Table 3. Parents' ratings of helpfulness of overall TDF by health behavior.

Domain	Physical activity, mean (SD)	Diet, mean (SD)	Sleep, mean (SD)
Capability	3.81 (0.88)	3.86 (0.93)	3.68 (0.98)
Motivational	3.83 (0.87)	3.75 (0.89)	3.69 (0.92)
Opportunity	3.56 (0.97)	3.52 (1.02)	3.41 (1.04)

Within the reinforcement domain, there was no significant difference in parents' ratings of rewards and reinforcement, suggesting parents viewed mHealth apps that provide rewards for engaging in one of the behaviors or provide encouragement as equally helpful for their child's physical activity, diet, and sleep. There was a significant difference in parents' ratings of incentives ($F_{2,546}$=4.34, P=.01). Parents rated physical activity (mean 4.10, SD 1.12, P=.01; Cohen d=0.31) significantly greater than sleep (mean 3.75, SD 1.15); this suggests parents may view mobile apps that provide a reward for completing a task of a health behavior to be helpful for physical activity.

Opportunity

Overall, parents rated the opportunity component as helpful for physical activity (63.3%,116/183 reported "somewhat helpful" to "really helpful"), diet (60.1%, 111/183 reported "somewhat helpful" to "really helpful"), and sleep (57.4%, 105/183 reported "somewhat helpful" to "really helpful"). There was no significant difference in parents' ratings of environmental context and resources domain. These results suggest parents are likely to view mHealth apps that provide information on resources related to health behaviors or provide information on barriers or facilitators to be equally helpful for their child's physical activity, diet, and sleep.

Within the social influences domain, there was no significant difference in parents' ratings of the social comparison construct. Therefore, parents are likely to view mHealth apps that encourage social comparison or social support to be helpful for their child's physical activity, diet, and sleep. There was no significant difference between physical activity and diet or sleep. Parents may view mHealth apps that allow friends, family, and health care providers to give encouragement and support to be slightly more helpful for their child's diet than sleep. Table 2 provides parents' overall mean ratings by theoretical domains for general health behaviors.

To determine whether there were differences in parental preferences for the capability, opportunity, and motivation domains, we conducted one-way ANOVAs with post hoc t tests. In each case, the only significant difference was that parents viewed the opportunity domain as less helpful than either the capability or motivation domains. Table 3 provides parents' overall mean ratings of the helpfulness of general theoretical domains by physical activity, diet, and sleep.

Discussion

Results from this study indicate that fewer than one-fifth of parents have downloaded an app on their child's cell phone to help manage their health behavior. Despite current low adoption, approximately two-thirds of parents indicated a willingness to use a mobile app to help manage their child's diet and physical activity, whereas approximately half of all parents reported they would use an app to manage their child's sleep. Within the TDF domain of capability, parents were most interested in increasing their child's capability as it related to physical activity and diet relative to sleep. However, these differences were relatively small (ie, less than one-third of a standard deviation) suggesting relatively strong interest from parents for using mobile apps to improve their child's capability to perform all three health behaviors. Within the domain of motivation, parents were again generally interested in using mHealth apps to improve their child's motivation for healthy behavior. The most notable difference among the health behaviors was for goal setting; findings suggested that parents may most readily associate goal setting with physical activity. Comparing the domains of motivation and opportunity, parents were less enthusiastic about using mHealth apps to improve the opportunity for their children to engage in health behavior across all three health behaviors assessed.

It is noteworthy that parents generally reported lower preference for TDF strategies targeting their child's sleep. This is an area where one can imagine some discrepancy between the views of a behavioral scientist and a parent. For instance, the behavioral scientist may be interested in teaching a child about sleep hygiene using a mHealth app. However, parents' responses suggest that they would be less amenable to using an app for this purpose. One explanation may be that parents are not as informed about good sleep practices. It is common for laypeople to have suboptimal knowledge about healthy sleep practices, and to experience poorer sleep as a consequence [17]. In this case, it might be useful to not only work with parents to determine their preference for specific behavior change strategies, but also to educate potential users on good sleep hygiene before asking their opinions on how to change these behaviors. Based on our findings, recent calls for app developers and sleep experts to develop evidence-based guidelines for sleep apps should be expanded to include stakeholders because these groups are likely to see the function of apps through different lenses [18].

Parents were more enthusiastic about using mHealth apps to assist their children with behavioral regulation, with particular enthusiasm shown for physical activity behavior. Strategies such as self-monitoring and goal-setting skills for physical activity were among the most highly preferred for parents. These features are often well-integrated as activity reminders or prompts to set goals in existing mHealth platforms [5]. Parents were less enthusiastic about leveraging social influences to motivate their child's activity. These findings are consistent with previous work in young adults that found participants are likely to judge self-regulation behavior change techniques as central to physical activity app efficacy, while relatively devaluing the importance of social features [19].

Despite our efforts to recruit a nationally representative sample using Amazon MTurk, our study overrepresents Native Americans and Asians and underrepresents African Americans relative to the 2015 US Census. This is a limiting factor for the current findings because there is well-documented disadvantage conferred on ethnic minority groups in terms of both disproportionate risk and lack of available interventions [20,21]. mHealth approaches can be an equalizing force because of the high adoption of mobile phone technology in minority groups [22]; however, representative samples from stakeholder groups are a necessary first step toward building an equitable intervention framework. Our sample is diverse, but not fully representative. Moreover, there is the potential for selection bias as our sample is only representative of high-performing workers on the MTurk platform who are also parents. Our study did not consider developmental age or cognitive ability as a driver of parent preferences.

Our study did not explicitly assess important factors such as usability. It is relatively well-established that usability modifies the degree to which a user perceives a given behavior change technique as central to the efficacy of an app [23,24]. Consequentially, our study is likely biased upward because respondents are likely assuming that apps are highly usable and that they would engage with them regularly. Our study did not involve stakeholders in the development of our survey. It is important to note that this report is not a definitive assessment of parent preferences. It is one low-cost approach to incorporating an end user and theory into the mHealth app development lifecycle.

It is well documented that evidence-based digital health solutions are not typically disseminated widely after validation. Perhaps even more concerning, the lessons learned from research studies or recommendations provided by expert consensus groups are not typically adopted in commercially available apps [5-7]. It has been argued that there needs to be an interdisciplinary bridge between for-profit technology companies and scientific labs to maximize the uptake of digital health interventions [1]. However, there are many barriers to such collaborations and research labs may need to do some preliminary work before forging a partnership with technologists.

This study presents a low-cost strategy for involving a large number of stakeholders in the discussion of how health behavior theory should be applied in a mHealth intervention. Our approach is innovative in that it took a scientific framework (ie, TDF) and made it digestible to parents so that they could then provide their opinions about features that might appear in a future app. Our survey items discriminated between various health behaviors allowing stakeholders to communicate the different health behaviors that they would like a TDF feature to change. Moreover, we were able to develop a set of consumer opinions about features that were directly linked to elements of the TDF. Similar approaches to app development may help to ensure that stakeholder opinions are included in theoretically sound app development. If successful, it may be possible to develop effective interventions with higher uptake among the target users because theoretically sound features that are rated as desirable by consumers can be included, while undesirable features (even if theoretically sound) could be excluded.

Acknowledgments

The authors would like to acknowledge Christina Amaro and Andrea Garcia for their assistance with this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Domain and survey questions.

[DOCX File, 28KB - mhealth_v6i12e192_app1.docx]

References

1. Cushing CC. eHealth applications in pediatric psychology. In: Roberts MC, Steele RG, editors. Handbook of Pediatric Psychology. New York: Guilford Press; 2017:201-212.
2. Fedele D, Cushing C, Fritz A, Amaro C, 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]
3. Cushing CC, Brannon EE, Suorsa KI, Wilson DK. Systematic review and meta-analysis of health promotion interventions for children and adolescents using an ecological framework. J Pediatr Psychol 2014 Sep;39(8):949-962 [FREE Full text] [doi: 10.1093/jpepsy/jsu042] [Medline: 24934246]
4. Mummah S, Robinson T, King A, Gardner C, 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]

5. Brannon EE, Cushing CC. A systematic review: is there an app for that? Translational science of pediatric behavior change for physical activity and dietary interventions. J Pediatr Psychol 2015 May;40(4):373-384. [doi: 10.1093/jpepsy/jsu108] [Medline: 25502745]
6. 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]
7. Breton ER, Fuemmeler BF, Abroms LC. Weight loss-there is an app for that! But does it adhere to evidence-informed practices? Transl Behav Med 2011 Dec;1(4):523-529 [FREE Full text] [doi: 10.1007/s13142-011-0076-5] [Medline: 24073074]
8. Rapoff M. Editorial: Assessing and enhancing clinical significance/social validity of intervention research in pediatric psychology. J Pediatr Psychol 2010 Mar;35(2):114-119. [doi: 10.1093/jpepsy/jsp102] [Medline: 19914955]
9. Kazdin AE. A model for developing effective treatments: progression and interplay of theory, research, and practice. J Clin Child Psychol 1997 Jun;26(2):114-129. [doi: 10.1207/s15374424jccp2602_1] [Medline: 9169373]
10. Schleider J, Weisz J. Using mechanical Turk to study family processes and youth mental health: a test of feasibility. J Child Fam Stud 2015 Jan 22;24(11):3235-3246. [doi: 10.1007/s10826-015-0126-6]
11. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [FREE Full text] [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
12. McGrady ME, Ryan JL, Brown GA, Cushing CC. Topical review: Theoretical frameworks in pediatric adherence-promotion interventions: research findings and methodological implications. J Pediatr Psychol 2015 Sep;40(8):721-726 [FREE Full text] [doi: 10.1093/jpepsy/jsv025] [Medline: 25817881]
13. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
14. Michie S, van Stralen M, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
15. Cushing C, Suorsa K. Intensive treatments for pediatric obesity: an acceptability analogue study. Child Health Care 2014 Aug 18;43(3):169-185. [doi: 10.1080/02739615.2013.837823]
16. Foster SL, Mash EJ. Assessing social validity in clinical treatment research: issues and procedures. J Consult Clin Psychol 1999 Jun;67(3):308-319. [Medline: 10369051]
17. Brown FC, Buboltz WC, Soper B. Relationship of sleep hygiene awareness, sleep hygiene practices, and sleep quality in university students. Behav Med 2002;28(1):33-38. [doi: 10.1080/08964280209596396] [Medline: 12244643]
18. Shin JC, Kim J, Grigsby-Toussaint D. Mobile phone interventions for sleep disorders and sleep quality: systematic review. JMIR Mhealth Uhealth 2017 Sep 07;5(9):e131 [FREE Full text] [doi: 10.2196/mhealth.7244] [Medline: 28882808]
19. Belmon L, Middelweerd A, Te Velde Saskia J, Brug J. Dutch young adults ratings of behavior change techniques applied in mobile phone apps to promote physical activity: a cross-sectional survey. JMIR Mhealth Uhealth 2015 Nov 12;3(4):e103 [FREE Full text] [doi: 10.2196/mhealth.4383] [Medline: 26563744]
20. Ogden C, Carroll M, Lawman H, Fryar CD, Kruszon-Moran D, Kit BK, et al. Trends in obesity prevalence among children and adolescents in the United States, 1988-1994 through 2013-2014. JAMA 2016 Jun 07;315(21):2292-2299. [doi: 10.1001/jama.2016.6361] [Medline: 27272581]
21. Barr D. Health Disparities in the United States: Social Class, Race, Ethnicity, and Health. Baltimore: JHU Press; 2014.
22. Pew Research Center. 2017 Jun 28. Roughtly three-quarters of Americans own a smartphone URL: http://www.pewresearch.org/fact-tank/2017/06/28/10-facts-about-smartphones/ft_17-06-27_smartphone_owners/ [accessed 2018-10-13] [WebCite Cache ID 738vTOyCu]
23. Hoj TH, Covey EL, Jones AC, Haines AC, Hall PC, Crookston BT, et al. How do apps work? An analysis of physical activity app users' perceptions of behavior change mechanisms. JMIR Mhealth Uhealth 2017 Aug 03;5(8):e114 [FREE Full text] [doi: 10.2196/mhealth.7206] [Medline: 28778846]
24. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and physical activity apps: perceived effectiveness by app users. JMIR Mhealth Uhealth 2016 Apr 07;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]

Abbreviations

HIT: Human Intelligence Task
mHealth: mobile health
TDF: Theoretical Domains Framework

Edited by G Eysenbach; submitted 09.01.18; peer-reviewed by J Pierce, A Psihogios, N Leonard, I Montagni; comments to author 15.02.18; revised version received 12.07.18; accepted 16.08.18; published 21.12.18.

Please cite as:
Cushing CC, Fedele DA, Brannon EE, Kichline T
Parents' Perspectives on the Theoretical Domains Framework Elements Needed in a Pediatric Health Behavior App: A Crowdsourced Social Validity Study
JMIR Mhealth Uhealth 2018;6(12):e192
URL: https://mhealth.jmir.org/2018/12/e192/
doi:10.2196/mhealth.9808
PMID:30578173

Original Paper

Usability of a Personal Air Pollution Monitor: Design-Feedback Iterative Cycle Study

Glen E Duncan[1], PhD; Edmund Seto[2], PhD; Ally R Avery[3], MS; Mike Oie[4], MS; Graeme Carvlin[2], PhD; Elena Austin[2], PhD; Jeffry H Shirai[2], MS; Jiayang He[2], MS; Byron Ockerman[2], BS; Igor Novosselov[2], PhD

[1]Washington State University, Spokane, WA, United States
[2]University of Washington, Seattle, WA, United States
[3]Washington State University, Everett, WA, United States
[4]Washington State University, Seattle, WA, United States

Corresponding Author:
Glen E Duncan, PhD
Washington State University
Box 1495
Spokane, WA, 99210-1495
United States
Phone: 1 5093587875
Email: glen.duncan@wsu.edu

Abstract

Background: There is considerable evidence that exposure to fine particulate matter (PM2.5) air pollution is associated with a variety of adverse health outcomes. However, true exposure-outcome associations are hampered by measurement issues, including compliance and exposure misclassification.

Objective: This paper describes the use of the design-feedback iterative cycle to improve the design and usability of a new portable PM2.5 monitor for use in an epidemiologic study of personal air pollution measures.

Methods: In total, 10 adults carried on their person a prefabricated PM2.5 monitor for 1 week over 3 waves of the iterative cycle. At the end of each wave, they participated in a 30-minute moderated focus group and completed 2 validated questionnaires on usability and views on research. The topics addressed included positives and negatives of the monitor, charging and battery life, desired features, and changes to the monitor from each previous wave. They also completed a log to record device wear time each day. The log also provided space to record any issues that may have arisen with the device or for general comments during the week of collection.

Results: The major focus group topics included device size, noise, battery and charge time, and method for carrying the device. These topics formed the basis of iterative design changes; by the final cycle, the device was reasonably smaller, quieter, held a longer charge, and was more convenient to carry. System usability scores improved systematically across each wave (median scores of 50-66 on a 100-point scale), as did median daily wear time (approximately 749-789 minutes).

Conclusions: Both qualitative and quantitative measures showed an improvement in device usability over the 3 waves. This study demonstrates how the design-feedback iterative cycle can be used to improve the usability of devices manufactured for use in large epidemiologic studies on personal air pollution exposures.

(JMIR Mhealth Uhealth 2018;6(12):e12023) doi:10.2196/12023

KEYWORDS

air pollution; methods; particulate matter; twins

Introduction

Considerable evidence proves that exposure to fine particulate matter (PM2.5) air pollution is related to various adverse health outcomes [1-9]. Despite regulation, PM2.5 remains a serious problem in the United States, where currently 20 counties affecting over 23 million people do not meet the federal PM2.5 standards [10], and even short-term increases in the PM2.5 levels may cause tens of thousands of excess deaths per year [9,11,12].

Although progress has been made in elucidating the biological mechanisms for PM2.5-related health effects, the cost, complexity, and burden on study subjects have made it difficult to conduct personal exposure assessments to better understand effects of acute and chronic exposures, where and when exposures occur, and how individual lifestyle factors affect exposures in humans under free-living conditions. Indeed, most epidemiologic studies have not measured true personal exposures. Instead, they have relied upon measurements made at central monitoring sites as exposure "surrogates," resulting in considerable concern within the exposure assessment community as to the impact of such exposure error on disease estimates [13]. More sophisticated geospatial models have been used to try to capture spatial variations within urban areas [14] but often, the assumption is that the modeled ambient concentration at a subject's residential address is a reasonable estimate of personal exposure, which is also incorrect because individuals are exposed to multiple locations in the course of daily living; for example, in the few studies that have used personal exposure monitoring instruments, substantial variations were found among individuals living within the same urban area and even within the same neighborhood [13,15,16]. Moreover, individuals tend to spend close to 90% of their time in indoor environments [17], and this is often not considered in air pollution epidemiologic studies. Recent meta-analyses of the issue have concluded that characteristics of the participants and their microenvironments can greatly affect the representativeness of such proxies and that greater attention is needed to evaluate the effects of measurement error [18,19].

The purpose of this study was to use the design-feedback iterative cycle to improve the usability of a portable PM2.5 monitor. This methodological paper describes the testing and refinement of the device for use in an epidemiologic study of personal air exposure measurements and clinical and biological outcomes in a large sample of twins recruited from a community-based registry. Although this study only reports on the usability aspects of the personal air pollution monitor, members of our team have published on the performance attributes of the sensor components for measuring PM2.5 and other endotoxins [20-25].

Methods

Recruitment

Participants for this study were drawn from the community-based Washington State Twin Registry. We chose to use twin pairs for the design-feedback iterative cycle portion of the study because twin pairs would ultimately be recruited for the epidemiologic cohort portion of the study. Only twin pairs who did not reside at the same address were eligible for the study; this ensured differential environmental exposures. For the present iterative cycle portion of the study, only pairs within the Puget Sound (King, Snohomish, Pierce, and Kitsap counties) were recruited because in-person assessments were required. The larger epidemiologic cohort portion of the study will enroll twin pairs from a wider geographic extent.

Three different versions of the personal air pollution monitor were given to participants across 3 cycles from February 2016 to August 2017. The same subjects took part in each of the 3 cycles; this ensured systematic feedback on the changes in the device over time. After each cycle, the research team used the feedback from the focus groups to modify the design of the monitor for use in the subsequent cycle.

Device Description

Briefly, all versions of the air pollution monitor were designed by researchers at the University of Washington for personal monitoring of PM2.5 exposures. The design requirements were to create a relatively small battery-operated wearable monitor that could provide continuous timestamped and geocoded data on PM2.5 exposures during the day (ie, at least 12 hours). Although the form factor and electronic design evolved with each version of the monitors tested in the study, generally, all monitors included an onboard real-time optical particle sensor, Global Positioning System (GPS), real-time clock, data logging to a memory card, 3-axis accelerometer for physical activity tracking, and sound sensor for noise exposure monitoring. Monitors also include a microblower and plastic cartridge assembly used to collect time-integrated sample particles throughout the period that the monitor is powered on.

Measures and Procedures

For the first cycle of the study, the twins came in for a study visit during which they received oral and written instructions on device use and for completing a daily wear log. The wear log was used to record wear time (start and stop times) for each day of the weeklong collection period. The log also provided space to record any issues that may have arisen with the device or any open-ended comments during the week of collection. They were also provided 2 validated questionnaires to assess device usability (the System Usability Scale, SUS [26]) and general views on research (Research Attitudes Questionnaire, RAQ [27,28]). The SUS contains 10 questions addressing the ease of use of the device. RAQ is a 7-question survey that measures the participants' attitude toward medical research. Both SUS and RAQ use a 5-point Likert scale for each question with 1 being "strongly disagree" and 5 being "strongly agree." For scoring SUS, 1 is subtracted from each odd item response and 5 from each even-numbered response. The converted responses are added up and multiplied by 2.5, providing a range of possible values from 0 (most negative experience) to 100 (most positive experience). The RAQ scale is created by summing the 7 questions to get a score that ranges from 7 to 28.

Each twin carried on their person the air pollution monitor, an ActiGraph accelerometer (ActiGraph WGT3X-BT, Pensacola, FL) for objective measurement of physical activity, and a GPS monitor (Qstarz BT-Q100XT, Taipei, Taiwan) to place exposures within a space and time framework [29]. The physical activity and GPS data are not reported in this paper because the main purpose of wearing those devices in the iterative cycle portion of the study was to ensure that it was feasible to wear all 3 devices in the epidemiologic cohort portion of the study. Participants collected data for 1 week, at which time participants returned the devices and participated in an in-person focus group. The second and third waves proceeded in the same way; however, the study materials were sent to participants via FedEx,

and the focus groups were conducted via teleconference. The focus groups lasted approximately 30 minutes and were moderated by 2 trained research staff. Topics addressed included positives and negatives of the air pollution device, device charging and battery life, desired features for subsequent versions, and pros and cons of changes to the device that occurred since the previous wave.

Results

The first focus group was completed by 5 twin pairs, the second by 4, and the final by 3 with the same participants carried over through each group. One pair was no longer interested in participating after the first wave of collection owing to increased responsibilities at work, and 1 pair was unavailable to participate in the final wave of collection because they were preparing for a multi-week vacation. The initial group of 10 consisted of 7 women and 3 men, the second group of 7 women and 1 man, and the final group of 5 women and 1 man. In total, 80% (8/10) participants were white, 80% (8/10) were aged >50 years, 80% (8/10) had completed some college, and 60% (6/10) were from households with incomes over the Washington state median based on Census data. Four of the twin pairs were monozygotic. A script for conducting the focus group was used to facilitate discussion. Notes were taken by a member of the research team, and the focus groups were recorded and transcribed by the research coordinator.

Text from both the focus group transcripts and the open-ended comments in the data collection log was analyzed to identify the main topics of importance. We created a corpus of both positive and negative open-ended comments and then converted the corpus to a document term matrix to determine which words were used most frequently. As shown in Table 1, the topics that emerged from the analysis across the 3 cycles of collection were device size, noise level, battery and charge time, and the method for carrying the device (belt clip, lanyard, etc). Focus group comments showed relatively slight improvements between cycle 1 and 2 with larger improvements from cycles 2 to 3; for example, the research team made the device substantially smaller from cycles 1 to 3 (initial size 5.21×0.91×2.76 inches, final size 4.25×0.88×2.56 inches; Figure 1). A muffler was added after cycle 2 to reduce the noise level of the device (from the microblower used to collect particles). In addition to the muffler, the cycle 3 device was programmed for intermittent use of the microblower for particle sampling instead of constant sampling, which also reduced the overall noise level.

The battery and charging protocol improved over the 3 cycles as well. After cycle 1, lights were added to the device to make it easier to determine when it was on and charging. The twins remarked on the helpfulness of the lights. After cycle 2, the battery was changed to extend its life and simplify the charging process. The new battery was expected to last for 16 hours, meaning the twins were able to run the device all day and then charge it overnight.

A final major point of discussion was how the device was carried. In cycles 1 and 2, participants were provided with different types of belt clips. All participants had issues with the belt clips for both cycles; several participants dropped the device owing to the clips or because they could not figure out how to best attach the clips to the device. For cycle 3, the research team moved away from the belt clips and provided a lanyard for the device (Figure 2). The lanyard could be worn around the neck or attached to a bag. Reactions to the lanyard were positive, though 1 device dropped after falling out of the lanyard. The research team has already addressed this issue by securing the devices within the lanyard with a cable tie or attaching the lanyard clip through a corner strap hole on the monitor's enclosure.

Figure 3 illustrates a box plot of the median wear time over the 3 cycles, showing an increase from 749 minutes per day (range 122-931 minutes) in cycle 1 to 789 minutes per day (range 594-847 minutes) in cycle 3. Figure 4 illustrates a box plot of median SUS scores over each cycle with usability scores improving from 50 in cycle 1 to 66.2 in cycle 3. In both figures, variability is shown as 1.5 times the inter-quartile range (1.5×interquartile range).

The RAQ scores demonstrated a favorable view of research in general; the average score for all participants was 24.3 (range 19-28). Moreover, focus group feedback demonstrated an interest in learning more about the purpose of the device and seeing the data being collected. Most of the participants did not feel they would want to own a device like this; however, they did bring up a few specific scenarios when the device would be helpful, such as if one had a respiratory issue or if one lived in an area with high levels of air pollution.

Table 1. Major thematic topics and comments from focus groups over 3 consecutive design-feedback iterative cycles for testing a personal air pollution monitor.

Major focus group topics	Cycle 1	Cycle 2	Cycle 3
Noise	Extremely irritating; concerned about others hearing it; hissing; obnoxious	Worse than first time; varied among devices; improved; hissing	Intermittent; positive reaction to cycled sampling; quieter; could still be quieter
Size	Clunky; bulky; sharp edges; cumbersome	Bulky; sharp edges	Smaller; reasonable size; liked clear device cover
Battery and charge	Couldn't tell if device is on or off; did not stay on; unclear whether it is charged and charging	Light helpful for charging; unreliable; stopped working	Stayed charged all day; simple to charge overnight
Method for carrying	Belt clip did not work; dropped device	Better, but still not ideal; attach clip before sending	Liked the lanyard; device dropped a few times

Figure 1. Size comparison of a personal air pollution monitor in cycle 2 (left) to cycle 3 (right) of the design feedback iterative cycle.

Figure 2. Lanyards were used in cycle 3 to allow participants to carry the air pollution monitor around the neck or attached to a bag.

Figure 3. Median System Usability Scale scores (100 scale) over 3 consecutive design-feedback cycles. SUS: System Usability Scale.

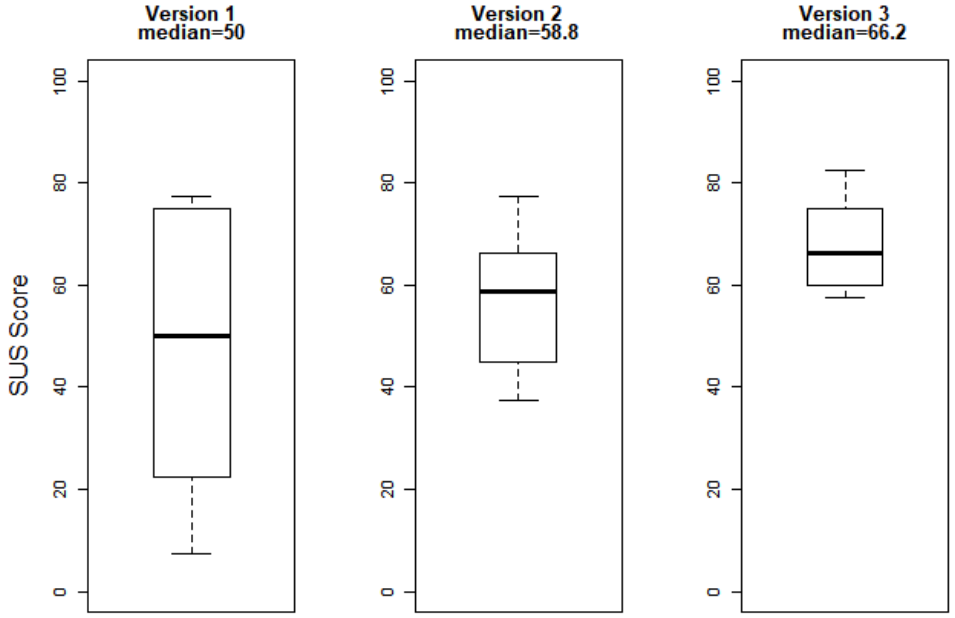
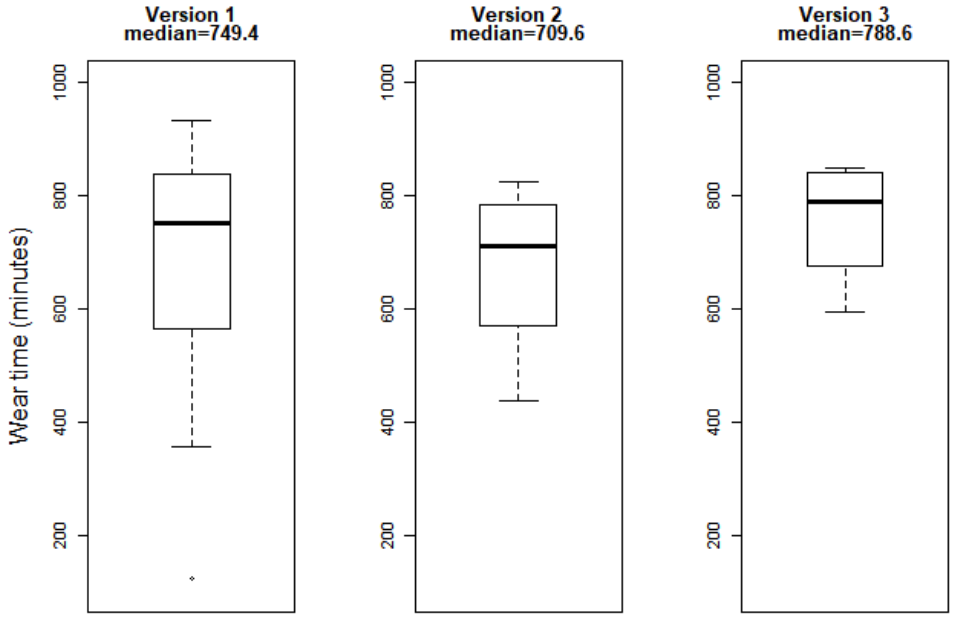


Figure 4. Median wear time (minutes per day) over 3 consecutive design-feedback cycles.



Discussion

Principal Findings

This methodological paper described how the design-feedback iterative cycle was used to improve the usability of a personal air pollution monitor. In general, data gathered from the focus groups and questionnaires showed an increase in satisfaction and usability with the device over each successive iterative cycle. The device will subsequently be deployed in a large epidemiologic study enrolling 150 twin pairs to examine associations between air pollution exposures measured in time and space on markers of inflammation and cardiometabolic risk. Because twin pairs will be used in the larger epidemiologic study, we chose to use twin pairs for device testing and feedback in this study. Although it is possible that the twins may have discussed data collection with each other and thus potentially influenced compliance and survey and focus group responses, we do not believe this was a major concern. First, it is important to note that the twins did not live together, a condition of

enrollment in the epidemiologic study because we are specifically interested in how "place" influences air pollution exposure. Thus, this level of intimate personal contact would not have occurred. Second, there was within-pair variability in wear time and in the comments provided on the questionnaires and during the focus groups, demonstrating that each twin member had his or her own unique experience using the device. Although twins generally agreed with one another on issues such as the size of the device and how much noise it made, these comments were made by others in the group as well. The focus group feedback and questionnaire data suggest that the device could still undergo an additional 1 or 2 iterative cycles to further improve usability, but any further changes could potentially result in a loss of functionality of the device; for example, the questionnaire data showed an increase in system usability scores over the 3 cycles. The SUS scores indicate that the user experience by the end of the design-feedback cycle had improved. Although this demonstrates a large relative improvement in usability, the final score still suggests further room for improvement. However, further changes to reduce size and noise level (still among the chief complaints) would likely have a negative effect on our ability to obtain valid and reliable data on PM2.5 air pollution. It is also important to note that participants were probably using their mobile phone or similar device as a size comparison, instead of similarly sized and commercially available personal air exposure monitors such as the MicroPEM (1.5×1.75×5 inches) or SidePak AM520 (5.1×3.7×3.1 inches).

Prior use and experience with a system or a tool can impact SUS scores on subsequent follow-up testing [30] and thus, it could be argued that participants provided better scores over time because of greater knowledge about the monitor (ie, the learning process). However, it is important to note that none of the participants had any experience with the device at baseline, so any learning that would have occurred would have been consistent across all participants. Introducing the device to novel participants at each wave would have introduced a number of important differences in subject characteristics (eg, age, sex, race or ethnicity, education level, and income level) that would have likely confounded the SUS results; therefore, we chose to keep the subjects consistent across waves. Finally, substantive changes were made to the device after each wave, including the addition of battery lights and a muffler, improved battery life, and a different carrying system for the device. All of these changes were repeatedly mentioned as positives in the focus groups and would have contributed to increased SUS scores, yet none of these changes could be attributed to increased knowledge of or experience with the device; rather, they were new additions specifically intended to improve usability.

We are not aware of any other studies that used SUS to evaluate performance in personal air pollution monitors. Thus, it is difficult to put our usability results in the context of other air pollution monitors. We identified another study that built a personal exposure monitor to measure particles as well as activity and location like ours, but it was only used for 6 hours in 1 individual [31]. Thus, once again, it is difficult to contextualize the results. The *Air Sensor Guidebook* (US Environmental Protection Agency [32]) notes that sensor performance requirements differ according to the specific application, making it difficult to compare our device to others. With respect to the increase in wear time, this is an important finding to us because our larger epidemiologic study will measure context-specific physical activity in addition to air pollution and a "valid monitoring" day in the physical activity literature is generally considered 600 minutes per day [33]. Thus, we are confident that our participants will comply with the wearing time aspect of the protocol and provide adequate physical activity data. The minimum wear time for a valid monitoring day in air pollution exposure studies is unknown, but we are confident that we will exceed the 600 minutes per day threshold, which should provide more than adequate data on personal air pollution exposures on a given day.

During the third cycle of collection, the Seattle region was experiencing an increase in smoke from wildfires in British Columbia, which was mentioned by the participants in the third focus group. They were curious as to what impact the wildfires would have on the data they were collecting. This speaks to the utility of the device; not only is the device a robust instrument for collecting PM2.5 air pollution data for research but it may also serve as a personal health monitoring tool to assess the impact of current air quality conditions in individuals with compromised health (eg, asthma). An example of the reports sent to participants is included in the Additional File, illustrating PM2.5 outdoor and personal air pollution during the data collection period for 2 members of a twin pair who are discordant for exposure. That the device may serve as a robust instrument for air pollution research is also supported by the participant's generally favorable view of the research study with an average RAQ score of 24.3 (on a 28-point scale).

Limitations

The main limitation of this study is the small sample size. Starting with only 5 pairs, we lost 1 pair in the second wave of the study, and 2 pairs in the third wave of the study. The study required both twins to participate and therefore, the loss of 1 member of a pair resulted in the loss of the full pair because we did not have the singleton twins collect data in the second and third waves. We also had a limited number of devices available for initial testing, which made pair-wise recruitment more difficult.

Conclusions

We used the design-feedback iterative cycle to improve the usability of a personal air pollution monitor for subsequent deployment in an epidemiologic cohort study. As attempts are made to decrease overall regional concentrations of PM2.5, identifying hotspots of exposure both in time and space, indoors and outdoors, will become increasingly relevant to protect public health. The availability of a low-cost, validated personal monitor that can measure multiple aspects of exposure, behavior, and context may greatly enhance our future ability to study the health impacts of these policy and planning changes.

Acknowledgments

The authors thank the staff and members of the Washington State Twin Registry for their participation and dedication. This work was supported by a grant from the National Institutes of Health (NIH; R21ES024715). The NIH played no role in the design of the study and collection, analysis, and interpretation of data, or in writing the manuscript and the decision to submit it for publication.

Conflicts of Interest

None declared.

References

1. Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, et al. An association between air pollution and mortality in six U.S. cities. N Engl J Med 1993 Dec 09;329(24):1753-1759. [doi: 10.1056/NEJM199312093292401] [Medline: 8179653]
2. Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, et al. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. Res Rep Health Eff Inst 2009 May(140):5-114; discussion 115. [Medline: 19627030]
3. Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 2002 Mar 06;287(9):1132-1141 [FREE Full text] [Medline: 11879110]
4. Pope CA, Turner MC, Burnett RT, Jerrett M, Gapstur SM, Diver WR, et al. Relationships between fine particulate air pollution, cardiometabolic disorders, and cardiovascular mortality. Circ Res 2015 Jan 02;116(1):108-115. [doi: 10.1161/CIRCRESAHA.116.305060] [Medline: 25348167]
5. Lipfert F, Wyzga R, Baty J, Miller J. Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: Long-term mortality in a cohort of US veterans. Atmospheric Environment 2006 Jan;40(1):154-169. [doi: 10.1016/j.atmosenv.2005.09.027]
6. Peng RD, Dominici F, Pastor-Barriuso R, Zeger SL, Samet JM. Seasonal analyses of air pollution and mortality in 100 US cities. Am J Epidemiol 2005 Mar 15;161(6):585-594. [doi: 10.1093/aje/kwi075] [Medline: 15746475]
7. Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. N Engl J Med 2000 Dec 14;343(24):1742-1749. [doi: 10.1056/NEJM200012143432401] [Medline: 11114312]
8. Zeger SL, Dominici F, McDermott A, Samet JM. Mortality in the Medicare population and chronic exposure to fine particulate air pollution in urban centers (2000-2005). Environ Health Perspect 2008 Dec;116(12):1614-1619 [FREE Full text] [doi: 10.1289/ehp.11449] [Medline: 19079710]
9. Pope CA, Dockery DW. Health effects of fine particulate air pollution: lines that connect. J Air Waste Manag Assoc 2006 Jun;56(6):709-742. [Medline: 16805397]
10. United States Environmental Protection Agency. 2018 Oct 31. PM-2.5 (2012) Designated Area/State Information URL: https://www3.epa.gov/airquality/greenbook/kbtc.html [accessed 2018-11-13] [WebCite Cache ID 73u9tg0VK]
11. Brook R, Rajagopalan S, Pope C, Brook JR, Bhatnagar A, Diez-Roux AV, American Heart Association Council on EpidemiologyPrevention, Council on the Kidney in Cardiovascular Disease,Council on Nutrition, Physical ActivityMetabolism. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation 2010 Jun 01;121(21):2331-2378. [doi: 10.1161/CIR.0b013e3181dbece1] [Medline: 20458016]
12. Simkhovich BZ, Kleinman MT, Kloner RA. Air pollution and cardiovascular injury epidemiology, toxicology, and mechanisms. J Am Coll Cardiol 2008 Aug 26;52(9):719-726 [FREE Full text] [doi: 10.1016/j.jacc.2008.05.029] [Medline: 18718418]
13. Sarnat JA, Wilson WE, Strand M, Brook J, Wyzga R, Lumley T. Panel discussion review: session 1--exposure assessment and related errors in air pollution epidemiologic studies. J Expo Sci Environ Epidemiol 2007 Dec;17 Suppl 2:S75-S82. [doi: 10.1038/sj.jes.7500621] [Medline: 18079768]
14. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 2005 Mar;15(2):185-204. [doi: 10.1038/sj.jea.7500388] [Medline: 15292906]
15. Sarnat JA, Koutrakis P, Suh HH. Assessing the relationship between personal particulate and gaseous exposures of senior citizens living in Baltimore, MD. J Air Waste Manag Assoc 2000 Jul;50(7):1184-1198. [Medline: 10939211]
16. Monn C. Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. Atmospheric Environment 2001 Jan;35(1):1-32. [doi: 10.1016/S1352-2310(00)00330-7]
17. Klepeis N, Nelson W, Ott W, Robinson J, Tsang A, Switzer P, et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J Expo Anal Environ Epidemiol 2001;11(3):231-252 [FREE Full text] [doi: 10.1038/sj.jea.7500165] [Medline: 11477521]

18. Avery C, Mills K, Williams R, McGraw K, Poole C, Smith R, et al. Estimating error in using ambient PM2.5 concentrations as proxies for personal exposures: a review. Epidemiology 2010 Mar;21(2):215-223 [FREE Full text] [doi: 10.1097/EDE.0b013e3181cb41f7] [Medline: 20087191]
19. Schmidt HG, Leffringhausen W. [Therapy and results of infections of the large joints without bone injuries using septopal]. Aktuelle Traumatol 1985 Oct;15(5):222-231. [Medline: 2866684]
20. Austin E, Novosselov I, Seto E, Yost M. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor. PLoS One 2015;10(9):e0137789 [FREE Full text] [doi: 10.1371/journal.pone.0137789] [Medline: 26367264]
21. Gao M, Cao J, Seto E. A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. Environ Pollut 2015 Apr;199:56-65 [FREE Full text] [doi: 10.1016/j.envpol.2015.01.013] [Medline: 25618367]
22. He JY, Novosselov IV. Design and evaluation of an aerodynamic focusing micro-well aerosol collector. Aerosol Science and Technology 2017 May 11;51(9):1016-1026. [doi: 10.1080/02786826.2017.1329515]
23. Holstius D, Pillarisetti A, Smith K, Seto E. Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. Atmos. Meas. Tech 2014 Apr 30;7(4):1121-1131. [doi: 10.5194/amt-7-1121-2014]
24. Nieuwenhuijsen MJ, Donaire-Gonzalez D, Rivas I, de Castro M, Cirach M, Hoek G, et al. Variability in and agreement between modeled and personal continuously measured black carbon levels using novel smartphone and sensor technologies. Environ Sci Technol 2015 Mar 03;49(5):2977-2982. [doi: 10.1021/es505362x] [Medline: 25621420]
25. He J, Beck NK, Kossik AL, Zhang J, Seto E, Meschke JS, et al. Evaluation of micro-well collector for capture and analysis of aerosolized Bacillus subtilis spores. PLoS One 2018;13(5):e0197783 [FREE Full text] [doi: 10.1371/journal.pone.0197783] [Medline: 29847559]
26. Brooke J, McClelland I. Usability Evaluation in Industry. In: Jordan P, Thomas B, Weerdmeester BA, McClelland AL, editors. SUS - A quick and dirty usability scale. London: Taylor & Francis; 1996:189-194.
27. Kim SYH, Kim HM, McCallum C, Tariot PN. What do people at risk for Alzheimer disease think about surrogate consent for research? Neurology 2005 Nov 08;65(9):1395-1401. [doi: 10.1212/01.wnl.0000183144.61428.73] [Medline: 16275826]
28. Rubright JD, Cary MS, Karlawish JH, Kim SYH. Measuring how people view biomedical research: Reliability and validity analysis of the Research Attitudes Questionnaire. J Empir Res Hum Res Ethics 2011 Mar;6(1):63-68 [FREE Full text] [doi: 10.1525/jer.2011.6.1.63] [Medline: 21460589]
29. Hurvitz PM, Moudon AV, Kang B, Saelens BE, Duncan GE. Emerging technologies for assessing physical activity behaviors in space and time. Front Public Health 2014;2:2 [FREE Full text] [doi: 10.3389/fpubh.2014.00002] [Medline: 24479113]
30. Bangor A, Kortum P, Miller J. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
31. Zhuang Y, Lin F, Yoo E, Xu W. AirSense: A portable context-sensing device for personal air quality monitoring. In: AirSense: A portable context-sensing device for personal air quality monitoring. In: MobileHealth '15; Hangzhou, China: ACM; 2015 Presented at: MobileHealth '15; 2015; Hangzhou, China p. A.
32. Williams R, Vasu K, Snyder E, Kaufman A, Dye T, Rutter A, et al. United States Environmental Protection Agency. 2014. Air Sensor Guidebook URL: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NERL&dirEntryId=277996&simpleSearch=1&searchAll=air+sensor+guidebook
33. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]

Abbreviations

GPS: Global Positioning System
PM2.5: fine particulate matter
RAQ: Research Attitudes Questionnaire
SUS: System Usability Scale

Edited by G Eysenbach; submitted 23.08.18; peer-reviewed by D Donaire-Gonzalez, N Arora, M Das; comments to author 27.09.18; revised version received 10.10.18; accepted 13.10.18; published 21.12.18.

Please cite as:
Duncan GE, Seto E, Avery AR, Oie M, Carvlin G, Austin E, Shirai JH, He J, Ockerman B, Novosselov I
Usability of a Personal Air Pollution Monitor: Design-Feedback Iterative Cycle Study
JMIR Mhealth Uhealth 2018;6(12):e12023
URL: http://mhealth.jmir.org/2018/12/e12023/
doi:10.2196/12023
PMID:30578204

Original Paper

The Impact of a Maternal Education Program Through Text Messaging in Rural China: Cluster Randomized Controlled Trial

Ri-Hua Xie[1,2,3], RN, PhD; Hongzhuan Tan[4], PhD; Monica Taljaard[5,6], PhD; Yan Liao[5], BEng; Daniel Krewski[3], PhD; Qingfeng Du[2], MD; Shi Wu Wen[5,6,7], PhD

[1]Department of Nursing, Nanhai Hospital, Southern Medical University, Foshan, China

[2]General Practice Center, Nanhai Hospital, Southern Medical University, Foshan, China

[3]McLaughlin Centre for Population Health Risk Assessment, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[4]Xiangya School of Public Health, Central South University, Changsha, China

[5]Clinical Epidemiology Program, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada

[6]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[7]Obstetrics Maternal Newborn Investigation Research Group, Department of Obstetrics & Gynecology, University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
Shi Wu Wen, PhD
Clinical Epidemiology Program
Ottawa Hospital Research Institute
University of Ottawa
501 Smyth Road
Ottawa, ON, ON K1H 8L6
Canada
Phone: 1 6137378899 ext 73912
Email: swwen@ohri.ca

Abstract

Background: In recent years, attempts have been made to use mobile phone text messaging (short message service, SMS) to achieve positive results for a range of health issues. Reports on the impact of maternal education programs based on this widely available, inexpensive, and instant communication tool are sparse.

Objective: This study aimed to explore the impact of a maternal education program through text messaging.

Methods: We conducted a cluster randomized trial in a remote region in the Chinese province of Hunan between October 1, 2011, and December 31, 2012. We used county as the unit of randomization (a total of 10 counties), with half of the counties randomly allocated to the intervention arm (with maternal education material adapted from the World Health Organization being delivered by text messaging to village health workers and pregnant women alike) and the other half to the control arm (normal care without text messaging). Data on maternal and infant health outcomes and health behaviors were collected and compared between the 2 arms, with maternal and perinatal mortality as the primary outcomes.

Results: A total of 13,937 pregnant women completed the follow-up and were included in the final analysis. Among them, 6771 were allocated to the intervention arm and 6966 were allocated to the control arm. At the county level, the mean (SD) of maternal mortality and perinatal mortality rate were 0.0% (0.1) and 1.3% (0.6), respectively, in the intervention arm and 0.1% (0.2) and 1.5% (0.4), respectively, in the control arm. However, these differences were not statistically significant. At the individual level, there were 3 maternal deaths (0.04%) and 84 perinatal deaths (1.24%) in the intervention arm and 6 maternal deaths (0.09%) and 101 perinatal deaths (1.45%) in the control arm. However, the differences were again not statistically significant.

Conclusions: Adequate resources should be secured to launch large-scale cluster randomized trials with smaller cluster units and more intensive implementation to confirm the benefits of the text messaging–based maternal education program suggested by this trial.

Trial Registration: ClinicalTrials.gov NCT01775150; https://clinicaltrials.gov/ct2/show/NCT01775150 (Archived by WebCite at http://www.webcitation.org/74cHmUexo)

(JMIR Mhealth Uhealth 2018;6(12):e11213) doi:10.2196/11213

KEYWORDS

maternal education; text messaging; maternal health; infant health; cluster trial

Introduction

Background

Although maternal and infant death rates in China are not as high as those in some developing countries [1], they are still very high, about 300 per 100,000 deliveries for maternal mortality and 40 per 1000 births for infant mortality in remote areas. Many of these maternal and infant deaths may be avoidable if mothers/local health workers can learn how to better detect and manage pregnancy complications. The World Health Organization (WHO)'s Promotion of Perinatal Care Program [2] contains teaching aids and texts for maternity care education. These WHO materials have been validated and widely implemented in many regions worldwide, with varying levels of success.

In recent years, attempts have been made to use mobile phone text messaging (short message service, SMS) to achieve positive results for a range of health issues including treatment management and adherence [3-5], quality of life and well-being assessment [6], weight management [7,8], suicide prevention [9], smoking cessation [10], and other public health issues [11,12]. Attempts have also been made to use mobile phone text messaging to address issues related to maternal and child health, including interventions to support postabortion contraception [13], infertility treatment [14], prevention of mother-to-child HIV transmission [15], pregnancy nutrition intervention [16], management of gestational diabetes [17], adherence to postpartum care [18], lactation management [19], and infant feeding [20]. Although the overarching goal of mobile phone text messaging seeks to promote behavioral changes in both health care providers and the target population of interest, text messaging interventions evaluated to date have met with varying degrees of success [11,12].

Reports on the impact of maternal education programs based on this widely available, inexpensive, and instant communication tool in low- and middle-income countries are sparse. To our knowledge, none of the studies have yet tried to integrate WHO's maternity care education material with text messaging as a local maternal and infant health promotion tool in remote rural areas in China. In a systematic review, Amoakoh-Coleman et al identified 19 studies (10 intervention studies and 9 descriptive studies) of mobile health (mHealth) on various maternal and child health issues; they found that none of these studies attempted to integrate the WHO education materials in the education program and none of these studies directly assessed the effect of mHealth on maternal and neonatal mortality [21]. Mobile phones are popular in China (more than 50% of the population has a mobile phone) and accessible (with wireless networks spanning most remote areas), and mobile phone text messaging is affordable (<5 cents per message), making China an opportune place to implement a large-scale maternal education program using this text messaging–based health communication tool.

Objective

To obtain the empirical data needed to explore this cost-effective novel health promotion opportunity, we designed a cluster randomized trial to evaluate the potential benefits of implementing the WHO maternal education program using text messaging in a remote area in China. We chose a cluster randomized trial because it could be implemented in large scale at a lower cost and it might help to reduce contamination [22].

Methods

Ethical Approval, Trial Registration, and Reporting

We obtained approval from the Ottawa Hospital Research Ethics and Confidentiality Committee before commencing the proposed study (REB # 2011467-01H). We registered this trial in the ClinicalTrials.gov Protocol Registration System (registration number: NCT01775150). We followed the Consolidated Standards of Reporting Trials 2010 statement: extension to cluster randomized trials [23] in the reporting of the trial.

Study Region

This study was conducted in the northwestern region of the Chinese province of Hunan, a mountainous area comprising about 5 million residents. Basic maternity care in this area is provided by village health workers. There are several unique features that make the northwestern region of Hunan province the ideal location for a cluster randomized trial to evaluate the impact of a maternal health education program. First, the authority governing the whole region agrees to participate in this study, so that no further negotiation with local authorities is needed. As a result, possible bias introduced by selective participation is reduced [22]. Second, the region is quite homogeneous, thereby increasing the chance of obtaining a balanced randomization result. Third, half of the village birth attendants in the region have no formal training, and the other half have inadequate or outdated training [24]. Therefore, there is room for improvement through the proposed maternal education program. Fourth, although the region is not well developed, penetration of mobile phone is high (>70%), rendering an education program relying on health communications by mobile phone feasible.

Development of the Health Education Tool

On the basis of the WHO education materials, we developed a health education tool with mobile phone text messaging for village health workers and pregnant women alike. The original WHO education materials are written documents with both electronic and paper versions [2]. These documents are comprehensive, with 28 education modules on details regarding various maternity care including safe motherhood, parenthood, concept of risk and appropriateness for prenatal care visits and testing, labor and delivery, postpartum care, and breastfeeding.

Although the WHO education tool is well founded and validated, it is developed in English and is too long and too detailed to be sent by text messaging effectively. To make it a user-friendly,

text messaging–based tool acceptable by village health workers and pregnant women in the study region, it needs to be translated into Chinese and to be shortened and modified. First, we formed a multidisciplinary expert panel comprising a maternity care specialist, a midwife, 2 epidemiologists, a psychologist, and a nutrition scientist. The expert panel made decisions on every step of the translation and modification of the WHO education materials. Second, the original English version of WHO education materials was translated into Chinese by 2 researchers independently. Third, the 2 Chinese translation versions were compared with the original English version by the expert panel. The expert panel revised the inconsistent and inaccurate items to reach the final version of the translated Chinese version. The translated Chinese version was then shortened and modified by the expert panel. Finally, the 28 modules of the original WHO educational materials were packaged into 4 periods: first trimester, second trimester, third trimester, and postpartum. For postpartum, materials were packaged for maternal care and baby care separately. For each period, up to 7 text messages with specific educational instructions were included. The main contents of the text messages are displayed in Multimedia Appendix 1 (details available upon request). In addition, a few modifications were made to suit the local culture and lifestyle. For example, with respect to nutrition items, beef and dairy products were replaced with high-protein foods such as pork, fish, chicken, and egg, which are frequently consumed by local people in the region. The whole process followed a previously developed protocol that included rigorous and accurate translations and appropriate appraisal.

Sample Size

In the original design, we planned to use village as the unit of randomization. The sample size calculation determined that we required 1130 villages (565 villages for each arm) and 10 births per village for the 12 months of the trial (11,300 total births) to achieve 90% power to detect a relative reduction of 30% in the primary outcome (maternal and perinatal death rates) from a control arm rate of 4% using a 2-sided test at the 5% level of significance [22]. These calculations were based on an assumed intraclass correlation of 0.02 for the villages. With a reduction of 30% as the acceptable magnitude of effect for consideration by researchers and/or policy makers, the available study sample is sufficient to answer the study question.

Recruitment of Study Participants and Randomization

Due to logistical difficulties and budget constraints, we could not use village as the unit of randomization for the cluster trial and we had to use county instead. We, therefore, selected 10 counties in the region and randomly allocated half of the counties to the intervention arm (with text messaging instructions to be sent by county maternal and child health bureaus) and the other half to the control arm (routine care with no text messaging). An independent statistician unrelated to this trial generated the random allocation sequence, and the investigator in charge (RHX) allocated the 10 participating counties to the intervention arm and the control arm accordingly. Village health workers in the 10 participating counties were requested to monitor women of reproductive age who planned to have a baby during the study period, and once a woman was confirmed to be pregnant, she was considered eligible for this study and was recruited into the study by the village health worker. However, any woman who was unable to read or access text messaging through her own phone or her husband's or family member's phone was excluded. Recruitment was started on October 1, 2011, and ended on August 31, 2012.

Delivery of Educational Material

We worked with local mobile phone carriers and maternal and child health bureaus of the intervention counties to install the adapted WHO education material into their wireless telecommunication systems. Text messages containing education materials were delivered to village health workers and pregnant women in the 5 intervention counties according to the pregnancy period recorded by staff at county maternal and child health bureaus.

Data Collection

Data on mothers' residence (rural vs urban), gravidity, parity, pregnancy risk status (according to the Chinese national guideline), prenatal visit, prenatal screening, syphilis test, hepatitis B test, folic acid supplementation, mode of delivery, obstetric hemorrhage, maternal death, infant sex, birth weight, perinatal death, thyroid test, phenylketonuria test, and hearing test were collected from study participants by village health workers using the data collection form developed by the research team. Data were collected at the beginning of the diagnosis of pregnancy and in the 42 days postpartum (to meet the definition of maternal death).

Statistical Analysis

We first compared the distribution of baseline characteristics and then compared maternal and infant outcomes between the 2 arms. Cluster-level analyses proceeded after comparing means and medians of maternal and infant outcomes as proportions for each cluster to ensure that cluster proportions were approximately normally distributed. We then compared mean differences of the maternal and infant outcomes between the 2 arms using a standard unweighted *t* test. Supplementary analysis at the individual level was also performed. In the analysis at the individual level, random effects logistic regression analysis was used. To account for clustering by county, the county was specified as a random effect. To adjust for the small number of clusters, the Kenward-Roger method was used [25]. Odds ratio and 95% CI were expressed as the effect measures, using the control arm (no text messaging) as the reference. The analysis was adjusted for the following baseline characteristics: gravidity, parity, rural residence, household income, high-risk pregnancy status, and infant gender.

Results

Participants

Between October 2011 and August 2012, a total of 25,236 pregnant women were recruited into the study (13,332 in the intervention arm and 11,904 in the control arm). Of these, 13,937 (55.2%) women completed the follow-up and were included in the final analysis. Among them, 6771 women were in the intervention arm and 6966 in the control arm. Most of

the remaining 11,299 pregnant women excluded from the final analysis did not complete pregnancy before the study closing date (December 31, 2012), rather than being lost to follow-up (Figure 1).

Baseline Characteristics

Table 1 shows the baseline characteristics of the study population. The study region is a typical rural area of China, with the majority (>90%) of the residents living in rural areas. Maternal and infant baseline characteristics between the 2 arms were generally comparable (Table 1).

As means and medians of county-specific outcome measures were very similar (data available upon request), we did not take log transformations of the data but used *t* tests to compare outcomes between the 2 arms at the county level.

Figure 1. Flowchart of study participants of the text messaging trial in Hunan, China.

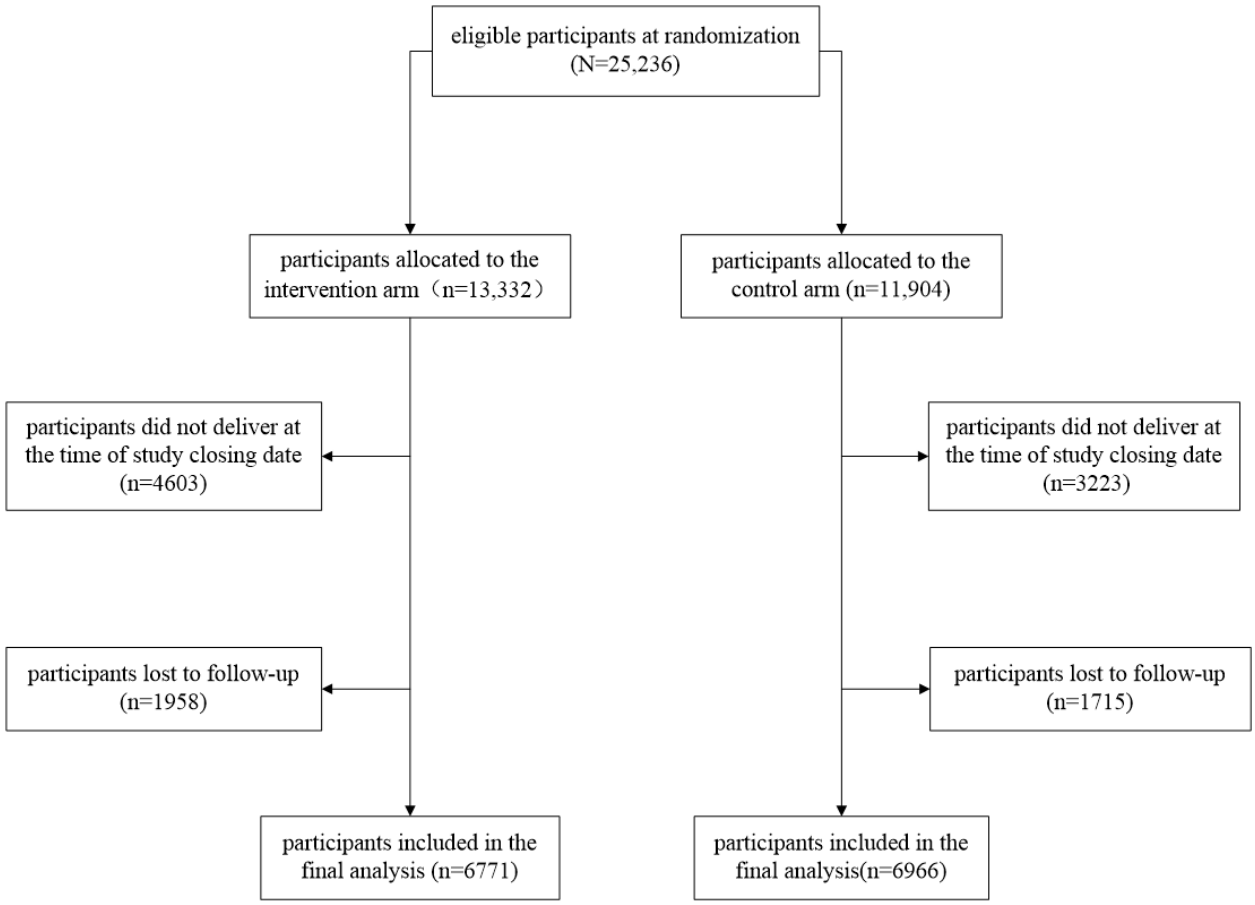


Table 1. Comparison of maternal and infant characteristics between the intervention and control arms at the individual level, Hunan, China, 2011-2012.

Characteristics	Intervention arm (N=6771), n (%)	Control arm (N=6966), n (%)
Gravidity		
>1	3285 (48.52)	3478 (49.93)
1	3486 (51.48)	3488 (50.07)
Parity		
>1	3883 (57.35)	4316 (61.96)
1	2253 (33.27)	2039 (29.27)
Rural resident	6388 (94.34)	6389 (91.72)
High-risk pregnancy	2488 (36.75)	2841 (40.78)
Fetal gender		
Female	3057 (45.15)	3639 (52.24)
Male	3684 (54.41)	3304 (47.43)

Table 2. Comparison of maternal and infant outcomes between the intervention and control arms at the county level, Hunan, China, 2011-2012.

Outcome	Intervention arm, % mean (SD)	Control arm, % mean (SD)	Percentage mean difference (95% CI)	*P* value (based on *t* test)
Early pregnancy visit	94.8 (2.3)	95.0 (1.1)	0.2 (−5.8 to 6.2)	.94
Prenatal screening	52.5 (12.7)	41.3 (14.3)	−11.2 (−30.9 to 8.5)	.23
Syphilis test	93.2 (7.1)	96.8 (3.3)	3.6 (−4.5 to 11.6)	.34
Hepatitis B test	94.9 (5.4)	98.3 (1.8)	3.4 (−2.4 to 9.3)	.21
Folic acid supplementation	75.6 (13.6)	78.4 (6.3)	2.7 (−12.8 to 18.2)	.69
Cesarean delivery	37.4 (6.8)	42.8 (16.0)	5.4 (−12.5 to 23.3)	.50
Obstetric hemorrhage	0.7 (0.5)	1.2 (0.8)	0.5 (−0.5 to 1.4)	.32
Maternal death	0.0 (0.1)	0.1 (0.2)	0.1 (−0.1 to 0.3)	.47
Perinatal death	1.3 (0.6)	1.5 (0.4)	0.1 (−0.6 to 0.9)	.66
Birth weight <2500 g	3.0 (0.5)	3.7 (2.1)	0.6 (−1.6 to 2.9)	.54
Birth weight >4000 g	1.5 (0.6)	1.6 (0.6)	0.1 (−0.8 to 1)	.80
Thyroid test	86.5 (10.2)	88.1 (5.0)	1.7 (−10.1 to 13.4)	.75
Phenylketonuria test	86.4 (10.2)	88.2 (4.9)	1.8 (−9.9 to 13.5)	.73
Hearing tests	2.3 (0.9)	2.1 (0.9)	−0.2 (−1.5 to 1.1)	.69

Table 3. Comparison of maternal and infant outcomes between the intervention and control arms at the individual level, Hunan, China, 2011-2012 (adjusted for gravidity, parity, residence, household income, high-risk pregnancy status, and gender of infant).

Outcomes	Intervention arm (N=6771), n (%)	Control arm (N=6966), n (%)	Crude odds ratio (95% CI)	Adjusted odds ratio (95% CI)
Early pregnancy visit	6308 (93.16)	6644 (95.38)	0.99 (0.98-1.00)	0.99 (0.99-1.00)
Prenatal screening	3291 (48.6)	2381 (34.18)	1.34 (1.28-1.39)	1.25 (1.21-1.31)
Syphilis test	6121 (90.4)	6625 (95.11)	0.98 (0.97-0.98)	0.98 (0.97-0.99)
Hepatitis B test	6229 (92)	6767 (97.14)	0.97 (0.96-0.98)	0.98 (0.97-0.99)
Folic acid supplementation	4733 (69.9)	5431 (77.96)	0.93 (0.91-0.95)	0.92 (0.90-0.93)
Cesarean delivery	2488 (36.75)	2927 (42.02)	0.88 (0.84-0.92)	0.95 (0.91-0.98)
Obstetric hemorrhage	47 (0.69)	77 (1.11)	0.64 (0.45-0.92)	0.42 (0.22-0.80)
Maternal death	3 (0.04)	6 (0.09)	0.52 (0.13-2.10)	__[a]
Perinatal death	84 (1.24)	101 (1.45)	0.96 (0.72-1.27)	0.73 (0.50-1.06)
Birth weight <2500 g	196 (2.9)	210 (3.02)	0.98 (0.81-1.19)	1.20 (0.97-1.47)
Birth weight >4000 g	97 (1.43)	105 (1.51)	0.97 (0.74-1.28)	1.01 (0.76-1.36)
Thyroid test	3977 (58.74)	4156 (59.66)	0.96 (0.95-0.98)	0.96 (0.95-0.98)
Phenylketonuria test	3911 (57.76)	4162 (59.75)	0.96 (0.95-0.98)	0.96 (0.95-0.98)
Hearing test	3935 (58.12)	3989 (57.26)	0.97 (0.95-0.98)	0.97 (0.95-0.99)

[a]Not estimable.

County Level Comparison

Table 2 compares outcomes between the 2 arms at the county level. Mean (SD) maternal mortality rates were 0.0% (0.1) and 0.1% (0.2), respectively, in the intervention arm and control arm. The corresponding means (SD) for the perinatal mortality rate were 1.3% (0.6) and 1.5% (0.4), respectively, in the intervention arm and control arm. However, these differences were not statistically significant (Table 2).

Individual Level Comparison

Table 3 displays results of analysis at the individual level. Of the 6771 participants, there were 3 maternal deaths (0.04%) and 84 perinatal deaths (1.24%) in the intervention arm, and of the 6996 participants, there were 6 maternal deaths (0.09%) and 101 perinatal deaths (1.45%) in the control arm. However, the differences were not statistically significant. For secondary outcomes, cesarean delivery (2488/6771, 36.75% cesarean deliveries) and obstetric hemorrhage (47/6771, 0.7% hemorrhage cases) rates were significantly lower in the intervention arm

than those in the control arm (2927/6966, 42.02% cesarean deliveries and 77/6966, 1.11% hemorrhage cases), both statistically and clinically (Table 3). No important differences between the 2 arms for other outcomes were observed (Table 3).

Discussion

Principal Findings

Our cluster randomized trial in a rural area in Hunan, China, found that it was feasible to deliver maternal education materials by text messaging through a mobile phone to village health workers and pregnant women simultaneously. The results did show some reduction in maternal mortality (3/6771, 0.04%) and perinatal mortality (84/6771, 1.24%) in the intervention arm as compared to the control arm (maternal death rate: 6/6996, 0.09% and perinatal death rate: 101/6996, 1.45%). For secondary outcomes, the rates of cesarean delivery and obstetric hemorrhage were lower in the intervention arm than those in the control arm, both statistically and clinically. This is also expected, as educated women should be better prepared; therefore, the need for cesarean delivery and the incidence of obstetric hemorrhage should be reduced. Although the observed associations between maternal education and maternal and infant outcomes were weak and not demonstrable after taking the cluster effect into consideration, these preliminary results are encouraging and deserve further investigation.

Strengths and Limitations

To our knowledge, this is the first study that evaluated the impact of WHO's maternity care education materials for local maternity care education in the remote rural area in China with text messaging. Through mobile phone–based text messaging, we are able to deliver the education materials to a large number of village health workers and pregnant women instantly. The cluster randomized trial is the appropriate design to assess the effect of a maternal education program, as it can be implemented with high efficiency and reduce the chance of contamination [22].

There are several reasons that may explain why our study failed to find an impact of a promising education tool delivered by an efficient method. First, because of implementation difficulties and budgetary constraint, we had to use the county as the unit of randomization. Originally, we planned to use villages as the units of randomization. We realized later that this would have made the trial cost prohibitive: with limited funding, we had to negotiate with local carriers for free text messaging service for this project, which the carriers agreed to only at the county level. There were also logistical considerations: villages lacked the manpower and expertise to deliver education material through text messaging. For these reasons, it was not feasible to use smaller units for randomization. As there were only 10 clusters (counties), we elected to use a more robust cluster-level method of analysis. A disadvantage of analysis at the cluster level is that there could be a loss in power because of data aggregation at the county level. Second, only about half of the recruited women were included in the final analysis. Most of the women were excluded not because they were lost to follow-up but because they had not yet delivered at the time of study termination (again because of budgetary constraints). The loss to follow-up is unlikely to introduce bias because both the intervention and control arms terminated at the same time. However, the substantial loss of study subjects resulted in lower power. Third, because of limited funding, we were not able to vigorously promote, implement, and monitor the maternal education program. For example, we did not track whether or not the village health workers and pregnant women received the text messages, actually read the messages, and if they found the text messages helpful. We did not have the capacity to provide additional assistance to the village health workers and pregnant women if they had difficulties understanding the messages or how to apply them to their own situations. As a result, the program may have not been implemented to the maximum extent possible, thus limiting its impact. Previous studies have suggested that to ensure the success of text messaging–based interventions, efforts should be made to intensively engage with the targeted population [19,20]. Fourth, we have based power calculation on maternal and perinatal mortality rates that were published more than 10 years ago [1]. Maternal and infant health has been improved substantially in the past decade in China, including rural areas [26], which further limited the study power of this trial.

Implications

Much of the mortality and morbidity in developing countries may be attributable to avoidable risk factors such as unhealthy diets, poor personal hygiene, unsafe delivery by birth attendants, and unintentional injuries; almost all these factors are modifiable [27-29]. For example, postpartum hemorrhage has been identified as one of the most important causes of maternal death in developing countries [27]. On the other hand, evidence generated from clinical investigations, mostly from the industrialized countries, has demonstrated that active management of the third stage of labor can substantially reduce the incidence of severe postpartum hemorrhage [30]. It is, therefore, reasonable to infer that if deliveries in developing countries were managed in the same manner as in industrialized countries, maternal deaths related to postpartum hemorrhage in these countries could be largely prevented. As another example, higher perinatal mortality in developing countries can be attributed in part to the lack of access to high-quality perinatal care for at-risk mothers, fetuses, and newborns [31-33]. Due to the emergent nature of the management of obstetric and neonatal complications and due to the difficulties in transferring at-risk mothers to nearby medical centers in a timely fashion in remote rural areas, instantly accessible information by mobile phone text messaging could provide a helpful tool for village health workers to manage obstetric and neonatal complications locally.

The mobile phone text messaging–based health education tool has been advocated by researchers and health organizations alike [11,12]. The scope and extent of use of this tool with respect to important population health issues have been expanded rapidly, with various trials being designed or launched on repeat suicidal episodes prevention [34], type 2 diabetes prevention [35], detection and management of hypertension in indigenous people [36], diabetes self-management in low- or middle-income countries [37], secondary prevention of coronary heart disease and diabetes [38], and increasing acceptability and

use of effective contraception among young women [39]. However, the impact of such a tool in reducing maternal and infant mortality and severe morbidity in low- and middle-income countries such as China has not yet been well documented. The key reason may be lack of rigorous evaluations from randomized trials, and it is crucial to evaluate its beneficial effects using cluster randomized trials.

Although our cluster randomized trial failed to find a statistically significant impact of the maternal education program delivered by text messaging through mobile phone on improved maternal and infant health and health behaviors, the advantages of text messaging in the field of maternal education should not be overlooked. It is able to deliver precisely packaged material to a massive population at a low cost. To send text messages to a massive population through a mobile phone, the senders need to work with local carriers. Therefore, these types of text messages were usually created and distributed by authoritative sources. On the other hand, messages delivered through social media platforms such as Facebook or WeChat, which could be distributed by anyone in the self-established social groups without scrutiny by experts, were often incorrect or even misleading. There is a general agreement that we need rigorous regulations, protocols, and ethical guidelines to correctly apply new technologies (mobile phone apps and text messages included) in the health care environment. Poorly validated information, often created by nonexperts, and a lack of updated data have been mentioned as concerning issues related to health smartphone apps. As such, authors urge different strategies that will provide higher quality evidence for smartphone apps' effectiveness and contents. This means that nonscientific or not evidence-based information spreading by text messages could be potentially dangerous to patients. Moreover, smartphones are needed to use social media platforms, which are often not affordable for people in remote areas.

Conclusions

In summary, a cluster randomized trial in a rural area in Hunan, China, suggests that it is possible to deliver maternal education material through text messaging to a massive population at a low cost. Although this exploration trial failed to demonstrate a statistically significant reduction of maternal and perinatal mortality or change in health behavior by maternal education through text messaging, several lessons learned from this exercise could help in the design and execution of future cluster randomized trials evaluating this intervention on maternal and infant health and other health issues. First, the choice of cluster unit for randomization requires balanced consideration. On the one hand, using smaller units such as villages is more efficient in terms of statistical analysis and study power. On the other hand, using larger units such as counties is much easier in the implementation of the trial at a much lower cost. However, using a larger unit of cluster will sacrifice statistical efficiency and study power. Second, to ensure the success of this type of intervention, vigorous promotion, implementation, and monitoring are needed. Finally, refined design considerations such as spreading the text messages after a small face-to-face meeting to explain the goals of text messaging intervention; developing a follow-up source to be sure the women can read and understand the received text messages; using images, videos, or other types of media format that may make the key concepts easier to understand; and assistance for those who may be in need could strengthen the intervention.

Acknowledgments

The authors thank the participating women, health care providers, local carriers, and local governments, particularly Mr Rong Xiang of the Huaihua Health Bureau, for their support to this project. This study was funded by Grand Challenges Canada (Grant#0029-01-04-01-01).. Additional support was provided by Canadian Institute of Health Research (FND-148438). DK is the Natural Sciences and Engineering Research Council of Canada Chair in Risk Science at the University of Ottawa.

Authors' Contributions

RHX, HT, and SWW conceptualized and designed the study; RHX, HT, MT, QD, DK, YL, and SWW made major contributions in the acquisition of data and analysis and interpretation of the data; RHX drafted the paper; SWW, HT, MT, and DK critically reviewed and revised the paper; all authors approved the final version of the manuscript. Each author certified that he or she had participated sufficiently in the work to believe in its overall validity and to take public responsibility for appropriate portions of its content.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The main contents of text messages sent to the health care providers and the women in Hunan, China during the four periods.

[PDF File (Adobe PDF File), 18KB - mhealth_v6i12e11213_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 3MB - mhealth_v6i12e11213_fig.pdf]

References

1. WHO. World Health Organization. 2005. Make every mother and child count URL: http://www.who.int/whr/2005/whr2005_en.pdf [accessed 2018-12-05] [WebCite Cache ID 74RTzYNt3]
2. Chalmers B, Mangiaterra V, Porter R. WHO principles of perinatal care: the essential antenatal, perinatal, and postpartum care course. Birth 2001 Sep;28(3):202-207. [Medline: 11552969]
3. Badawy SM, Thompson AA, Kuhns LM. Medication adherence and technology-based interventions for adolescents with chronic health conditions: a few key considerations. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e202 [FREE Full text] [doi: 10.2196/mhealth.8310] [Medline: 29273573]
4. Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth interventions to support self-management in HIV: a systematic review. Open AIDS J 2017 Nov 21;11:119-132 [FREE Full text] [doi: 10.2174/1874613601711010119] [Medline: 29290888]
5. Moore DJ, Jain S, Dubé MP, Daar ES, Sun X, Young J, et al. Randomized controlled trial of daily text messages to support adherence to preexposure prophylaxis in individuals at risk for human immunodeficiency virus: the TAPIR study. Clin Infect Dis 2018 May 02;66(10):1566-1572. [doi: 10.1093/cid/cix1055] [Medline: 29228144]
6. 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]
7. Griffin JB, Struempler B, Funderburk K, Parmer SM, Tran C, Wadsworth DD. My quest, an intervention using text messaging to improve dietary and physical activity behaviors and promote weight loss in low-income women. J Nutr Educ Behav 2018 Jan;50(1):11-18. [doi: 10.1016/j.jneb.2017.09.007] [Medline: 29325657]
8. Job JR, Fjeldsoe BS, Eakin EG, Reeves MM. Effectiveness of extended contact interventions for weight management delivered via text messaging: a systematic review and meta-analysis. Obes Rev 2018 Apr;19(4):538-549. [doi: 10.1111/obr.12648] [Medline: 29243354]
9. Berrouiguet S, Larsen ME, Mesmeur C, Gravey M, Billot R, Walter M, HUGOPSY Network, et al. Toward mHealth brief contact interventions in suicide prevention: case series from the suicide intervention assisted by messages (SIAM) randomized controlled trial. JMIR Mhealth Uhealth 2018 Jan 10;6(1):e8 [FREE Full text] [doi: 10.2196/mhealth.7780] [Medline: 29321126]
10. Noonan D, Silva S, Njuru J, Bishop T, Fish LJ, Simmons LA, et al. Feasibility of a text-based smoking cessation intervention in rural older adults. Health Educ Res 2018 Feb 01;33(1):81-88. [doi: 10.1093/her/cyx080] [Medline: 29309599]
11. Bogoch II, Lundin J, Lo NC, Andrews JR. Mobile phone and handheld microscopes for public health applications. Lancet Public Health 2017 Aug;2(8):e355 [FREE Full text] [doi: 10.1016/S2468-2667(17)30120-2] [Medline: 29253476]
12. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
13. Smith C, Ly S, Uk V, Warnock R, Edwards P, Free C. Process evaluation of a mobile phone-based intervention to support post-abortion contraception in Cambodia. Contracept Reprod Med 2017 May 01;2:16 [FREE Full text] [doi: 10.1186/s40834-017-0043-8] [Medline: 29201421]
14. Ronen K, Unger JA, Drake AL, Perrier T, Akinyi P, Osborn L, et al. SMS messaging to improve ART adherence: perspectives of pregnant HIV-infected women in Kenya on HIV-related message content. AIDS Care 2018 Apr;30(4):500-505 [FREE Full text] [doi: 10.1080/09540121.2017.1417971] [Medline: 29254362]
15. Fairbanks J, Beima-Sofie K, Akinyi P, Matemo D, Unger JA, Kinuthia J, et al. You will know that despite being HIV positive you are not alone: qualitative study to inform content of a text messaging intervention to improve prevention of mother-to-child HIV transmission. JMIR Mhealth Uhealth 2018 Jul 19;6(7):e10671 [FREE Full text] [doi: 10.2196/10671] [Medline: 30026177]
16. Huda TM, Alam A, Tahsina T, Hasan MM, Khan J, Rahman MM, et al. Mobile-based nutrition counseling and unconditional cash transfers for improving maternal and child nutrition in Bangladesh: pilot study. JMIR Mhealth Uhealth 2018 Jul 18;6(7):e156 [FREE Full text] [doi: 10.2196/mhealth.8832] [Medline: 30021707]
17. Rasekaba T, Lim K, Blackberry I, Gray K, Furler J. Telemedicine for gestational diabetes mellitus (TeleGDM): a mixed-method study protocol of effects of a web-based GDM support system on health service utilization, maternal and fetal outcomes, costs, and user experience. JMIR Res Protoc 2016 Aug 09;5(3):e163 [FREE Full text] [doi: 10.2196/resprot.6044] [Medline: 27507708]
18. Himes KP, Donovan H, Wang S, Weaver C, Grove JR, Facco FL. Healthy beyond pregnancy, a web-based intervention to improve adherence to postpartum care: randomized controlled feasibility trial. JMIR Hum Factors 2017 Oct 10;4(4):e26 [FREE Full text] [doi: 10.2196/humanfactors.7964] [Medline: 29017990]
19. Martinez-Brockman JL, Harari N, Pérez-Escamilla R. Lactation advice through texting can help: an analysis of intensity of engagement via two-way text messaging. J Health Commun 2018;23(1):40-51. [doi: 10.1080/10810730.2017.1401686] [Medline: 29236569]

20. Litterbach EK, Russell CG, Taki S, Denney-Wilson E, Campbell KJ, Laws RA. Factors influencing engagement and behavioral determinants of infant feeding in an mHealth program: qualitative evaluation of the growing healthy program. JMIR Mhealth Uhealth 2017 Dec 18;5(12):e196 [FREE Full text] [doi: 10.2196/mhealth.8515] [Medline: 29254908]
21. Amoakoh-Coleman M, Borgstein AB, Sondaal SF, Grobbee DE, Miltenburg AS, Verwijs M, et al. Effectiveness of mHealth interventions targeting health care workers to improve pregnancy outcomes in low- and middle-income countries: a systematic review. J Med Internet Res 2016 Aug 19;18(8):e226 [FREE Full text] [doi: 10.2196/jmir.5533] [Medline: 27543152]
22. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. United States: John Wiley & Sons Inc; 2000.
23. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. CONSORT 2010 statement: extension to cluster randomised trials. Br Med J 2012 Sep 04;345:e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
24. Annual report on the maternal and child health status in Hunan. Hunan: Hunan Ministry of Health; 2001.
25. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997 Sep;53(3):983-997. [Medline: 9333350]
26. Xi B, Zhou C, Zhang M, Wang Y, Xu L. Maternal and child mortality in China. Lancet 2014 Mar 15;383(9921):953-954. [doi: 10.1016/S0140-6736(14)60481-5] [Medline: 24629296]
27. Graham WJ. Now or never: the case for measuring maternal mortality. Lancet 2002 Feb 23;359(9307):701-704. [doi: 10.1016/S0140-6736(02)07817-0] [Medline: 11879885]
28. World Health Organization. Geneva: WHO; 1985 Nov. Prevention of maternal mortality: report of a WHO interregional meeting URL: https://tinyurl.com/y7pm7bru [accessed 2018-12-07] [WebCite Cache ID 74TtUscFS]
29. World Health Organization. Reducing risks, promoting healthy life. Geneva: WHO; 2002. URL: https://www.who.int/whr/2002/en/ [accessed 2018-12-07] [WebCite Cache ID 74Tty7QvH]
30. WHO recommendations for the prevention and treatment of postpartum haemorrhage. Geneva: World Health Organization; 2012. URL: https://tinyurl.com/yd9mrkpd [accessed 2018-12-05] [WebCite Cache ID 74RUwpOcZ]
31. Conde-Agudelo A, Belizán JM, Díaz-Rossello JL. Epidemiology of fetal death in Latin America. Acta Obstet Gynecol Scand 2000 May;79(5):371-378. [Medline: 10830764]
32. Langhoff-Roos J, Larsen S, Basys V, Lindmark G, Badokynote M. Potentially avoidable perinatal deaths in Denmark, Sweden and Lithuania as classified by the Nordic-Baltic classification. Br J Obstet Gynaecol 1998 Nov;105(11):1189-1194. [Medline: 9853768]
33. Sheiner E, Hallak M, Shoham-Vardi I, Goldstein D, Mazor M, Katz M. Determining risk factors for intrapartum fetal death. J Reprod Med 2000 May;45(5):419-424. [Medline: 10845177]
34. Larsen ME, Shand F, Morley K, Batterham PJ, Petrie K, Reda B, et al. A mobile text message intervention to reduce repeat suicidal episodes: design and development of reconnecting after a suicide attempt (RAFT). JMIR Ment Health 2017 Dec 13;4(4):e56 [FREE Full text] [doi: 10.2196/mental.7500] [Medline: 29237584]
35. Muralidharan S, Mohan V, Anjana RM, Jena S, Tandon N, Allender S, et al. Mobile health technology (mDiab) for the prevention of type 2 diabetes: protocol for a randomized controlled trial. JMIR Res Protoc 2017 Dec 12;6(12):e242 [FREE Full text] [doi: 10.2196/resprot.8644] [Medline: 29233806]
36. Yeates K, Campbell N, Maar MA, Perkins N, Liu P, Sleeth J, et al. The effectiveness of text messaging for detection and management of hypertension in indigenous people in Canada: protocol for a randomized controlled trial. JMIR Res Protoc 2017 Dec 19;6(12):e244 [FREE Full text] [doi: 10.2196/resprot.7139] [Medline: 29258978]
37. Abaza H, Marschollek M. SMS education for the promotion of diabetes self-management in low & middle income countries: a pilot randomized controlled trial in Egypt. BMC Public Health 2017 Dec 19;17(1):962 [FREE Full text] [doi: 10.1186/s12889-017-4973-5] [Medline: 29258499]
38. Huo X, Spatz ES, Ding Q, Horak P, Zheng X, Masters C, et al. Design and rationale of the Cardiovascular Health and Text Messaging (CHAT) study and the CHAT-Diabetes Mellitus (CHAT-DM) study: two randomised controlled trials of text messaging to improve secondary prevention for coronary heart disease and diabetes. BMJ Open 2017 Dec 21;7(12):e018302 [FREE Full text] [doi: 10.1136/bmjopen-2017-018302] [Medline: 29273661]
39. McCarthy OL, Osorio CV, Makleff S, Huaynoca S, Leurent B, Edwards P, et al. An intervention delivered by app instant messaging to increase acceptability and use of effective contraception among young women in Bolivia: protocol of a randomized controlled trial. JMIR Res Protoc 2017 Dec 18;6(12):e252 [FREE Full text] [doi: 10.2196/resprot.8679] [Medline: 29254910]

Abbreviations

mHealth: mobile health
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 03.06.18; peer-reviewed by E Dorronzoro, J Arrish, E Rincon; comments to author 06.10.18; revised version received 15.10.18; accepted 30.10.18; published 19.12.18.

Please cite as:
Xie RH, Tan H, Taljaard M, Liao Y, Krewski D, Du Q, Wen SW
The Impact of a Maternal Education Program Through Text Messaging in Rural China: Cluster Randomized Controlled Trial
JMIR Mhealth Uhealth 2018;6(12):e11213
URL: http://mhealth.jmir.org/2018/12/e11213/
doi:10.2196/11213
PMID:30567693

Original Paper

Preliminary Effectiveness of a Smartphone App to Reduce Depressive Symptoms in the Workplace: Feasibility and Acceptability Study

Mark Deady[1], BPsy (Hons), PhD; David Johnston[1,2], MClinPsych; David Milne[3], PhD; Nick Glozier[4], MA, MSc, MBBS, MRCPsych, FRANZCP, PhD; Dorian Peters[5], BHA, MMDes; Rafael Calvo[5], PhD; Samuel Harvey[1], MBBS, MRCGP, MRCPsych, FRANZCR, PhD

[1]Black Dog Institute, Faculty of Medicine, University of New South Wales, Sydney, Australia

[2]Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom

[3]School of Systems, Management and Leadership, Faculty of Engineering and IT, University of Technology, Sydney, Australia

[4]Brain and Mind Centre, Central Clinical School, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

[5]School of Electrical and Information Engineering, University of Sydney, Sydney, Australia

Corresponding Author:
Mark Deady, BPsy (Hons), PhD
Black Dog Institute
Faculty of Medicine
University of New South Wales
Hospital Road, Randwick
Sydney,
Australia
Phone: 61 293824517
Email: m.deady@unsw.edu.au

Abstract

Background: The workplace represents a unique setting for mental health interventions. Due to range of job-related factors, employees in male-dominated industries are at an elevated risk. However, these at-risk groups are often overlooked. *HeadGear* is a smartphone app–based intervention designed to reduce depressive symptoms and increase well-being in these populations.

Objective: This paper presents the development and pilot testing of the app's usability, acceptability, feasibility, and preliminary effectiveness.

Methods: The development process took place from January 2016 to August 2017. Participants for prototype testing (n=21; stage 1) were recruited from industry partner organizations to assess acceptability and utility. A 5-week effectiveness and feasibility pilot study (n=84; stage 2) was then undertaken, utilizing social media recruitment. Demographic data, acceptability and utility questionnaires, depression (Patient Health Questionnaire-9), and other mental health measures were collected.

Results: The majority of respondents felt *HeadGear* was easy to use (92%), easily understood (92%), were satisfied with the app (67%), and would recommend it to a friend (75%; stage 1). Stage 2 found that compared with baseline, depression and anxiety symptoms were significantly lower at follow-up (t_{30}=2.53; P=.02 and t_{30}=2.18; P=.04, respectively), days of sick leave in past month (t_{28}=2.38; P=.02), and higher self-reported job performance (t_{28}=−2.09; P=.046; stage 2). Over 90% of respondents claimed it helped improve their mental fitness, and user feedback was again positive. Attrition was high across the stages.

Conclusions: Overall, *HeadGear* was well received, and preliminary findings indicate it may provide an innovative new platform for improving mental health outcomes. Unfortunately, attrition was a significant issue, and findings should be interpreted with caution. The next stage of evaluation will be a randomized controlled trial. If found to be efficacious, the app has the potential to reduce disease burden and improve health in this at-risk group.

(JMIR Mhealth Uhealth 2018;6(12):e11661) doi:10.2196/11661

KEYWORDS

depression; workplace; mHealth; smartphone; eHealth; pilot

Introduction

Mental health conditions, and depression specifically, are leading causes of long-term disability globally [1,2]. Such disorders curtail and prohibit an individual's participation in basic activities of life including work [3]. The workplace has a complex relationship with mental well-being, as it is associated with both positive (eg, life satisfaction, personal autonomy, and confidence) [4,5] and negative mental health outcomes (eg, strain, stress, injury, and illness) [6]. Harvey et al's [7] model of psychosocial workplace risk factors highlights the complex relationship between work and the development of mental health problems as well as the potential for administering psychological interventions in the workplace. With recent Australian data indicating disability support payments for psychiatric conditions are on the rise with these conditions now being the leading cause of sickness absence [8], the development of effective interventions is a pertinent concern.

Due to the predominate role work has in individuals' lives, the workplace is increasingly being recognized as presenting a unique opportunity for both prevention and treatment of mental ill-health [9]. Although work strain is present across all industries, certain job-related factors make the issue more pertinent in some. Employees in male-dominated industries (MDIs; ie, those in which ≥70% workers are male, eg, agriculture, construction, mining, manufacturing, transport, and utilities [10]) have been found to be at heightened risk of mental health conditions [11,12]. This is likely due to a combination of job-related factors (eg, seasonal employment fluctuations leading to job insecurity, remote or isolated locations and family separation, and highly competitive, high-pressure work environments) [12] and the sociodemographic features of the employees themselves (eg, alcohol and substance abuse, low mental health literacy, and low rates of help seeking [13,14]). Related to—and compounding—both these areas is a traditional male attitude and workplace culture that has historically valued concepts of *toughness*, stoicism, and self-reliance [15,16]. Despite this need, little work has been specifically aimed toward these at-risk employees, with conventional prevention programs being poorly utilized by—or tailored to—these groups [17].

Electronic health (eHealth) and specifically mobile health (mHealth; health care practices supported by internet or mobile phone technologies) provide an opportunity to overcome some of the barriers present in traditional approaches to prevention and treatment [18]. Recent evidence suggests such interventions have utility in improving mental health outcomes in general [19-21], whereas workplace reviews have found eHealth interventions are effective at improving workers' psychological well-being, increase work effectiveness [22], and mental health and stress symptoms [23]. Although the dominant therapeutic approach in this area is cognitive behavioral therapy, there is increasing evidence that mindfulness and other approaches may hold distinct utility in this space [22-24]. Furthermore, the high rates of smartphone ownership increase the viability of mobile mental health care interventions [25]. However, this area is still in its infancy and little is known about the feasibility of such approaches in MDIs specifically.

Considering these findings and gaps in the knowledge base, we sought to develop a smartphone-based workplace intervention to reduce depressive symptoms and promote well-being, with a specific focus on MDIs. This paper presents a methodological framework, based on that of the Medical Research Council (MRC) [26]; elucidates the development and initial testing of the app; and details the 2-staged testing approach to finalizing the development of the program.

The aim of this study is to evaluate the usability, acceptability, feasibility, and preliminary efficacy of a newly developed app (*HeadGear*) designed to reduce depressive symptoms in an MDI working population.

Methods

Study Design

The model used to develop the app involved a process of research and analysis, development, implementation, and evaluation. In developing new technologies, it was important that the framework was systematic (clear steps following a logical order), systemic (all processes critical for success are incorporated), reliable (steps are clearly described so that they can be replicated by other designers in other projects), iterative (the cycle of analysis design development testing revision can be repeated a number of times), and empirical (data gathering is built into the process and decisions are made on the basis of data) [27]. The development process utilized a 3-step approach based on the intervention mapping protocol [28]. Similarly, processes have been used successfully for mHealth app-based interventions [29]. The predominate emphasis of this paper is that of the third step, as other steps have been reported elsewhere [30,31].

Employees in MDIs were specifically targeted. The process took place from January 2016 to April 2017. An interdisciplinary team of computer engineers, psychiatrists, psychologists, and design experts (user experience and graphic designers) collaborated in the design and development of the app.

Step 1: Defining the Problem

Although the effectiveness of eHealth and mHealth technologies for treating moderate levels of mental ill-health in general and clinical populations has been established, less is known about workplace eHealth interventions and eHealth prevention. The problem led the team to conduct a series of systematic reviews and meta-analyses to determine the effectiveness of workplace interventions for common mental disorders (CMD) [32], workplace depression prevention [33], the use of eHealth for prevention of CMD in general populations [34], and the use of eHealth tools for CMD in the workplace [23].

A recent meta-analysis of work-based depression prevention programs found such programs to be encouraging, with a number of different types of work-based interventions, particularly those based on cognitive behavioral models, demonstrating an ability to reduce depressive symptoms on unselected working populations [9].

Step 2: Participatory Engagement

To develop a relevant and engaging program, it is important to involve end users in participatory design and user experience research [35]. This stage comprised several components, including 6 focus groups (N=60) with industry partners and an in-depth survey of 1 specific industry partner (N=105). The findings of these are reported separately [30,31], and following feedback, we aimed to design and prototype an app that is engaging for men in the target workplaces.

Step 3: Design and Pilot Testing

Building on the outcomes of the initial 2 steps, the app content and design were finalized. The pilot testing of the app involved a 2-stage approach to test both initial utility and acceptability (using an alpha version [acceptance testing] of the app) and feasibility (engagement and perceived usefulness to users) and preliminary efficacy (using a beta version [operational testing] of the app). There were several reasons for this approach. Primarily, the costs involved in the creation of such technology are considerable. During the participatory engagement step, no prototype was used to generate unbiased input. However, it was necessary to be able to make modifications based on this testing and the usability of the app. Subsequently, modifications were made to the app between the 2 (alpha or beta) stages of the design and pilot testing step to refine usability elements and to test preliminary efficacy.

The App

HeadGear is a smartphone app–based intervention centered on behavioral activation and mindfulness therapy. The main therapeutic component of the *HeadGear* app takes the form of a 30-day challenge in which users' complete 1 *challenge* daily (approximately 5-10 min; Figure 1). These *challenges* include psychoeducational videos on coping skills or resiliency, mindfulness, and behavioral activation; mindfulness exercises; value-driven activity planning, goal-setting, and review; and coping skill development (problem solving, sleep, grounding, alcohol use, assertiveness, and training in adaptive forms of coping). The inclusion of these specific components was driven by the findings of stages 1 and 2 (specifically, [9,23,30,31,34]).

The first daily challenge involves the completion of a risk calculator, which assesses and provides participants with personalized feedback regarding their risk for future mental health issues. The risk calculator consists of 20 items developed from the Household, Income and Labour Dynamics in Australia Survey (HILDA) and has been validated in the Australian adult population [36]. The risk factor items are based on participant self-report. The HILDA risk items include age, gender, Aboriginal and Torres Strait Islander status, active career status, freedom to decide work, satisfaction with hours worked, satisfaction with employment opportunities, physical activity, alcohol use, episodes of distress in the previous 2 years, satisfaction with health, satisfaction with the neighborhood, satisfaction with partner, satisfaction with the way tasks are divided with partner, having someone to confide in, the feeling of being pushed around, and English as a second language. The HILDA questions and response items were replicated from the original items included in the HILDA survey, apart from age, which is measured here as a continuous measure. Users received personalized risk feedback immediately after completing the risk calculator. The personalized risk feedback involves an interactive icon array, which displays the calculated numerical risk estimate of developing anxiety and depression within the next year, along with a text description (Figure 2). Although much of the app is not specific to a workplace (or even MDI) setting and is likely to have utility to a general population, it was within these populations that development occurred. The outcomes of early development work [30,31] led to the inclusion of certain elements, determined to be the most relevant among these groups. Importantly, the risk algorithm was built from a working population sample and was fundamental to its working population delivery.

Other components of the app include a mood monitoring widget, a toolbox of skills (which is built from the challenges as they are completed), and support service helplines. Users had access to the app indefinitely. The app monitors use time and frequency and mission completion rates.

Stage 1: Alpha Testing—Utility and Acceptability

Participants

Participants (N=21) were recruited via email circulation and snowball recruitment from 3 industry partner organizations (agriculture, freight or postage, and mining). Study eligibility included Australian residency, aged between 18 and 65 years, valid email, ownership of an Apple- or Android-operating smartphone, ability to comfortably read English, and current employment. Consent was obtained electronically from all participants, and any identifiable data were encrypted to ensure confidentiality. The study acted in accordance with the Helsinki Declaration.

Procedure

Interested individuals were directed to the program's website to undergo screening and provide informed consent via the Web-based participant information statement. Participants completed a baseline questionnaire and were then invited to download the app. As an alpha version, iPhone users were required to download the app via a third-party app *Testflight*. Participants were encouraged to use the app for 30 days as often as they wanted. At the end of this period, users completed a follow-up questionnaire within the app (with 2 reminder emails sent to noncompleters). Daily engagement in the intervention was not incentivized, but successful completion of the posttrial questionnaire placed participants in the draw for an Aus $300 gift voucher. The study was approved by the University of New South Wales (UNSW) Human Research Ethics Committee (HC No: 16646).

Figure 1. Intervention component of the *HeadGear* app: the 30-day challenge (left), a behavioral activation day (middle), and a mindfulness day (right).

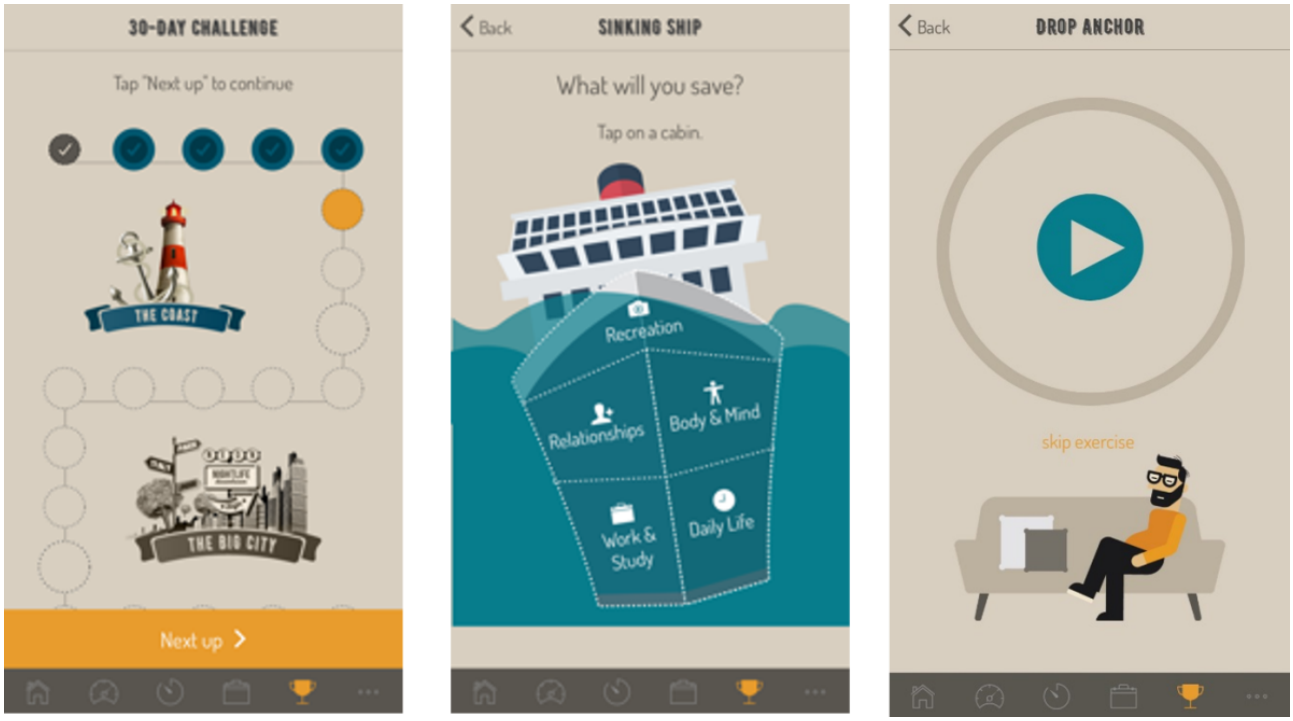


Figure 2. Additional features of the *HeadGear* app: mood widget (left), toolbox (middle), risk feedback (right).

Measures

Participants completed self-administered questionnaires within the app. Demographic information provided included age, sex, education, occupation, role, location, and industry group. The follow-up survey comprised the same measures as in the initial battery with the addition of a 26-item acceptability and usability questionnaire (comprising adapted items from the System Usability Scale [37]; Post Study System Usability Questionnaire [38]; Technology Assessment Model Measurement Scales [39]; and Usefulness, Satisfaction, and Ease questionnaire [40]), and this blended tool has been used successfully in previous research [41]. Participants were asked to rate their agreement with a series of statements about the intervention. Usage data were automatically collected by the app including time spent in app, number of logins, and specific responses to exercises.

Stage 2: Beta Testing—Feasibility and Preliminary Efficacy

Participants

Participants (N=84) were recruited via Facebook advertisements. All the advertisements were targeted (using Facebook's advertising platform) to males aged between 18 and 65 years, located within Australia, and employed in an MDI. Facebook allowed targeting of the following industries: agriculture, engineering, transport, forestry, mining, plumbing, and construction. Inclusion eligibility criteria were the same as those in stage 1.

Procedure

The advertising campaigns all ran simultaneously between July and August 2017. Advertisements were restricted to be shown only on mobile devices. Clicking anywhere on the Facebook advertisement directed interested individuals to the study website where they completed consent electronically. Confidentiality was assured via data encryption. After giving consent, individuals were asked to provide a mobile phone number. This number was verified by sending it a short message service (SMS) text message containing a random 4-digit code, which the individual was required to enter on the study website to continue. After a successful verification, the individual was sent (via SMS text message) a link that allowed them to download the *HeadGear* app via the *Google Play* or *iOS* app store, depending on their device. Participants then proceeded to an in-app questionnaire that collected demographic information and contained a number of study-specific measures (see below). At 5 weeks post baseline, participants were sent a text message (with 2 reminder texts sent to noncompleters), which directed them to the study data-collection site, and responded to a similar questionnaire (with the removal of demographic items and inclusion of some program feedback questions). Postintervention assessment occurred at 5 weeks post baseline to allow users 1 extra week to complete the 30-day program.

Daily engagement in the intervention was not incentivized, but successful completion of the posttrial questionnaire placed participants in the draw for a Aus $200 gift voucher. The study was approved by the UNSW Human Research Ethics Committee (HC17021).

Measures

The Patient Health Questionnaire-9 (PHQ-9) was used to measure depression symptoms [42]. The PHQ-9 is a reliable and valid 9-item measure of depression severity over the past 2 weeks [42,43]. Each of the 9 items of the PHQ-9 is scored as 0 (not at all), 1 (several days), 2 (more than half the days), or 3 (nearly every day). As a screening tool, summing the 9 item leads to a maximum score of 27 indicating all symptoms occurring nearly daily. The criterion and construct validity of the PHQ-9 have previously been demonstrated, with 73% sensitivity and 98% specificity in detecting major depression compared with clinician-based assessment [42,44], and regardless of diagnostic status, it typically represents clinically significant depression [42]. The measure has demonstrated excellent internal consistency (Cronbach alpha >.85 in multiple samples) and test-retest reliability of .84 [43].

Anxiety was measured using the 2-item Generalized Anxiety Disorder (GAD-2) scale [45]. The GAD-2 consists of the 2 core criteria for generalized anxiety disorder, which have also been shown to be effective screening items for panic, social anxiety, and posttraumatic stress disorders [45]. Equivalent to the parent scales, the PHQ-2 begins with the following stem question: "Over the last 2 weeks, how often have you been bothered by the following problems?" Response options are "not at all," "several days," "more than half the days," and "nearly every day," scored as 0, 1, 2, and 3, respectively (total ranging from 0 to 6). Scale scores of 3 or above are suggested as cut-off points between the normal range and probable cases of anxiety [45].

Resilience was measured by the Connor Davidson Resilience Scale (CD-RISC), a 10-item self-report scale demonstrated to be psychometrically sound with high internal consistency (Cronbach alpha=.89), construct validity, and test-retest reliability in the general population and in clinical settings [46]. Total scores range from 0 to 40 with higher scores corresponding to greater resilience. Validity is highly relative to other measures and reflects differentiation in resilience among diverse populations, showing that higher levels of resilience are consistent with lower levels of perceived stress vulnerability [46]. The CD-RISC has been shown to differentiate between individuals who function well after adversity from those who do not and measures the core features of resilience and the ability to tolerate experiences [47]. It is believed that increased resilience may reduce rates of mental ill-health [48].

Well-being was assessed using the 5-item World Health Organization Well-Being Index (WHO-5) [49,50]. Raw scores range from 0 to 25, where 0 indicates the worst possible quality of life and a score of 25 represents the best possible quality of life. A score less than or equal to 13 or an answer of 0 or 1 on any of the 5 items shows poor well-being. WHO-5 is a psychometrically sound measure of well-being with high internal consistency (Cronbach alpha=.84) and convergent associations with other measures of well-being [51].

Work performance was assessed using a modified version of the World Health Organization Health and Work Performance Questionnaire (WHO-HPQ) [52]. The WHO-HPQ is a self-report instrument designed to estimate the workplace costs of health problems in terms of self-reported sickness absence and reduced job performance (presenteeism). The absolute presenteeism score derived from this tool ranges from 0 (total lack of performance during time on the job) to 100 (no lack of performance during time on the job) with higher scores indicating less presenteeism. Absolute presenteeism was calculated, given it has been associated with better construct validity than the relative measure [53].

The WHO-HPQ was modified to simplify the absenteeism measure. Short-term absenteeism was assessed by asking "how many days/shifts have you missed over the past 4 weeks (28 days) due to sickness absence." If greater than 0, respondents were then asked, "how many of these sick days were due to mental health or emotional problems." For long-term absenteeism it was asked, "over the last 6 months have you had a continuous 1-week period of sickness absence." Following

this question respondents were asked, "if yes, was this due to mental health or emotional problems?"

Statistical Analysis

Sample Size

For stage 1, 25 individuals were sought to review the program. This number is not based on traditional power analysis calculations as our descriptive design precludes the ability to carry out power analyses. For this reason, we have drawn on previous studies in the field to guide in sample size determination.

Pilot studies (stage 2) tend to be underpowered to determine *proof-of-concept*. Additionally, the large sample size required for universal prevention work contribute to a lack of power in such pilot trials [54]. Despite this, for stage 2, using a 2-tailed test, with alpha set at *P*=.05 and power level of .80 (to detect a medium effect), a total of 40 participants was required. Due to expected high rates of dropout due to the unguided eHealth, general population, and nature of the study, an attrition rate of 50% was selected and a sample size of 80 was set.

Analysis Plan

All data were analyzed using IBM SPSS version 23.0 [55]. Stage 1 presents only descriptive statistics. In stage 2, descriptive statistics derived from participants' smartphone use data were used to characterize engagement and acceptability in the pilot study. Paired sample *t* tests were used to test for differences between pre- and posttrial clinical outcomes (eg, PHQ-9). Symptom change scores were computed, and linear regression was performed to test for the effect of time spent using the app, level of baseline risk, or the industry of employment on symptom change. Standardized effect sizes (Cohen *d* [56]) were calculated for outcomes of interest following the methods reported in the study by Lipsey and Wilson [57].

Results

Stage 1: Alpha Testing

In total, 21 participants downloaded the app, 12 of whom responded to the follow-up survey. However, 6 participants consented but did not download the app and were subsequently removed from the study. The average age of the participants was 37.86 years (SD=10.98); approximately half of the participants were female (n=12). The majority of the sample worked in freight and postage (n=11), followed by mining (n=6) and agriculture (n=2); however, 3 participants declined to provide their industry. Approximately half of the participants were working in a manager role (n=9) and the majority were based in an urban center (n=15).

Utility

Participants on average completed 5.71 challenge days (SD=9.02) and logged an average of 3.33 (SD=5.48) moods. Participants were asked to rate their agreement with a series of statements about the app's utility (see Table 1). Over 90% of participants reported that they believed most people would learn to use the app quickly and were satisfied with how easy the app was to use. Over 80% were comfortable using the app. The majority of negative feedback received came from 1 participant who only used the app to log 1 mood.

Acceptability

Table 2 shows respondents' rating of the app's acceptability. Over 90% of participants reported that they believed the information was easily understood and over 80% felt confident using the app. No respondent felt they needed to learn a lot of things before using the app. Over two-thirds of respondents were satisfied with the app, whereas 75% claimed it was fun to use, interactive, and that they would recommend it to a friend. Again, only 1 user reported substantial negative responses. There was a degree of concern about the utility of the app with only 40 to 50% of respondents claiming they would use it, or use it often, and 42% claiming the app worked the way they wanted it to. However, few actively disagreed with these statements.

Table 1. App utility questionnaire.

Statement	Disagree, n (%)	Neutral, n (%)	Agree, n (%)
I think that I would need the support of a technical person to be able to use the app	10 (83)	0 (0)	2 (17)
I found that the different parts of the app work well together	1 (8)	3 (25)	8 (67)
I thought there was too much inconsistency in the app	7 (58)	4 (33)	1 (8)
I would imagine that most people would learn to use the app very quickly	1 (8)	0 (0)	11 (2)
I found the app very awkward to use	10 (83)	0 (0)	2 (17)
Overall, I am satisfied with how easy it is to use the app	1 (8)	0 (0)	11 (92)
I was able to complete the "modules" quickly in the app	2 (17)	2 (17)	8 (67)
I felt comfortable using the app	1 (8)	1 (8)	10 (83)
Whenever I made a mistake using the app, I could recover easily and quickly	3 (25)	0 (0)	9 (75)
How things appeared on the screen was clear	1 (8)	2 (17)	9 (75)

Table 2. App acceptability questionnaire.

Statement	Disagree, n (%)	Neutral, n (%)	Agree, n (%)
I think that I would like to use the app often	1 (8)	6 (50)	5 (42)
I found the app to be very complicated	8 (67)	3 (25)	1 (8)
I felt very confident using the app	1 (8)	1 (8)	10 (83)
I needed to learn a lot of things before I could get going with the app	11 (92)	1 (8)	0 (0)
The information provided for the app was easy to understand	1 (8)	0 (0)	11 (92)
If I have access to the app, I will use it	1 (8)	5 (42)	6 (50)
I am satisfied with the app	1 (8)	3 (25)	8 (67)
I would recommend the app to a friend	1 (8)	2 (17)	9 (75)
The app is fun to use	1 (8)	2 (17)	9 (75)
The app helped me manage my symptoms	2 (17)	3 (25)	7 (58)
The app was interactive enough	1 (8)	2 (17)	9 (75)

Table 3. Demographics and app usage (N=84).

Characteristics	Statistics
Age in years, mean (SD)	38.62 (9.23)
Male, n (%)	84 (100)
Prior episode of mental ill-health, n (%)	48 (79)
Total active time in minutes, n (%)	58.24 (63)
Challenges completed, n (%)	9.11 (10)
Days used, n (%)	15.03 (16)
Industry, n (%)	
Male-dominated industry[a]	37 (45)
Nonmale dominated industry	45 (54)
Industry not provided	2 (2)
Role, n (%)	
General employee	55 (67)
Manager	19 (23)
Senior manager	8 (10)
Role not provided	2 (2)
Risk category [36], n (%)	
Low (≤4.5%; up to 25th percentile)	10 (12)
Average (4.6%-22%; 25th to 90th percentile)	21 (251)
High (≥23%; above 90th percentile)	30 (35)

[a]Agriculture or forestry or fishing, manufacturing, wholesale trade, mining, construction, other manual trade, transport or postal or warehousing, and first responder or defense or security.

Changes

A number of functionality and user interface and experience issues were resolved between stage 1 and 2. Additionally, changes were made to the app based on individual feedback. This included improved risk feedback (to better explain the feedback and direct users to elements in the challenge or external help), reminder functionality, a new booster session video added, goal-setting changes (to link values to both small and larger goals), and improvements to the skill toolbox (allowing for better integration with the challenge).

Stage 2: Beta Testing

The sample was entirely male, with almost half working in an MDI (Table 3). Participants had a mean age of 38 years (SD=9.23). On average, participants spent just under an hour in the app (mean=58.24 min; SD=62.98) and completed a third (mean=9.11; SD=10.25) of the challenge days. Over half (n=48)

of the participants reported a prior episode of mental ill-health and were considered high-risk on the HILDA-derived risk algorithm [36].

Preliminary Effectiveness

Although usage data were collected on all participants, only 34 (40.5%) completed follow-up questionnaires. No differences were found on any baseline data collected between responders and nonresponders; however, those responding to follow-up completed significantly more challenges ($t_{54.23}$=4.12; *P*<.001), app sessions ($t_{41.62}$=3.22; *P*=.002), and active time ($t_{41.49}$=3.38; *P*<.002). At 5-week follow-up, the *HeadGear* app was associated with significant reductions in depression symptoms (t_{30}=2.53; *P*=.02; Cohen *d*=0.39), anxiety symptoms (t_{30}=2.18; *P*=.04; Cohen *d*=0.38), and overall past month sick days (t_{28}=2.38; *P*=.02; Cohen *d*=0.22) and increases in self-reported workplace productivity (t_{28}=−2.09; *P*=.046; Cohen *d*=0.33). Trends toward improvement were found for well-being and mental health sick days, although these did not reach significance (Table 4).

Further analysis was conducted to determine whether improvement in depression and anxiety symptomatology was related to app usage. The results showed that there was a significant association between change in depression symptoms and time spent using app ($F_{1,31}$=6.08, *P*=.02; R^2=.164). Similarly, there was also a significant association between change in anxiety ($F_{1,29}$=5.35, *P*=.03; R^2=.174) and well-being ($F_{1,30}$=4.15, *P*=.049; R^2=.121) and time spent using the app. These results suggested that more time spent using the app was associated with a greater reduction in depression and anxiety symptomatology and a greater improvement in well-being. No other comparisons reached significance. Additional analysis indicated that the change in depression and anxiety symptomatology was not related to participants' level of risk category or industry type.

Feasibility and Feedback

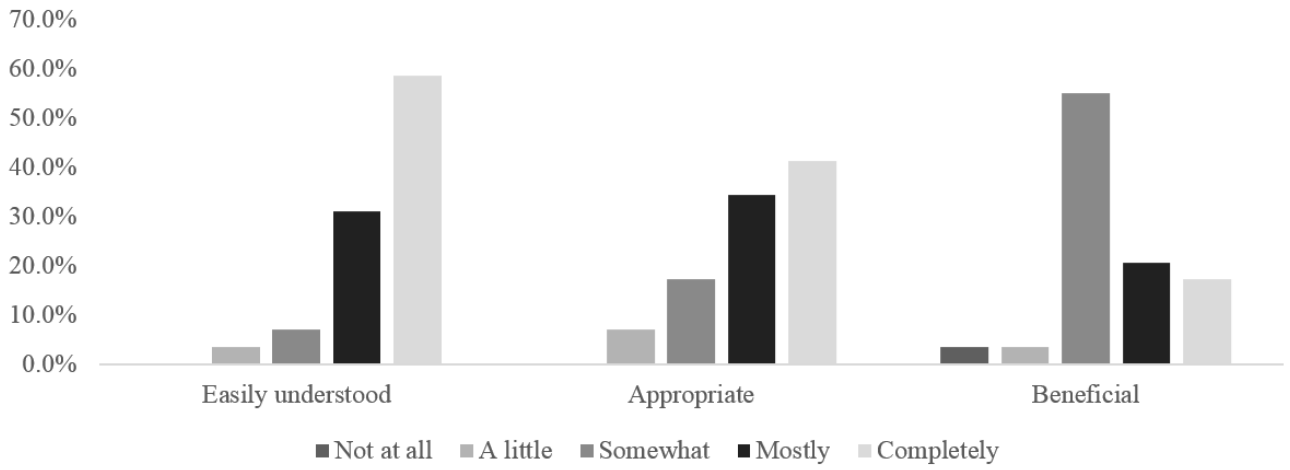
Figure 3 presents the basic feasibility of the program (n=34). Over three quarters (76%) of the respondents found the app to be mostly or completely appropriate for them, over 90% claimed it helped them improve their mental fitness (at least somewhat), and 90% found it mostly or completely understandable. Users were asked about the best and worst features of the app (stability, speed, look and feel, functionality, navigation, content, and other). Content was the most popular feature reported (46%), followed by both look and feel and functionality (23%). Navigation was the most highly ranked issue with the app (23%).

Table 4. Effectiveness outcomes. Italics indicates significance at the .05 level.

Outcome measure	Pretrial, mean (SD)	Posttrial, mean (SD)	Significance
Patient Health Questionnaire-9	12.00 (5.93)	9.68 (5.86)	*.02*
5-item World Health Organization Well-Being Index	9.29 (4.26)	10.00 (5.45)	.47
Connor Davidson Resilience Scale	23.57 (7.32)	23.27 (8.13)	.75
2-item Generalized Anxiety Disorder	2.77 (1.61)	2.16 (1.63)	*.04*
Absolute presenteeism[a]	53.79 (28.34)	63.10 (20.20)	*.046*
Sick days past month	2.31 (4.86)	1.24 (3.06)	*.02*
Mental health sick days past month	1.59 (4.87)	0.90 (2.85)	.10

[a]A score of self-reported workplace productivity (higher scores=greater productivity).

Figure 3. *HeadGear* pilot beta feedback (n=34).



Open feedback on the app was generally positive, with mindfulness and value-based goal setting highly regarded:

> *Improved my focus to make mindfulness a more consistent part of my day.*
>
> *Great app has really helped me look at all aspects of my life: work, relationships, interests, exercise, diet and mindfulness. This app has helped me manage my anxiety and depression.*

However, engagement and personal commitment were consistently raised as issues:

> *I wasn't able to sustain engagement with it. This was mostly through having some really good days. My mental health is constantly fluctuating. I think the content I saw was really good and I think if I had the time (didn't work so much) and was in a worse way [sic] would've used it more consistently.*

Some users reported disengaging from the longer challenges:

> *Disengaged from longer sessions, feeling like I wasn't acting on set actions without consequence.*

While creating time was also an issue:

> *I didn't make enough time to complete it,*
>
> *(You need to) Break up long sessions/ (have) time limited options.*

Only 8 respondents wanted to see additional features in the app; these features included a sleep tracker, ability to download and print, rescheduling of reminders (already present in the app), more reminders, a journal space, and longer mindfulness exercises.

Discussion

Principal Findings

This study provides both a framework for the development and testing of a new smartphone app intervention, *HeadGear*, and investigates the use and acceptability, along with the feasibility and preliminary effectiveness, of the app in a working population, specifically MDIs. The core features and functionality of the app were developed through a participatory design process, and the content of the app was based on current best available evidence-based theory. The research team encompassed computer engineers, psychiatrists, psychologists, and design (user experience and graphic design) experts allowing for a multidisciplinary approach to development. The pilot testing of the app incorporated a 2-stage process that utilized different samples and different outcomes measures to reflect the progression of the app from alpha to beta testing. Overall, the app was well received in both stages of the pilot testing, and preliminary testing indicated significant improvements in measures of psychopathology and workplace productivity.

The results from this feasibility and efficacy pilot trial suggest that an mHealth app can be an engaging, useful, and acceptable intervention. Across the 2 stages of the study, the majority of the participants acknowledged the utility, helpfulness, and overall ease and acceptability of use of the *HeadGear* app. With regards to preliminary efficacy of the intervention, the results are in line with previous findings that have shown mindfulness and behavioral activation to be effective in the treatment of mood and anxiety disorders [58,59], even in mobile app forms [24]. The dose-effect response seen between level of usage of the *HeadGear* app and improvements in both depression and anxiety symptoms was also encouraging; however, due to high attrition, findings need to be interpreted with caution.

Although improvements in well-being and resilience were found, these findings were not significant. As the sample included both *well* and *unwell* individuals, it is likely to have been underpowered to detect such changes; this underscores the need for a full-scale efficacy trial. Results also indicated that there were significant reductions in absenteeism and increases in worker productivity. This is especially encouraging given medical interventions in isolation have not shown as positive an effect on work-related outcomes when compared with workplace interventions [60]. This finding suggests the utility in incorporating evidence-based interventions in the workplace.

Strength and Limitations

Despite the positive reviews of stage 1, there was a low level of challenge days completed. This may reflect a number of functionality issues resolved for stage 2 and that the sample's characteristics were not representative of MDIs (from which this sample was taken and for which the app was designed) as 50% of participants were women. When contrasted with a technically improved iteration and a more representative population (stage 2), there was significantly more engagement with the app. Nevertheless, engendering motivation to complete the program was a concern raised in this review process. Although reasons for disengagement are complex and rarely only due to dissatisfaction [61] and somewhat unsurprising given the unguided nature of the trial [62,63], it does raise some feasibility concerns. O'Brien and Toms [64] suggest that engagement is not static but a process operating over a continuum; therefore, understanding this process more specifically across each of the challenge days might assist in improving adherence, which may be garnered through a larger trial. It was determined in earlier development steps [30,31] that end users were familiar with month-long health endeavors (eg, FebFast, Steptember, Dry July), and this played a role in the selection of the 30-day challenge period. Mobile apps, in general, suffer from poor rates of retention. Overall, 43% of global mobile users were still using apps (at least once) 1 month after download [65]. However, 23% will use an app only once, and only 1 in 3 will use an app at least 11 times [66]. Ultimately, this presents new obstacles in regards to engagement with a mental health and well-being app that need to be considered over the full intervention [67]. Encouragingly, results indicated the more time spent in the app was associated with more positive outcomes on the primary outcome and that participants used the app irrespective of their current symptom level, suggesting it has wider appeal than simply those with heightened symptomatology. Nevertheless, further research is required to better understand ways in which to enhance engagement.

A substantial strength of the study was the development process, which allowed for detailed and systematic analysis of a product

in multiple stages of testing. Additionally, the mobile-based delivery of the program holds a number of advantages over traditional methods particularly in MDIs [68]. Despite some limitations to generalizability, the study indicates that the intervention may have value in engaging this difficult-to-reach and at-risk group [69]. Indeed, tailoring the treatment to the feedback received during participatory user research meant that goal-directed and skill-based activities were the predominate focus of the intervention, which is in line with other recommendations for this group [70].

In addition to modest rates of intervention completion, the follow-up rate was also a limitation, and as mentioned, this has implications for the findings. Despite email (stage 1) and SMS text message (stage 2) reminders and incentives for assessment completion, follow-up rates were low compared with the literature [71]. Some reasons postulated for this include the source of recruitment (social media in stage 2), limited exclusion criteria (ie, those who were well may have had less motivation to engage), and a lack of personalized follow-up. However, a key factor which is unique to this trial is onboarding, whereby participants downloaded the app, consented, and subsequently completed baseline within the app. Therefore, users may have had little desire to participate in the trial but simply wanted access to the app. In an attempt to streamline the user experience (avoiding filtering participants through an arduous onboarding, which may lose all but the most conscientious participants), the study may, in fact, have recruited a less research-engaged (though perhaps more real-world) sample. Clearly, alternate and intensive strategies are required, as these low levels of retention raise feasibility concerns for a larger randomized controlled trial (RCT) trial. Additionally, low levels of mental health sick days were reported in the sample. This is unsurprising considering the small size and short follow-up; however, it limits what can be derived from this outcome. Sample size limited further investigation of change in outcomes based on baseline risk category or industry; again, larger RCT studies are required to explore these relationships. An additional limitation is despite targeting MDIs, there was significant interest from non-MDIs, and consequently, the conclusions that can be reached pertaining solely to MDIs are limited; conversely, the app may have wider utility. Finally, and perhaps most importantly, the lack of a control group limits any conclusions that can be made regarding the beneficial impact of this app; an RCT would help to ameliorate the biases inherent in uncontrolled trials.

Conclusions

The results from this pilot trial suggest that the *HeadGear* app can be an engaging, acceptable, and potentially effective intervention. Although preliminary results were encouraging, noted limitations in the pilot design highlight the need for a full-scale efficacy trial to better understand the utility of smartphone apps in the prevention and treatment of depression symptoms.

Acknowledgments

This study was developed in partnership with beyondblue with donations from the Movember Foundation. RC is funded by an Australian Research Council Future Fellowship FT140100824. SH and MD are supported by funding from the iCare Foundation and NSW Health.

Authors' Contributions

MD played a primary role in conceptualization, program development, data collection, analysis, write-up, and the editing of the manuscript. DJ had a role in data collection, analysis, and write-up of the manuscript. DM and RC were involved in program development; RC also edited the manuscript. NG and SH contributed to the conceptualization, program development, and editing of the manuscript. SH was also involved in writing the manuscript.

Conflicts of Interest

All authors declare that this is a beyondblue-funded study. All researchers have remained independent from the funders in the completion and submission of this work. MD, NG, DM, RC, and SH were involved in the development of the HeadGear app. The intellectual property is jointly owned by MD, NG, DM, RC, and SH; however, the authors do not currently receive any financial gain from the app. There are no other conflicts of interest to declare.

References

1. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61689-4] [Medline: 23245608]
2. Harvey SB, Henderson M, Lelliott P, Hotopf M. Mental health and employment: much work still to be done. Br J Psychiatry 2009 Mar;194(3):201-203. [doi: 10.1192/bjp.bp.108.055111] [Medline: 19252144]
3. Knudsen AK, Harvey SB, Mykletun A, Øverland S. Common mental disorders and long-term sickness absence in a general working population. The Hordaland Health Study. Acta Psychiatr Scand 2013 Apr;127(4):287-297. [doi: 10.1111/j.1600-0447.2012.01902.x] [Medline: 22775341]

4. Modini M, Joyce S, Mykletun A, Christensen H, Bryant RA, Mitchell PB, et al. The mental health benefits of employment: results of a systematic meta-review. Australas Psychiatry 2016 Aug;24(4):331-336. [doi: 10.1177/1039856215618523] [Medline: 26773063]
5. Stansfeld SA, Shipley MJ, Head J, Fuhrer R, Kivimaki M. Work characteristics and personal social support as determinants of subjective well-being. PLoS One 2013;8(11):e81115 [FREE Full text] [doi: 10.1371/journal.pone.0081115] [Medline: 24260545]
6. Stansfeld S, Candy B. Psychosocial work environment and mental health--a meta-analytic review. Scand J Work Environ Health 2006 Dec;32(6):443-462 [FREE Full text] [Medline: 17173201]
7. Harvey SB, Joyce S, Tan L, Johnson A, Nguyen H, Modini M, et al. Developing a mentally healthy workplace: A review of the literature. Sydney: National Mental Health Commission and the Mentally Healthy Workplace Alliance; 2014.
8. Harvey SB, Deady M, Wang MJ, Mykletun A, Butterworth P, Christensen H, et al. Is the prevalence of mental illness increasing in Australia? Evidence from national health surveys and administrative data, 2001-2014. Med J Aust 2017 Jun 19;206(11):490-493. [Medline: 28918719]
9. Tan L, Wang MJ, Modini M, Joyce S, Mykletun A, Christensen H, et al. Preventing the development of depression at work: a systematic review and meta-analysis of universal interventions in the workplace. BMC Med 2014 May 9;12:74 [FREE Full text] [doi: 10.1186/1741-7015-12-74] [Medline: 24886246]
10. Australian Bureau of Statistics. Canberra: Australian Bureau of Statistics; 2008. Australian and New Zealand standard industrial classification URL: https://tinyurl.com/y98qgfhx
11. Battams S, Roche AM, Fischer JA, Lee NK, Cameron J, Kostadinov V. Workplace risk factors for anxiety and depression in male-dominated industries: a systematic review. Health Psychol Behav Med 2014 Jan 1;2(1):983-1008 [FREE Full text] [doi: 10.1080/21642850.2014.954579] [Medline: 25750830]
12. Roche AM, Fischer J, Pidd K, Lee N, Battams S, Nicholas R. beyondblue. 2012. Workplace mental illness and substance use disorders in male-dominated industries: A Systematic Literature Review URL: https://www.beyondblue.org.au/docs/default-source/research-project-files/bw0203.pdf?sfvrsn=2 [accessed 2018-10-29] [WebCite Cache ID 73WzJzXBk]
13. Oliver MI, Pearson N, Coe N, Gunnell D. Help-seeking behaviour in men and women with common mental health problems: cross-sectional study. Br J Psychiatry 2005 Apr;186:297-301 [FREE Full text] [doi: 10.1192/bjp.186.4.297] [Medline: 15802685]
14. Rochlen AB, Paterniti DA, Epstein RM, Duberstein P, Willeford L, Kravitz RL. Barriers in diagnosing and treating men with depression: a focus group report. Am J Mens Health 2010 Jun;4(2):167-175 [FREE Full text] [doi: 10.1177/1557988309335823] [Medline: 19477750]
15. Courtenay WH. Constructions of masculinity and their influence on men's well-being: a theory of gender and health. Soc Sci Med 2000 May;50(10):1385-1401. [Medline: 10741575]
16. Stergiou-Kita M, Mansfield E, Bezo R, Colantonio A, Garritano E, Lafrance M, et al. Danger zone: Men, masculinity and occupational health and safety in high risk occupations. Saf Sci 2015 Dec 1;80:213-220 [FREE Full text] [doi: 10.1016/j.ssci.2015.07.029] [Medline: 27239098]
17. Lee NK, Roche A, Duraisingam V, Fischer JA, Cameron J. Effective interventions for mental health in male-dominated workplaces. Ment Health Rev J 2014 Dec 2;19(4):237-250. [doi: 10.1108/MHRJ-09-2014-0034]
18. 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]
19. 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 [FREE Full text] [doi: 10.1016/j.brat.2018.08.003] [Medline: 30125790]
20. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
21. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Dec 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
22. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. J Med Internet Res 2017 Jul 26;19(7):e271 [FREE Full text] [doi: 10.2196/jmir.7583] [Medline: 28747293]
23. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. PLoS One 2017;12(12):e0189904 [FREE Full text] [doi: 10.1371/journal.pone.0189904] [Medline: 29267334]
24. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]

25. 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;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
26. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655. [doi: 10.1136/bmj.a1655]
27. Hanington B, Martin B. Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions. Beverly, MA: Rockport Publishers; 2012:1-208.
28. Olson JS, Kellogg WA, editors. Ways of Knowing in HCI. New York, NY: Springer; 2014:XI 472.
29. Garnett C, Crane D, West R, Brown J, Michie S. The development of Drink Less: an alcohol reduction smartphone app for excessive drinkers. Transl Behav Med 2018 May 4:043. [doi: 10.1093/tbm/iby043] [Medline: 29733406]
30. Deady M, Peters D, Lang H, Calvo R, Glozier N, Christensen H, et al. Designing smartphone mental health applications for emergency service workers. Occup Med (Lond) 2017 Aug 1;67(6):425-428. [doi: 10.1093/occmed/kqx056] [Medline: 28535246]
31. Peters D, Deady M, Glozier N, Harvey S, Calvo RA. Worker preferences for a mental health app within male-dominated industries: participatory study. JMIR Ment Health 2018 Apr 25;5(2):e30 [FREE Full text] [doi: 10.2196/mental.8999] [Medline: 29695371]
32. Joyce S, Modini M, Christensen H, Mykletun A, Bryant R, Mitchell PB, et al. Workplace interventions for common mental disorders: a systematic meta-review. Psychol Med 2016 Mar;46(4):683-697. [doi: 10.1017/S0033291715002408] [Medline: 26620157]
33. Tan L, Wang MJ, Modini M, Joyce S, Mykletun A, Christensen H, et al. Preventing the development of depression at work: a systematic review and meta-analysis of universal interventions in the workplace. BMC Med 2014 May 9;12(1):74 [FREE Full text] [doi: 10.1186/1741-7015-12-74] [Medline: 24886246]
34. Deady M, Choi I, Calvo RA, Glozier N, Christensen H, Harvey SB. eHealth interventions for the prevention of depression and anxiety in the general population: a systematic review and meta-analysis. BMC Psychiatry 2017 Aug 29;17(1):310 [FREE Full text] [doi: 10.1186/s12888-017-1473-1] [Medline: 28851342]
35. Schuler D, Namioka A, editors. Participatory design: Principles and practices. Boca Raton, Florida: CRC Press; 1993:1-334.
36. Fernandez A, Salvador-Carulla L, Choi I, Calvo R, Harvey SB, Glozier N. Development and validation of a prediction algorithm for the onset of common mental disorders in a working population. Aust N Z J Psychiatry 2018 Jan;52(1):47-58. [doi: 10.1177/0004867417704506] [Medline: 28403625]
37. Brooke J. SUS-A quick and dirty usability scale. In: Patrick WJ, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. London: Taylor & Francis; 1996:189-194.
38. Lewis JR. Psychometric evaluation of the poststudy system usability questionnaire: the PSSUQ. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting.: Human Factors and Ergonomics Society; 1992 Presented at: Human Factor Society Annual Meeting; October, 1992; Santa Monica, CA. [doi: 10.1177/154193129203601617]
39. Venkatech V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
40. Lund A. Measuring usability with the USE questionnaire. Usability and User Experience 2001;8:8 [FREE Full text]
41. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253. [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
42. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
43. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
44. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. Psychiatric Annals 2002 Sep 1;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
45. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [Medline: 17339617]
46. Connor KM, Davidson JRT. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
47. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. J Trauma Stress 2007 Dec;20(6):1019-1028. [doi: 10.1002/jts.20271] [Medline: 18157881]
48. Southwick SM, Vythilingam M, Charney DS. The psychobiology of depression and resilience to stress: implications for prevention and treatment. Annu Rev Clin Psychol 2005;1:255-291. [doi: 10.1146/annurev.clinpsy.1.102803.143948] [Medline: 17716089]
49. Bech P. Measuring the dimensions of psychological general well-being by the WHO-5. QoL Newsletter 2004;32:15-16 [FREE Full text]

50. Bech P, Gudex C, Johansen KS. The WHO (Ten) Well-Being Index: validation in diabetes. Psychother Psychosom 1996;65(4):183-190. [doi: 10.1159/000289073] [Medline: 8843498]
51. Krieger T, Zimmermann J, Huffziger S, Ubl B, Diener C, Kuehner C, et al. Measuring depression with a well-being index: further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. J Affect Disord 2014 Mar;156:240-244. [doi: 10.1016/j.jad.2013.12.015] [Medline: 24412323]
52. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med 2003 Feb;45(2):156-174. [Medline: 12625231]
53. AlHeresh R, LaValley MP, Coster W, Keysor JJ. Construct validity and scoring methods of the World Health Organization: health and work performance questionnaire among workers with arthritis and rheumatological conditions. J Occup Environ Med 2017 Jun;59(6):e112-e1e8 [FREE Full text] [doi: 10.1097/JOM.0000000000001044] [Medline: 28598938]
54. Muñoz RF, Cuijpers P, Smit F, Barrera AZ, Leykin Y. Prevention of major depression. Annu Rev Clin Psychol 2010;6:181-212. [doi: 10.1146/annurev-clinpsy-033109-132040] [Medline: 20192789]
55. IBM SPSS Statistics for Windows, Version 23. Armonk, NY: IBM Corp; 2015. URL: http://www-01.ibm.com/support/docview.wss?uid=swg24038592 [accessed 2018-10-29] [WebCite Cache ID 73X2Hghc6]
56. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.
57. Lipsey MW, Wilson DB. Practical meta-analysis. Thousand Oaks, CA: Sage; 2001.
58. Ekers D, Webster L, Van Straten A, Cuijpers P, Richards D, Gilbody S. Behavioural activation for depression; an update of meta-analysis of effectiveness and sub group analysis. PLoS One 2014;9(6):e100100 [FREE Full text] [doi: 10.1371/journal.pone.0100100] [Medline: 24936656]
59. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183 [FREE Full text] [doi: 10.1037/a0018555] [Medline: 20350028]
60. Nieuwenhuijsen K, Bültmann U, Neumeyer-Gromen A, Verhoeven A, Verbeek J, van der Feltz-Cornelis CM. Interventions to improve occupational health in depressed people. Cochrane Database Syst Rev 2008 Apr 16(2):CD006237. [doi: 10.1002/14651858.CD006237.pub2] [Medline: 18425942]
61. Christensen H, Griffiths K, Groves C, Korten A. Free range users and one hit wonders: community users of an Internet-based cognitive behaviour therapy program. Aust N Z J Psychiatry 2006 Jan;40(1):59-62. [doi: 10.1111/j.1440-1614.2006.01743.x] [Medline: 16403040]
62. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
63. Lillevoll KR, Vangberg HC, Griffiths KM, Waterloo K, Eisemann MR. Uptake and adherence of a self-directed internet-based mental health intervention with tailored e-mail reminders in senior high schools in Norway. BMC Psychiatry 2014;14:14 [FREE Full text] [doi: 10.1186/1471-244X-14-14] [Medline: 24443820]
64. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
65. Localytics. Statista. 2017. Average three month user retention and churn rate of mobile apps worldwide as of 2nd half 2017 URL: https://www.statista.com/statistics/384224/monthly-app-launches-churn/ [accessed 2018-10-29] [WebCite Cache ID 73X2lkOe1]
66. O'Connell C. Localytics. 2016 Mar 26. 23% of Users Abandon an App After One Use URL: http://info.localytics.com/blog/23-of-users-abandon-an-app-after-one-use [accessed 2018-10-29] [WebCite Cache ID 73X2qkUOw]
67. Bickmore T, Schulman D, Yin L. Maintaining engagement in long-term interventions with relational agents. Appl Artif Intell 2010 Jul 1;24(6):648-666 [FREE Full text] [doi: 10.1080/08839514.2010.492259] [Medline: 21318052]
68. 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]
69. Seidler ZE, Rice SM, Oliffe JL, Fogarty AS, Dhillon HM. Men in and out of treatment for depression: strategies for improved engagement. Australian Psychologist 2017;53(5):405-415. [doi: 10.1111/ap.12331]
70. Ellis LA, Collin P, Hurley PJ, Davenport TA, Burns JM, Hickie IB. Young men's attitudes and behaviour in relation to mental health and technology: implications for the development of online mental health services. BMC Psychiatry 2013;13:119 [FREE Full text] [doi: 10.1186/1471-244X-13-119] [Medline: 23601273]
71. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]

Abbreviations

CBT: cognitive behavioral therapy
CD-RISC: Connor Davidson Resilience Scale
CMD: common mental disorders
eHealth: electronic health
HILDA: Household, Income, and Labor Dynamics in Australia Survey
MDI: male-dominated industries

mHealth: mobile health
MRC: Medical Research Council
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial
SMS: short message service
UNSW: University of New South Wales
WHO-5: 5-item World Health Organization Well-Being Index
WHO-HPQ: World Health Organization Health and Work Performance Questionnaire

Edited by C Dias; submitted 25.07.18; peer-reviewed by JE Bibault, M Wolf, R Picard, D Leightley; comments to author 02.09.18; revised version received 14.09.18; accepted 14.09.18; published 04.12.18.

Please cite as:
Deady M, Johnston D, Milne D, Glozier N, Peters D, Calvo R, Harvey S
Preliminary Effectiveness of a Smartphone App to Reduce Depressive Symptoms in the Workplace: Feasibility and Acceptability Study
JMIR Mhealth Uhealth 2018;6(12):e11661
URL: https://mhealth.jmir.org/2018/12/e11661/
doi:10.2196/11661
PMID:30514694

Original Paper

Creating Gameful Design in mHealth: A Participatory Co-Design Approach

Stian Jessen[1,2], MPhil Ed; Jelena Mirkovic[1], PhD; Cornelia M Ruland[1,2], RN, PhD

[1]Center for Shared Decision Making and Collaborative Care Research, Division of Medicine, Oslo University Hospital, Oslo, Norway

[2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Stian Jessen, MPhil Ed
Center for Shared Decision Making and Collaborative Care Research
Division of Medicine
Oslo University Hospital
Box 4950 Nydalen
Oslo, 0424
Norway
Phone: 47 980 49 759
Email: stian.jessen@rr-research.no

Abstract

Background: Gameful designs (gamification), using design pieces and concepts typically found in the world of games, is a promising approach to increase users' engagement with, and adherence to, electronic health and mobile health (mHealth) tools. Even though both identifying and addressing users' requirements and needs are important steps of designing information technology tools, little is known about the users' requirements and preferences for gameful designs in the context of self-management of chronic conditions.

Objective: This study aimed to present findings as well as the applied methods and design activities from a series of participatory design workshops with patients with chronic conditions, organized to generate and explore user needs, preferences, and ideas to the implementation of gameful designs in an mHealth self-management app.

Methods: We conducted three sets of two consecutive co-design workshops with a total of 22 participants with chronic conditions. In the workshops, we applied participatory design methods to engage users in different activities such as design games, scenario making, prototyping, and sticky notes exercises. The workshops were filmed, and the participants' interactions, written products, ideas, and suggestions were analyzed thematically.

Results: During the workshops, the participants identified a wide range of requirements, concerns, and ideas for using the gameful elements in the design of an mHealth self-management app. Overall inputs on the design of the app concerned aspects such as providing a positive user experience by promoting collaboration and not visibly losing to someone or by designing all feedback in the app to be uplifting and positive. The participants provided both general inputs (regarding the degree of competitiveness, use of rewards, or possibilities for customization) and specific inputs (such as being able to customize the look of their avatars or by having rewards that can be exchanged for real-world goods in a gift shop). However, inputs also highlighted the importance of making tools that provide features that are meaningful and motivating on their own and do not only have to rely on gameful design features to make people use them.

Conclusions: The main contribution in this study was users' contextualized and richly described needs and requirements for gamefully designed mHealth tools for supporting chronic patients in self-management as well as the methods and techniques used to facilitate and support both the participant's creativity and communication of ideas and inputs. The range, variety, and depth of the inputs from our participants also showed the appropriateness of our design approach and activities. These findings may be combined with literature and relevant theories to further inform in the selection and application of gameful designs in mHealth apps, or they can be used as a starting point for conducting more participatory workshops focused on co-designing gameful health apps.

(JMIR Mhealth Uhealth 2018;6(12):e11579) doi:10.2196/11579

KEYWORDS

gamification; gameful design; participatory design; co-design; mHealth; eHealth; self-management; chronic disease; patient requirements; patient participation; patient personal strengths

Introduction

Background

As smartphones and other mobile devices become increasingly ubiquitous, more and more mobile health (mHealth) tools and apps to support people living with chronic illnesses in self-management are becoming available. Although mHealth tools show great promise for supporting people with chronic illnesses [1,2], their success is often contingent on them being used as intended by the designers [3-5]. To increase the likelihood of users adhering to tools or services, borrowing design traits and approaches from the world of games, typically called either gamification or gameful design, has become increasingly popular over the past decade [6,7]. As opposed to serious games, which are games developed with an added instructional or normative purpose or takeaway [8], gameful designs refer to the use of game design approaches and techniques in otherwise nongameful or nongamelike situations, services, or tools to increase the user's enjoyment and motivation [8,9]. Typical applications and elements of gameful designs are, for example, competitions with either the app itself or other users, setting goals to accomplish, earning rewards such as points and badges, or having your own avatars [6,7,10,11]. Following Hamari et al's definition of gamification [9], whether or not something is to be considered gamefully designed is not connected to what specific elements one uses, but how these are applied and, in the end, experienced by the users. In this study, we approach this similarly and define gameful designs as using design approaches and implementations from the world of games (in our otherwise nongame tool) to add a sense of playfulness and increase users' overall enjoyment and engagement.

In the field of health and well-being, several gameful electronic health (eHealth) and mHealth tools for a range of different user groups and contexts have been created, such as smoking cessation [12], mental health [13], diabetes [14], medication adherence [15], and transitional care [16]. Still, some [10] also point to the limited number of gamified apps for health promotion in comparison with other fields such as education and business. Johnson et al [6], in a review of gamified tools for health and well-being, identified 19 empirical studies and reported that over half of the studies included had positive effects (59%), especially on behavioral outcomes such as physical activity, whereas the remaining 41% led to mixed or neutral outcomes. Even though many of these tools target changes in behavior, and there is an overlap with gameful design techniques and behavior change techniques [10,12], these are however not the same. If we consider again the definition of gameful designs as proposed by Hamari et al [9], this comes down to whether or not this is experienced as gameful by the users. Furthermore, and as reported by Johnson et al [6], gameful eHealth or mHealth tools also have the added possibility and potential to increase wellness and well-being by, for instance, providing pleasant designs and user experiences. The authors also found that the positive benefits of gamified mHealth tools are greater for users without preexisting motivation, compared with those already motivated to use the tools. Despite this, these findings should be interpreted with caution due to the relatively small number of studies currently published and their methodological limitations.

Design Guidelines for Developing Gameful Designs

At present, there is a dearth of guidelines, principles, or frameworks for designing and developing gameful designs that are empirically validated or evidence based [17]. From reviewing design frameworks for gamification, Mora et al [18] identified 40 frameworks, of which only 1 is in the field of health care [19], and specifically concerns the design of rehabilitation systems. Here, the authors proposed a detailed workflow of the overall design process, in addition to outlining specific suggestions for activities with stakeholders. This framework has, to our knowledge, not been evaluated. "The Wheel of Sukr" [20] is another set of guidelines, concerning the design of gameful mHealth apps for the self-management of diabetes. Even though this has been evaluated through a questionnaire regarding its content, it has not yet been practically tested [21].

Discussing design frameworks in general, Deterding [17] argues that these mostly consist of selecting typical gameful elements or parts, such as points, badges, or competitions from a predefined list, and fitting these to your design or solution. This makes them generic, thus not taking into account the well-known fact that the experience of gameful designs is context-dependent [17]. As such, there is no one-size-fits-all solution [22,23], and the gamefully designed tools need to fit both the users and the context in which they will be used [7,17,24]. Finally, and as with the health care–specific frameworks mentioned above, evaluations of the frameworks themselves are rarely conducted.

Thus, we can surmise that currently there are no validated frameworks for designing eHealth or mHealth tools gamefully [18] or that the road to success for gamefully designed tools is not found by following formulaic approaches, but is rather highly dependent on both the users' preferences and needs as well as the different contexts in which they are using the tools [17]. Even though there has been some investigation into people's preferences of gameful designs [11], such findings are typically decontextualized, and knowledge about users' specific needs and preferences for gameful and engaging designs is still mostly lacking [17].

User Participation in Design Processes

Even though there is a lack of evaluated frameworks for gameful designs, most proposed guidelines or frameworks as well as literature concerning gameful and engaging eHealth or mHealth tools, emphasize the importance and value of keeping the design processes user-centric [17,18,24-26].

User-centered design processes focus on the needs, interests, and requirements of the users [27]. These processes can be

placed in a continuum from expert-minded processes, which view users more as passive objects to be tapped for information, to participatory-minded processes, which include the users as co-designers [28]. *Participatory design* is firmly placed in the latter end of this continuum. More than just a design methodology, participatory design [29] takes the position that those whose future we are designing should not only have a voice but also a say in this process [30]. To achieve this, the approach is not only focused on the outcomes of design processes but also on the process itself, as it is a vehicle for enabling the co-designers' meaningful participation. Supporting this, the following are among the core tenets of participatory design: (1) mutual learning between participants and designers to better understand each other and the real-life situations in which the designs will eventually be used, (2) equalization of power relations by providing a voice to those who often do not have one in the society, and (3) using and designing tools and techniques that enable and support the participatory practices necessary to allow the participants to communicate and collaborate in the design processes [31], for example, by enacting real-life situations, playing design games, or exploratory prototyping [32].

Using Gamelike Design Activities in Co-Design Workshops

Previous literature has shown that framing design tasks in a gamelike manner can be well suited to support participants' easier understanding of the activities at hand by giving them clear rules and game-pieces as well as promoting their collaboration and creativity during the co-design processes [33-35]. For instance, Nicholas et al [34] used a version of the game Snakes and Ladders as a basis for the participants' design work. In another study, Brandt et al [36] describe a design game in which the players combine cards with pictures of situations, with cards presenting descriptive words to create stories about a persona. In general, design games typically share a randomized and open-ended nature, which can make it easier for the participants to create new and novel ideas [33,34].

A participatory design approach with gamelike activities should, therefore, be well suited for a design process that is not only sensitive to both the design goals of designers but also to the different preferences and needs of users as well as the different contexts in which the tool will be used. Even though there are published work related to using participatory approaches in the design of mHealth tools, rehabilitation games, and serious games [37-39], to our knowledge, little has been done in terms of co-designing gameful mHealth tools for people living with chronic illnesses.

Study Aims

This study is part of a larger research project funded by the Research Council of Norway, "The Power of Personal Strengths—using gamification to support patients in chronic illness management." The project's goal is to design and develop a gameful mHealth tool to help people living with chronic illnesses (long-term physical and psychological health challenges) [40] identify and use their own personal strengths to manage their everyday challenges of living with chronic conditions. The concept of personal strengths has its foundation in positive psychology [41] and can be defined as people's "positive traits reflected in thoughts, feelings, and behaviors" [42]. Simply put, a focus on strengths means emphasizing what is possible, valuable, and doable as opposed to only the deficit and problem focus one traditionally finds in medicine [43]. Previous research has shown that strength-based interventions among other can contribute positively to better moods and happiness [41] and increased general health and well-being [44]. Therefore, the main goal of the tool developed through this project is to, in a gameful and motivating fashion, help its users find and use their own personal strengths in overcoming their everyday challenges and how technology could help them do so.

Previously, we have reported on users' and stakeholders' needs and requirements of functionalities for the potential strength-based tool [45]. As the next step in our research project, this study describes co-design activities undertaken to inform and inspire the gameful and engaging designs of the self-management tool. Thus, the aims of this paper are twofold: (1) to explore new approaches for using participatory design methods in co-design sessions for designing a gameful mHealth intervention and (2) to identify user requirements and ideas for a gameful self-management for people living with chronic illnesses. As much of the existing publications concerning methods for gameful designs are terse in their descriptions of the creative design phases [24], this study's presentation will provide the reader with a detailed description of the workshop's activities, materials, and their rationale.

Methods

In this paper, we report on the methods applied to, and the outcomes from a series of 2 connected participatory co-design workshops exploring users' preferences and potential contexts of use for a gameful strength-based self-management tool for people with chronic illnesses.

Participants

For the workshops, participants were recruited through 2 hospital educational centers in the northern and southern parts of Norway, as well as the youth council at a hospital in the Oslo region. The criteria for participation were being fluent in Norwegian, having a long-term health challenge, and being over the age of 16 years. This study was approved by the privacy ombudsman at Oslo University Hospital, and all participants, or their legal guardians, signed informed consent forms before taking part. The participants each received a gift card valued at Norwegian krone 250 (approximately US $30) as compensation for participating in each of the 2 workshops.

In total, 22 participants, 14 female and 8 males, aged between 17 and 64 years (mean age 35.5 years) took part in the workshops. Due to illness and scheduling, not all participants from the first workshop were able to participate in the second, and 3 new participants were recruited (see Table 1 for the participants' background information and their distribution per workshop). All but 1 of the participants used a smartphone, and they on average rated themselves to 3.5 out of 5 on the question "how experienced are you with smartphones and or tablets."

Most used their phones for many more services other than just talking and messaging, and 13 of the 22 installed new apps at least monthly.

Process: the Workshops

We conducted 2 connected co-design workshops with each of the 3 different participant groups during the summer and autumn of 2017. The workshops were held at the premises of each of the 3 participating institutions and facilitated by the first author (SJ). Working primarily as an observer, a researcher or research assistant from the project group supported the facilitator and took notes and photographs. The first workshop focused on exploring ideas on how to use gameful techniques and approaches in the design of mHealth-related technologies. The second workshop both continued from and built on the output of the former, with an emphasis on helping users find and mobilize their personal strengths and how the technology could be used to support this process in an engaging and meaningful manner.

The workshops were designed in line with the ideals of participatory design [31,32], and as such, a significant amount of time was spent on learning activities to enable the participants to meaningfully take part. As we are introducing the participants to many new and advanced concepts, the learning activities were designed and organized along the lines of modern science classes, by starting with what people already know on a topic, presenting new information and organizing the new and old information, before finally reflecting on and applying the new knowledge [46]. For our workshops, this meant beginning with learning focused activities and gradually transitioning toward more open and design-focused activities as the workshops progress. Furthermore, the main design tasks were themselves designed to be gamelike activities, as this has been shown to both engage and put participants in a creative and innovative state of mind [33-35]. To keep the participants both engaged and active for the entirety of the workshops and not overload them cognitively with new and demanding concepts, ideas, and tasks [47], the workshops were planned to last for around 2.5 hours. As we are working with people with chronic illnesses, keeping the workshops shorter would also make participating less of a burden to them. In addition, having 2 shorter workshops as opposed to 1 long workshop would allow the participants to reflect on the content and concepts between the 2 gatherings.

Workshop 1

Introduction

The first workshop started with a round of introductions where everyone presented themselves before we gave a short presentation of our project and the reasoning behind it. We then explained the idea of using what makes games fun and motivating to create gameful designs. Through examples, we presented a range of games from different genres, fields, and contexts aiming to cover some games everyone liked (such as *Super Mario*, *Pitching Pennies*, *Crossword*, and *Monopoly*), stopping for further discussion when the participants had reflections or thoughts on what we were discussing.

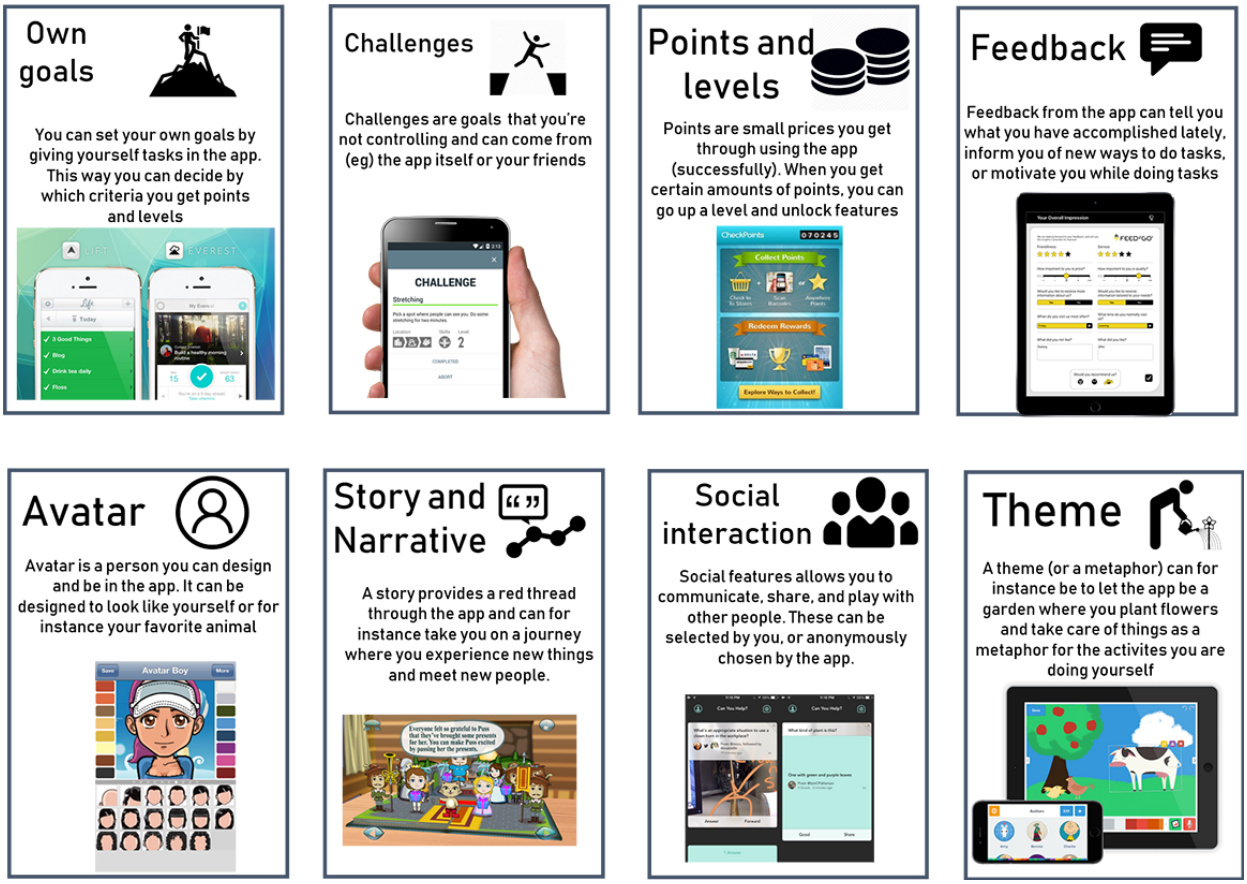
Sticky Notes Exercise

Next, we did a sticky notes activity where we asked the participants to note down on separate sticky notes the games they liked, why they liked them, and what feelings they evoked, and share this in group afterward. This task provided us with both the participants' overall preferences of games and was the first step in thinking of games as a mix of smaller design pieces that together create the user experience. After discussing what the participants reported, we asked for games they did not like, why they did not like it, and how they would improve it. This latter activity gave the participants a taste of designing and putting together new ideas. Ending the first half of the workshop, we summarized what we had accomplished thus far and presented 8 different categories of game elements (see Figure 1 and Multimedia Appendix 1), based on Hamari et al [7], using games suggested during the sticky notes exercises as examples.

Table 1. Participant information.

Site	Workshop	Number of participants (n)	Mean age (range), years	Diagnosis (n)	How experienced are you with smartphones and tablets?	Highest completed education (n)
A	1	7	36 (21-58)	Attention deficit hyperactivity disorder (4); bipolar disorder (2); bipolar disorder and eating disorders (1)	3.7	Secondary school (5); university (2)
A	2	4	37 (21-58)	Attention deficit hyperactivity disorder (2); bipolar disorder (2)	3.7	Secondary school (3); university (1)
B	1	5	19 (17-21)	Chronic fatigue syndrome (1); Crohn disease (1); depression (1); chronic intestinal pseudo-obstruction, gastroparesis, spinal cord injury (1); not reported (1)	3.8	Secondary school (5)
B	2	4	20 (17-21)	Crohn disease (1); cerebral palsy (1); chronic regional pain syndrome (1); chronic intestinal pseudo-obstruction, gastroparesis, spinal cord injury (1)	4	Secondary school (4)
C	1	7	48 (27-64)	Chronic fatigue syndrome (2); spinal cord injury (2); fibromyalgia and posttraumatic stress disorder (1); hearing impairment (1); multiple sclerosis (1)	3.3	Primary school (1), secondary school (5); university (1)
C	2	6	50 (32-64)	Chronic fatigue syndrome (2); spinal cord injury (2); hearing impairment (1); multiple sclerosis (1)	3.0	Secondary school (5); university (1)

Figure 1. Presentation of game elements from workshop.



Design Game 1

The second half of the workshop consisted of a 2-part design game. For the first part, we prepared a set of cards with (1) personas and design challenges and (2) game design elements (Figure 2). Personas are descriptive models or representations of unique users [48], and we created 3 personas that had generally known chronic illnesses with commonly known symptoms, challenges, and issues. The back of the persona cards featured a small design challenge specific for the persona. These challenges were based on tool functionality ideas identified in earlier work on the project [45]. All 3 personas are presented in full in Multimedia Appendix 2.

The game element cards, 8 in total, had the title of the element on the front, and a small descriptive icon and explanatory text with a few general examples of use on the back. We also provided a "wild card," which could be whatever design element or approach the participants chose, to lessen the chance of the participants running out of ideas for their task and promote creativity.

The participants were split into 2 smaller groups of 3 to 4 participants each, who worked independently of each other. Each group at random drew a card with a persona and a design challenge, 2 cards containing game elements, and got a game element wild card. The overall idea for this activity was to create an idea solving the challenge on the persona card by using the game elements cards. The facilitators were always available for discussion but did not partake in the groups' work. To structure their work, the groups were given a poster to write down their ideas on (see poster A in Figure 3). After working for 20 min, the groups presented their ideas, the facilitator asked a few reflecting questions, and there was a short plenary discussion on the different game elements' uses and ideas.

Figure 2. Cards from design game.

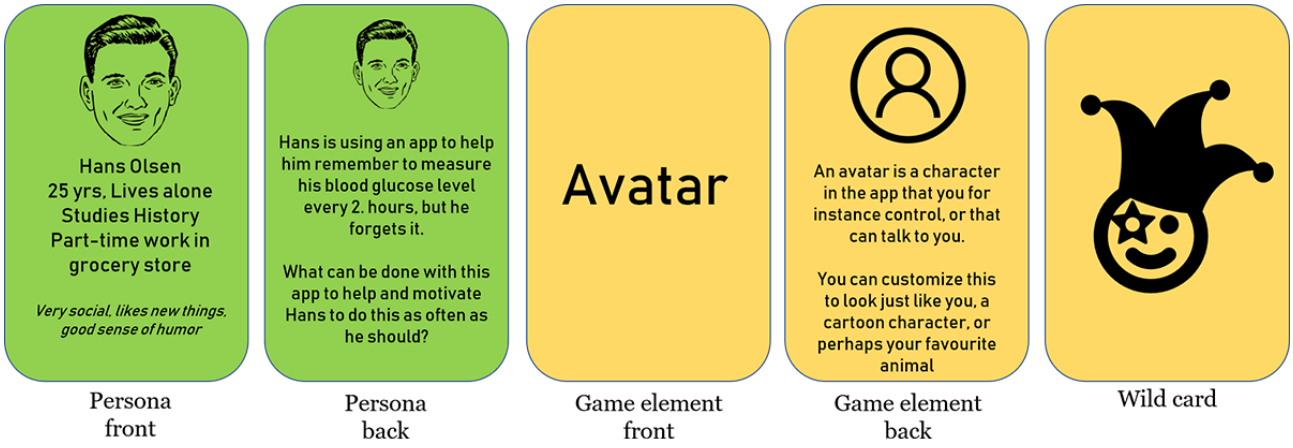
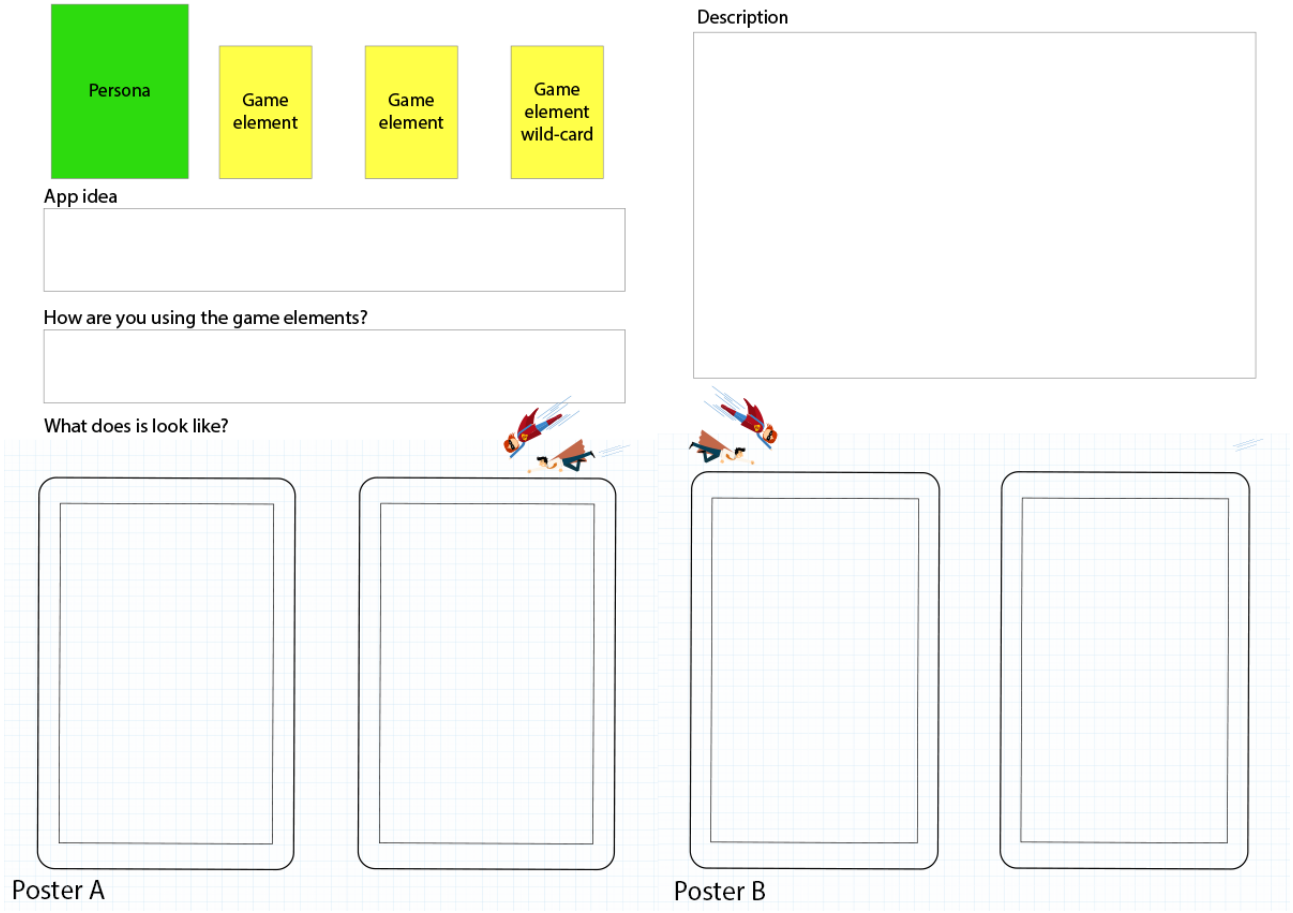


Figure 3. Posters for design proposals.



Design Game 2

The second design game followed the same outline as the previous. The groups kept their persona but received a new design challenge card (see Figure 4) with a larger and more complex design challenge. As the groups now were familiar with how this design game was played, we removed the constraints of the groups having to use the design elements they had drawn, and they were free to design whatever and however they wanted. To write down and present their ideas, we provided an additional poster (poster B in Figure 3) with more room for describing and drawing their ideas and proposals. The groups got 30 min to work and then presented what they had come up with. As before, the facilitator then led a short plenary discussion on the different game elements' uses and ideas.

Figure 4. Cards for design game 2.

End of Workshop 1

When the discussion had ended, the facilitator briefly summed up the workshop activities and thanked everyone for their participation. We gave the participants the gift cards and a small notebook and asked them to write down experiences of gameful designs they would have until the next workshop, where we would discuss these. Before leaving, we held a short plenary discussion for the participants to provide feedback on the workshop.

Planning and Design for Workshop 2

After the first round of workshop, the ideas and inputs from the participants were sorted and preliminarily analyzed. We then combined these design ideas and inputs with the previously identified and wanted functionality of our mHealth tool [45] and created a paper-based, low-fidelity prototype of our mHealth tool. This contained functionality for assessing your own strengths, setting goals, selecting strengths to help achieve your goals, and to collaborate with a friend. On the basis of the feedback and our own experiences, we also tried to make activities in the second workshop more concise to allow more time for group work. Finally, as it became clear that new people had to be recruited to the second workshop, we made the recap of the first workshop more detailed and comprehensive.

Workshop 2

The second workshop was held with the same user groups at the same settings, approximately a month after the first. The main theme for this workshop was the design of the tool supporting the discovery and use of personal strengths. Maintaining the same overall structure as the first workshop, the second was built upon the results from the first and added the concept of personal strengths in the same manner as gameful designs in the first workshop.

Introduction

We started with a recap of our project's aims, what gameful designs are, and our goal of designing the mHealth tool gamefully can make it more engaging to use. Thereafter, everyone had the opportunity to either present what they had written down in their notebooks or other thoughts and reflections they had since the last workshops regarding gameful designs. We then presented the concept of personal strengths and how basing care and self-management around your own strengths can improve quality of life and overall well-being. The participants then did a strengths identification exercise by selecting strengths items from a list of 30 personal strengths items that participants in previous studies have reported [43], such as "I am a social person" and "I like to try new things." The participants volunteered to present their strengths and stories of situations in which they had used these. This was followed by a discussion concerning the exercise and reflections on the process. These tasks were performed to help the participants better understand the concept of strengths, experience how the strengths-identification process would look and feel, and create an overall positive atmosphere by reminding the participants of their own strengths.

Redesign Activity

We introduced the participants to the paper prototype of the app (Figure 5) and asked them to redesign it to make it better suited

for finding and using more of their own personal strengths. The prototypes were printed on A4 size paper, clipped together to allow for easy reorganization, with ample room for notes and drawing. We also supplied *empty* wireframes for new drawings.

The participants worked for 20 min in groups and then presented their results. Thereafter, the facilitator led a short discussion, asking reflecting questions regarding the redesigned prototypes and the participants' implementation of the strengths concept in these.

Design Game

For the final task, we used cards with the personas from the previous workshop. This time they were slightly rewritten, removing the design challenges and instead listing 5 of their strengths. We also provided cards (Figure 6) presenting a context in which the user would use the app (at home, at work, at school, with friends, at the doctor, and engaging in a recreational activity).

Each group drew a persona and a context card. The task was to describe how their persona would use their modified app in that context. After working for approximately 30 min, the groups presented their solutions as use scenarios. The discussion then continued on how the participants themselves could use such an app in their own context.

Ending Workshop 2

After the discussion, we summed up both workshops and presented the project's future development plan. We then briefly discussed the participants' experiences of the workshops before we thanked everyone for their participation, gave them gift cards, and ended the workshop.

Data

The workshops were audio and video-recorded, totaling approximately 15 hours. Both the facilitator and the observer present took notes as well as photos during the workshops, and we collected all written materials created during the workshops. This provides us with 4 types of data (see Table 2). The recordings and written materials form the core data for our analysis, whereas the photos and notes add context and framing.

Figure 5. Paper prototypes for workshop 2.

Figure 6. Cards for design game in second workshop.

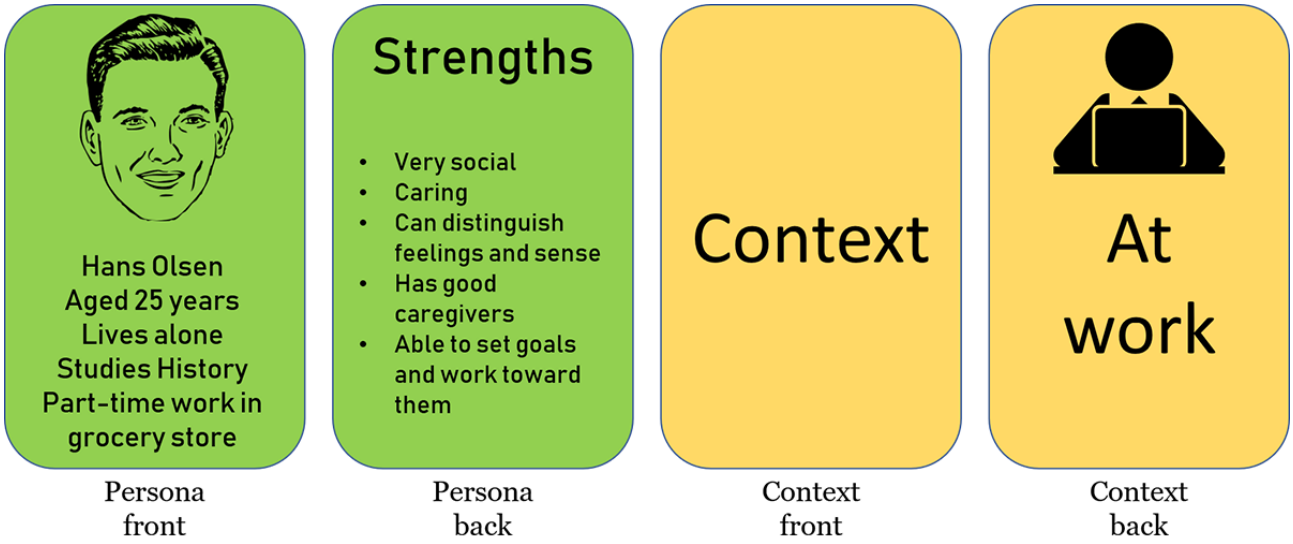


Table 2. Types of data.

Type	Description	Status
Audio/video recordings	Transcribed recordings of all 6 workshops, including individual groups	Core data
Written materials	Drawings, notes, sticky notes, and written ideas from the workshops	Core data
Photos	Photos taken during the workshops	Contextualizing data
Facilitators' notes	Our own notes, written down during and after the workshops	Contextualizing data

Analytical Approach

The data collected were analyzed thematically, guided by the 6 steps described by Braun et al [49]. The recordings were transcribed by the first author (SJ) and imported into QSR NVIVO 11 [50]. The first author made a first pass of coding, seeking inputs regarding use and experiences of gameful designs. New codes were generated as needed. To ensure that all data were coded with the same set of codes, the whole material was gone over a second time. When codes were related, overall categories were created, resulting in an hierarchy with 2 levels such as the category *Goals and Competitiveness*, which contains the 3 codes *Challenges from users or the app*, *Competition*, and *Setting goals/challenging yourself*. The codes and categories were discussed with the second author (JM), and inconsistencies or disagreements were discussed until an agreement was reached. The first author then made another pass to ensure that the entire corpus was coded from the updated codes and categories. The first and second author then together reviewed and agreed on a final set of categories and codes.

Reflections on Reliability

In this study, reliability issues are addressed by following several of the strategies suggested by Creswell [51]. First, the workshops are held at 3 different sites and are conducted using both individual and collaborative methods and activities. Second, in the presentation of the results, we present both decontextualized extracts and examples of our coding, as well as 2 examples of the participant's whole app concepts. Third, as the second workshop builds upon the output of the first, it also functions as a form of member checking. Despite this, although we follow several suggested strategies for reliability in qualitative research, it is important to not view the output from the workshops as generalizable but as products of the situated activities.

Results

Overview

The coded data cover aspects of design such as gamelike features, look, feel, and overall user experience and were separated into 6 overall categories during analysis. In addition to the coded data, this section also presents 2 complete design concepts for self-management mHealth apps that 2 of the participant groups created during the design activities. These 2 concepts provide a macro view of the participants' preferences, use, and combinations of different game elements.

Points, Progress, and Rewards

Points, progress, and rewards or combinations of these were mentioned by all the groups. The ideas included getting recognition from the app for finishing tasks or being able to acquire points to unlock new functionality:

> *To get points when you have done something positive is kind of the easiest. You know, to get some recognition when you have done well.* [Site C, WS1 male, 64, spinal cord injury]

Some of the participants suggested having different types of points that are aligned with the users' situation or context. For example, the user could obtain points for doing nothing or taking a break, as resting is important for many patients with chronic conditions such as fatigue. Some of the participants also

suggested granting users control over what forms of rewards that will be used:

> *Maybe you can decide for yourself what you want as a reward then?* [Site A WS1, woman 58, bipolar disorder]

Points were popular with the participants, even going as far as one group suggesting awarding 100 points a day for each small task completed in the app. However, rewarding points were also discussed with trepidation, as pursuing more points can be both stressful and addictive:

> *But, points, doesn't that stress you out when you should be relaxing with the app?* [Site C WS1, male 47, hearing impairment and tinnitus]

Exchanging points into rewards was an important topic of discussion. Some of the groups discussed linking rewards in the app to rewards in the real world, by having a gift shop at a hospital, where users can choose and exchange real-world rewards and presents for the points in the app (which is similar to how people donating blood in Norway are rewarded). Contrary to the idea of real-world rewards, several groups discussed having rewards in the virtual world, such as trophies:

> *If she wins something, it should be something she gets in the app, and not in the real world.* [Site C WS1, female 37, fibromyalgia, posttraumatic stress disorder]

Goals, Challenges, and Competition

Most of the groups thought that users should be able to set their own goals and break these down to more manageable subgoals. Some suggested entering a goal when starting to use the app and then creating subgoals to achieve it.

Other ideas included tailoring the goals and challenges based on the users' preferences situation. It was proposed that the app could do this automatically or by having someone working on the back end:

> *We think the app gives you challenges based on the goals you have set. Say you need to get better at feeling when your body needs rest, and then it [the app] will give you challenges that make you think about it.* [Site B WS2, woman 21, chronic intestinal pseudo-obstruction, gastroparesis, spinal cord injury]

Several ideas for increasing engagement revolved around the app enabling users to connect to and compete with others. However, participants also raised important concerns regarding the use of competitive elements in this context:

> *It can be tricky to let people compete or compare themselves against each other [referring to a prototype picture of two people climbing a mountain together]. This is fine if you're alone, but to see others being better than you or you being poorer/worse can be hard if you're lagging. There will always be someone at the back, and they may well be struggling the most.* [Site 1 WS2, woman 29, attention deficit hyperactivity disorder]

Avatars and Feedback

Avatars were also proposed by many of the participants. An avatar could, for example, function as a tutor or guide to give the user feedback on activities or show and explain how to do certain exercises. Presenting information through avatars was discussed as potentially making it more meaningful:

> *I believe in that about avatars, having someone talk to you. Maybe not your parents, but perhaps yourself or a friend or something...Because I think having someone talk to you is more effective than just reading it on screen.* [Site C WS1, male 64, spinal cord injury]

Another use of avatars was to have it represent the user herself and for instance, visualize the users' progression through the app or using the avatar to show how to do forms of exercises. With respect to the appearance of the avatar, it was suggested that users should be able to choose from a gallery of predefined avatars or create new ones on their own (for example, an avatar that can resemble the user or his/her favorite animal).

The participants also discussed the content of feedback users would receive, and the suggestions ranged from having the app delivering automatic predefined feedback to receiving it from peers who also use the app. When discussing feedback in general, many participants agreed that it should be mostly positive and productive, such as informing you of your accomplishments or providing help or guidance.

Concerning feedback in the form of notifications, the participants said they often view these as irritating and suggested that they should have a more meaningful purpose than to just remind users to use the app:

> *Not an app that gives lots of notifications like, you haven't done this and that, but more like, Good, you did this! But not reminding of the negative, so that you get more energy out of it.* [Site A WS1, woman 21, attention deficit hyperactivity disorder]

Social Features

Being able to share experiences, communicate, or collaborate with others are recognized by the participants as being powerful in terms of motivating and supporting an app's users. Ideas included the ability to connect with others by communicating through chat rooms or forums. One idea was to enable others to cheer you on in your progress by one-directional messages of support. Having a button to easily ask others in similar situations for help is suggested by several groups. Some participants also suggested that a user could have one specific partner to collaborate with closely while using the app.

Keeping a positive focus was also important in this context. One group was so concerned with this that they suggested that users only should be able to send content from a set of predefined texts, icons, or emoticons to ensure all communication is of a positive nature and that there are no negative comments:

> *Being able to push and motivate. But it should not be that you can send negative messages to each other, so it could be an alternative to only be able to send pre-written messages like good, heart, stars and stuff.*

> [Site B WS1, female 21, chronic intestinal pseudo-obstruction, gastroparesis, spinal cord injury]

Several participants also discussed issues surrounding privacy, such as enabling the user to decide who to share what with or whether to share things at all:

> *Cause', if it is private it is much easier to be totally honest. Sometimes you have strengths you might not want to share with anyone else.* [Site B WS1, female 17, did not report diagnosis]

Themes, Stories, and Narratives

Several groups suggested overall themes such as designing the app as a journey to exploring countries or continents and gradually unlocking new locations and activities. Other themes were a 400-meter sports-track with hurdles and other obstacles representing smaller goals or challenges. One group suggested climbing mountains as a theme that can both represent the users' goal and progression. Similarly, another group proposed having a theme of being in nature (which is a very common recreational activity in Norway), for example, moving through a forest in an unfolding story as a narrative:

> *It could be that you walk into a forest where something exciting is going to happen, maybe a story, and for each morning you go further in there.* [Site A WS1 woman 58, bipolar disorder]

Combining the idea of a narrative with the rewards in the app, another group suggested theming the app as a mystery story with rewards that unlock new chapters.

Regarding the personalization and fit of themes, one group also suggested that users should be able to choose between different themes after their own liking:

> *The app could be related to something you like. For instance, he likes working on cars, so perhaps instead of climbing mountains he gets a car in pieces he has to assemble.* [Site A WS2, male 40, attention deficit hyperactivity disorder]

Engaging Visuals, Sounds, and Texts

The participants' suggestions ranged from very specific needs (for example, sizes or look of specific buttons), to the general need and guiding principles for the app's "cool look." For instance, one group suggested that the home screen could be a "boasting wall," showing off the users' successes and strengths. Focusing on the positive, using images of times of success and happiness was mentioned by several groups as being powerful reminders and positive boosts during negative periods.

Other ideas included using a scrolling wheel instead of drop-down lists to adjust dates and times or having variation in the content and notification provided by the app. Interestingly, several groups also discussed the need for the app to be something new and innovative, not just copying features of other tools or services:

> *It's starting to be very similar to Facebook now, and it should not be that similar to other apps.* [Site A WS2, male 40, attention deficit hyperactivity disorder]

Although engaging design elements were proposed and described, the participants were very cautious about how these could affect the apps' usability and intuitiveness. For example, one group said the app should not have too many different buttons and menus, as this could be confusing:

> *It should be simple, and with easy and quick access. If there are people around 60 and 70…they might not understand everything, and may not find out how to use it.* [Site A WS2, male 21, attention deficit hyperactivity disorder]

Several groups also added to the immersion of the themes by suggesting sounds or music in the app that are topically proper, such as sounds of the forest or a windy mountain. Regarding textual content, the participants mostly agreed that the design elements should be presented as audio or video rather than just text. Some also discussed the different experiences that text and video can provide:

> *Yea, and if it's video then you get more the feeling that it's talking directly to you then if you're reading it.* [Site B WS2, male 17, cerebral palsy]

Two Design Concepts

This section presents summaries of 2 app concepts that were generated by the participants during the workshops. These serve both to highlight the complexity of the systems the participants created during the workshops and provide a macro level and more contextualized view on how the participants assembled the various microlevel, design elements.

Idea 1: A Journey Toward Mindfulness (Site C Workshop 1)

The goal of this app concept is to help the user perform mindfulness exercises. The app uses a journey to different parts of the world as a metaphor. The user can travel to different areas and countries with levels or stages that can be gradually unlocked. Each stage contains new exercises especially themed and tailored to the area. For instance, with India as the destination, the app uses Indian-themed symbols, sound effects, music, and provides mindfulness and yoga exercises based on the given region. When the user goes to another place, the content is themed for the new location. Elaborating on this during the second design game, the participants suggested adding a feature that lets a person from the users' personal network, such as a partner or a parent, provide support with encouraging messages through the journey. The app should also allow the user to add pictures of happy times and situations that can be used both as rewards and reminders, for instance in periods when one is feeling depressed. The participants also suggested a feature that enables the user to communicate with others in the same situation by sharing new places discovered on your journey in the app as well as documenting and sharing physical places that provide meaning, joy, or relaxation in their real life.

Figure 7. Participant drawing from workshop 1, and refined prototype wireframe for workshop 2.

Idea 2: Climbing Mountains of Challenges (Site B Workshop 1)

The other large design concept developed by the participants involved helping a user reach their goals. The apps' main metaphor is mountain climbing (see drawing in Figure 7). The app allows a user to design his or her own avatar and the movement of this avatar across the mountain serves as a visualization of the users' movement toward his or her goals. The avatar can also show and explain various tasks or exercises the users will encounter along their journey up the mountain. These mountains are modeled on real mountains, and their height is relative to the users' progression in the app. This app was also designed to provide the users with sounds of nature to add to the immersion. Users also have the option to communicate with others and add them to their app and climb together. The app would then visualize how both users are scaling the mountain in relation to each other.

Discussion

Participatory Methods

In this study, we organized a series of participatory design workshops with people living with chronic illnesses to jointly explore preferences, requirements, and ideas for gameful mHealth tools. The results of the study showed that engaging the participants with gamelike activities supported them to be collaborative, effective, and creative, especially by applying activities that set particular rules to their interaction (such as the rules of the game itself and the restriction of design elements). In addition, this approach provided the participants with a direction for their exploration of new ideas through, for example, the personas with their connected design challenges and the game elements cards. These findings are in line with previous studies that explore using gamelike participatory activities in design processes [33-35].

Using card-based design tools and activities are common to both design in general [39,52] and to game and gameful designs [17,53]. Even though it is important to not only focus on the

application of specific elements but also adapting them specifically to the target users and their context, this is often overlooked and not addressed properly in design processes [17]. In our study, we addressed this by the following: (1) having the persona card provide a clear context for both the use of the future tool and the user and (2) defining design challenges in the context of system functionality and requirements identified by users in earlier phases of the project. The game design element cards were also designed to be broad, leaving it to the participants to decide the specific interpretation and use of the elements. Thus, the participants were staying within the specific context and scope of our project while having room to freely ideate and be creative. Interestingly, it seems that the participants used the design cards more as what is termed *inspiration cards* [54], which often carry images, words, or short statements that are used as a point of departure for further discussion and exploration. Our participants used these cards as a starting point in the process of discovering possible design ideas and solutions. For instance, when some of the groups discussed using rewards in their idea, they not only discussed the reward as any generic reward but also how it can be used to (1) provide social activity by being a free dinner out with friends or (2) give positive boosts to the user by showing photos uploaded by either the user or their partner. As such, the participants used the game element cards freely and creatively and thus somewhat opposite to their often very specific and more prescriptive uses in design activities [17].

However, we also experienced that facilitating a process that not only supports both openness for the participants to be creative and innovate but also provides rules to keep their work within the boundaries of our project can be challenging. One example of how we addressed this issue is by limiting the number of game element cards the participants drew during the activities but also giving them a wild card that provided the freedom to use any gamelike approach they liked. In this manner, the participants were simultaneously provided with (1) a gamelike experience where they received different game elements by chance, (2) support in choosing the design concepts to use as a starting point, and (3) the possibility to freely explore features and elements other than those they had randomly drawn. The 3 different groups of participants in this study all had a different character and behavior, and although it is important to find a good balance of openness and rules during the planning of the design activities, some adjustments still had to be made on the fly during the workshops.

User involvement in the design process can play an integral part in *widening the design space* by contributing choices and ideas to the design project, stemming from their own imaginations of future uses of such tools [55]. The participants in our study came up with a great range of ideas and design proposals, showcasing a collective creativity that greatly adds to and extends that of the professional designers, developers, and researchers in our project team. Some of these ideas include having a gift shop in which you can exchange virtual points for real-world gifts, being able to add own photos that can be used as rewards and positive reminders, or the button you could push to easily get in contact with people in similar situations. However, as a wide range of ideas and suggestions were reported and discussed by the participants during the workshops, some of these are at times at odds with each other, such as the ideas of having competitions in the app and the wish to not visibly lose to someone else. Some ideas are also counter to evidence and design principles. For example, one group suggested awarding 100 points for each of the 8 completed small tasks during a day, for a total of 800 points as a score for completion. Although rewarding points are one of the more popular gameful design elements [6,7], it is also known that one does not engage users more by inflating the rewards by as suggested, giving 100 instead of a single point [17]. Therefore, even though involving users is both important and valuable, one must still make sure design decisions are made in accordance with relevant literature and evidence concerning both the design and content of the tool that is being made.

Overall, we can conclude from the vast variety of user inputs that the workshops were successful in generating new and creative concepts and ideas for mHealth tools. It served as a vehicle for the participants to gain new knowledge from this domain and communicate their requirements and needs. However, giving the participants the freedom to interpret their own tasks also allowed them to veer in directions that can be unproductive (as with the example of awarding 100 points at a time) or impossible to implement. At the same time, it is hard to correct participants when they veer outside our topic without seeming critical or negative, and in these few cases, we mostly let them continue. Despite this, even though such diversions may be unproductive in terms of creating design ideas, they still expand the knowledge base and overall output of the design activities. One thing that did not work as intended was the notebooks given to the participants after workshop 1. Many had misplaced or simply forgotten about these between the 2 workshops, and in future studies, we will consider either using text messages or social media to remind the participants of such tasks. The participants and the facilitators alike found the workshops to be both productive and enjoyable. In fact, when getting feedback at the end of the first workshop, all 3 groups wanted to spend more time on the next workshop.

Design Ideas and Requirements

As presented in the Results section, the workshops yielded a range of ideas and requirements for designing mHealth tools gamefully. For example, the use of metaphors, which is a well-known valuable design approach to increase motivation and use of mHealth tools [25,56], was frequently proposed by participants in the study. As exemplified by the 2 design concepts, “Journey towards mindfulness” and “Climbing mountains of challenge,” the participants confirmed that the use of an overarching theme or metaphor can be a suitable approach to designing mHealth tools. However, we also noticed that many of the proposed metaphors are culturally and context specific, which limits their overall generalizability. For example, hiking mountains and being outdoors in nature is a popular recreational activity in Norway but possibly not equally appealing for people living without easy access to nature. Similarly, chapters of a story as rewards could be engaging only for users interested in the story. Therefore, we can conclude that although metaphors can be a powerful and engaging design element, they need to be fitting to the target users, and one way

to ensure this is, as also suggested, to allow users to choose between several themes or styles of metaphors.

Social functionalities are commonly employed in mHealth tools to provide collaboration or communication between users [3,6,57]. Such features were also often suggested by the participants as a powerful way for both getting support from, and being motivated through interaction with others. However, it is important for designers to be careful about how they implement such features for sensitive groups, such as people living with chronic illnesses [13,58,59]. This was discussed on multiple occasions by the participants, and they voiced both opportunities and concerns. On the beneficial side, having others to communicate with can be a great source of inspiration and support during hard times. Being part of a group or community with others in similar situations or with the same diagnosis was also mentioned as a good way for obtaining advice. This is in line with, amongst others, findings from research on the *Patients like me* network [60] that showed users advising and supporting others in similar situations based on their own personal experiences. On the other hand, being able to compete or compare one's own progression with other users of the app was also mentioned as being detrimental to motivation and joy in general. This is similar to what the study by Chen Y et al [61] reported, which stated that competition between people with different abilities and performance could be experienced as demotivating. During the workshops, the participants also often touched upon the changing shape or mood people living with chronic illness experience and highlighted the importance of taking this into consideration when designing features for mHealth tools that, for instance, offer social comparison or competition with others.

As mentioned, awarding points and rewards are among the most commonly used gameful design elements [6,7], and were also one of the more popular design features suggested by the participants. Besides awarding points in the app, most groups also discussed the possibility of rewards outside of the app. This is in line with Nicholson [23] who argued that rewarding not only virtually but also with something tangible can be experienced as more meaningful by the users. In addition, approaches such as pursuing rewards have also been reported as unfit in some of the health-related contexts such as mental health and mindfulness [13]. This is also reported by the participants, who often voice concerns regarding using designs that rely heavily on collecting points and trophies or rewarding use with streaks. Thus, we can conclude that combining both external rewards such as points in the app with more personal and intrinsic rewards (such as a real-world gift of your choice or going to dinner with your friends) can be a promising approach for providing rewards as part of gameful designs in this context.

Previous research has shown that personalization and allowing users to customize their own gameful tools might be a way of alleviating the issues of one-size-fits-all designs [17]. Personalization of the mHealth services by, for example, tailoring messages or allowing the users to customize the appearance or behavior of the service, can be an important mediator for user satisfaction and enjoyment of services [25]. In addition, personalization can broaden the reach of metaphors, social features, competitive elements, and rewards by having the users adapt these to their own preferences [25], something also suggested by the participants. In terms of designing for positive and more engaging user experiences, the participants proposed a range of relevant ideas and solutions, such as adding one's own music, designing one's own avatar, or setting one's own goals.

Every Stone Is a Keystone

For gameful designs in general, the results gained from this series of workshops highlight the complexity of both designing and experiencing gameful tools. We saw that the participants used the different design elements in highly interconnected ways and sometimes had them build on each other to form overall concepts or ideas for a tool. One example is the idea of climbing mountains with a friend. The overall idea was providing a sense of a competition (2 users compete for reaching the top of the mountain), but it is also designed as a social feature (you compete with someone) and a way of monitoring progress (climbing the mountain visualizes both users' progress). This highlights how the experience of gameful designs is not a product of individual elements such as trophies or avatar but rather a product of the interaction with the gameful tool or service as a whole—something that is also often discussed in existing literature [9,62,63].

For both this and future work on co-designing gameful tools and apps, it is thus important to be considerate when combining pieces from different proposals or ideas coming from co-designers, as when you combine pieces from different ideas, you also create new and different overall user experience. Including end users throughout the design process and being open to their needs, requirements, and inputs are therefore important for ensuring that users find the final tools both meaningful and valuable [25].

Strengths and Limitations

Both the methods used and the findings from this study can serve as a backing for future work and research on creating gameful designs for and with people with chronic illnesses. However, due to the explorative nature of this study, any generalizations as to what gameful approaches or designs people living with chronic illnesses enjoy or want is neither possible nor intended. Yet, we can conclude that the chosen methods worked well with 3 different groups and may be applicable to others as well. Furthermore, the heterogeneity of the participant group may be considered a limitation as the size of any participant subgroup (be it group by age, gender, or illness) is small. Nonetheless, this also allowed us to get feedback and input from many different viewpoints. It should be noted that the participants were all recruited from active users of the hospital youth council or education centers and most were also active in patient organizations. As such, this participant group is possibly somewhat biased in that they are resourceful and able to manage their life well with a chronic illness and not necessarily representative of the overall population of chronic patients. However, using *empowered* users is common in this phase of design projects, as they typically have more experience in addressing the existing problems and may have more reflective thoughts about their situation. Moreover, in a study

like this, recruiting a convenience sample is often necessary to find participants willing to take part. Finally, the gender balance among the participants is skewed with 14 women and 8 men, which may have influenced the ideas from the workshops.

Conclusions

In this study, we used participatory design methods to jointly explore, together with people living with chronic illnesses, their preferences, requirements, and ideas for designing gameful and engaging mHealth tools. Through gamelike design activities, the participants were both engaged, creative, and voiced a wide range of ideas and requirements. Much of the reported input and ideas are in line with previous research and provide important contextualization and nuance to these design choices from the users' perspective, although we cannot generalize from the findings. As such, both the participants' needs and requirements as well as the applied methods and activities add to a growing body of literature in the field of designing mHealth and eHealth tools in engaging ways by implementing gameful design features.

Both the methods used in and the results from this study could be used as a starting point for future studies exploring requirements of gameful designs in depth with other user groups, and we invite others to both further develop, adapt, and build on these activities for their contexts.

Acknowledgments

Funding for this project is provided by the Norwegian Research Council (Grant #248026). The authors would like to thank their larger project team; Kurt Stange, Shirley M Moore, Luis Fernandez Luque, Hein De Vries, and Christina Bode. The authors would also like to thank all staff and participants at the learning and mastery centers at the Hospital of Southern Norway and Norland Hospital as well as the Youth Council at Akershus University Hospital for their time, creativity, and effort.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Game elements.

[PDF File (Adobe PDF File), 11KB - mhealth_v6i12e11579_app1.pdf]

Multimedia Appendix 2

Personas.

[PDF File (Adobe PDF File), 15KB - mhealth_v6i12e11579_app2.pdf]

References

1. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
2. Eland-de KP, van Os-Medendorp H, Vergouwe-Meijer A, Bruijnzeel-Koomen C, Ros W. A systematic review of the effects of e-health on chronically ill patients. J Clin Nurs 2011 Nov;20(21-22):2997-3010. [doi: 10.1111/j.1365-2702.2011.03743.x] [Medline: 21707807]
3. 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;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
4. Kelders SM. Involvement as a Working Mechanism for Persuasive Technology. Chicago, USA: Springer, Cham; 2015 Presented at: International Conference on Persuasive Technology; 3-5 June 2015; Chicago, USA p. 3-14 URL: https://link.springer.com/chapter/10.1007%2F978-3-319-20306-5_1 [doi: 10.1007/978-3-319-20306-5_1]
5. Triberti S, Kelders S, Gaggioli A. User Engagement. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development. London & New York: Routledge; May 15, 2018:271-289.
6. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002]
7. Hamari J, Koivisto J, Sarsa H. Does gamification work? - A literature review of empirical studies on gamification. In: Proceedings of the 47th Annual Hawaii International Conference on System Sciences.: IEEE; 2014 Mar 10 Presented at: The 47th Annual Hawaii International Conference on System Sciences; 6-9 Jan. 2014; Waikoloa, HI, USA p. 3025-3034. [doi: 10.1109/HICSS.2014.377]
8. 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. New

York, NY, USA: ACM; 2011 Presented at: MindTrek '11; September 28 - 30, 2011; Tampere, Finland p. 9-15 URL: https://dl.acm.org/citation.cfm?id=2181040 [doi: 10.1145/2181037.2181040]
9. Huotari K, Hamari J. A definition for gamification: anchoring gamification in the service marketing literature. Electron Mark 2016 Jan 15;27(1):21-31 [FREE Full text] [doi: 10.1007/s12525-015-0212-z]
10. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiyagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Oct 04;6(10):e012447 [FREE Full text] [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
11. Tondello GF, Mora A, Nacke LE. Elements of Gameful Design Emerging from User Preferences. In: Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY 17. New York, USA: ACM Press; 2017 Presented at: The Annual Symposium on Computer-Human Interaction in Play - CHI PLAY 17; October 15 - 18, 2017; Amsterdam, The Netherlands URL: https://dl.acm.org/citation.cfm?id=3116627 [doi: 10.1145/3116595.3116627]
12. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game On? Smoking cessation through the gamification of mHealth: a longitudinal qualitative study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]
13. Ahtinen A, Mattila E, Välkkynen P, Kaipainen K, Vanhala T, Ermes M, et al. Mobile mental wellness training for stress management: feasibility and design implications based on a one-month field study. JMIR Mhealth Uhealth 2013;1(2):e11 [FREE Full text] [doi: 10.2196/mhealth.2596] [Medline: 25100683]
14. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
15. LeGrand S, Muessig KE, McNulty T, Soni K, Knudtson K, Lemann A, et al. Epic allies: development of a gaming app to improve antiretroviral therapy adherence among young HIV-positive men who have sex with men. JMIR Serious Games 2016 May 13;4(1):e6 [FREE Full text] [doi: 10.2196/games.5687] [Medline: 27178752]
16. Wilson AS, McDonagh JE. A gamification model to encourage positive healthcare behaviours in young people with long term conditions. EAI Endorsed Transactions on Game-Based Learning 2014 May 22;1(2):1-10 [FREE Full text] [doi: 10.4108/sg.1.2.e3]
17. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Hum-Comput Interact 2015 May 15;30(3-4):294-335 [FREE Full text] [doi: 10.1080/07370024.2014.993471]
18. Mora A, Riera D, González C, Arnedo-Moreno J. Gamification: a systematic review of design frameworks. J Comput High Educ 2017 May 23;29(3):516-548 [FREE Full text] [doi: 10.1007/s12528-017-9150-4]
19. Charles D, Mcdonough S. A participatory design framework for the gamification of rehabilitation systems. In: Proceedings of the 10th Intl Conf Disability, Virtual Reality & Associated Technologies.: The University of Reading; 2014 Presented at: 10th Intl Conf Disability, Virtual Reality & Associated Technologies; 2–4 Sept. 2014; Gothenburg, Sweden p. 293-296 URL: http://centaur.reading.ac.uk/37397/
20. AlMarshedi A, Wills G, Ranchhod A. Gamifying self-management of chronic illnesses: a mixed-methods study. JMIR Serious Games 2016 Sep 09;4(2):e14 [FREE Full text] [doi: 10.2196/games.5943] [Medline: 27612632]
21. AlMarshedi A, Wills G, Ranchhod A. Guidelines for the gamification of self-management of chronic illnesses: multimethod study. JMIR Serious Games 2017 May 12;5(2):e12 [FREE Full text] [doi: 10.2196/games.7472] [Medline: 28500018]
22. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Dec;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
23. Nicholson S. A User-Centered Theoretical Framework for Meaningful Gamification. : WITC Press; 2012 Presented at: Games+Learning+Society 8.0; 8-10 July 2012; MADISON, WI, USA.
24. Morschheuser B, Werder K, Hamari J, Abe J. How to gamify? A method for designing gamification. In: Proceedings of the 50th Annual Hawaii International Conference on System Sciences (HICSS). 2017 Presented at: The 50th Annual Hawaii International Conference on System Sciences; January 4-7, 2017; Hawaii, USA p. 1298-1307. [doi: 10.24251/HICSS.2017.155]
25. Ludden GD, Van Rompay TJ, Kelders SM, Van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. J Med Internet Res 2015;17(7):e172 [FREE Full text] [doi: 10.2196/jmir.4201] [Medline: 26163456]
26. 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]
27. Abras C, Maloney-Krichmar D, Preece J. User-Centered Design. In: Bainbridge WS, editor. Berkshire Encyclopedia of Human-Computer Interaction. Great Barrington, MA, USA: Berkshire Publishing Group; 2004:763-768.
28. Sanders L. An evolving map of design practice and design research. Interactions 2008 Nov 01;15(6):13-17 [FREE Full text] [doi: 10.1145/1409040.1409043]
29. Robertson T, Simonsen J. Participatory Design: an Introduction. In: Simonsen J, Robertson T, editors. Routledge international handbook of participatory design. New York, NY, USA: Routledge; 2013:1-17.

30. Robertson T, Wagner I. Ethics: Engagement, representation and politics-in-action. In: Simonsen J, Robertson T, editors. Routledge international handbook of participatory design. New York, NY, USA: Routledge; 2013:64-85.
31. Kensing F, Greenbaum J. Heritage: Having a Say. In: Simonsen J, Robertson T, editors. Routledge international handbook of participatory design. New York, NY, USA: Routledge; 2013:21-36.
32. Brandt E, Binder T, Sanders EB. Tools and techniques: Ways to engage telling, making and enacting. In: Simonsen J, Robertson T, editors. Routledge International Handbook of Participatory Design. New York, NY, USA: Routledge; 2013:145-181.
33. Brandt E. Designing Exploratory Design Games: A Framework for Participation in Participatory Design? In: Proceedings of the ninth Participatory Design Conference 2006. New York, NY, USA: ACM; 2006 Presented at: Participatory Design Conference 2006; July 31- August 5, 2006; Trento, Italy p. 57-66. [doi: 10.1145/1147261.1147271]
34. Nicholas M, Hagen P, Rahilly K, Swainston N. Using participatory design methods to engage the uninterested. In: Proceedings of the 12th Participatory Design Conference: Exploratory Papers, Workshop Descriptions, Industry Cases - Volume 2. New York, NY, USA: ACM; 2012 Presented at: The 12th Participatory Design Conference; August 12 - 16, 2012; Roskilde, Denmark p. 121-124. [doi: 10.1145/2348144.2348183]
35. Brandt E, Messeter J. Facilitating collaboration through design games. In: PDC 04 Proceedings of the eighth conference on Participatory design: Artful integration: interweaving media, materials and practices - Volume 1. New york, NY, USA: ACM; 2004 Presented at: PDC 04 The eighth conference on Participatory design; July 27 - 31, 2004; Toronto, Ontario, Canada p. 121-131. [doi: 10.1145/1011870.1011885]
36. Brandt E, Messeter J, Binder T. Formatting design dialogues - games and participation. CoDesign 2008;4(1):51-64. [doi: 10.1080/15710880801905724]
37. Kayali F, Peters K, Kuczwara J, Reithofer A, Martinek D, Wölfle R, et al. Participatory Game Design for the INTERACCT Serious Game for Health. In: First Joint International Conference, JCSG 2015, Huddersfield, UK, June 3-4, 2015, Proceedings.: Springer, Cham; 2015 Presented at: First Joint International Conference on Serious Games, JCSG 2015; June 3-4 2015; Huddersfield, UK p. 13-25. [doi: 10.1007/978-3-319-19126-3_2]
38. Ruland CM, Starren J, Vatne TM. Participatory design with children in the development of a support system for patient-centered care in pediatric oncology. J Biomed Inform 2008 Aug;41(4):624-635 [FREE Full text] [doi: 10.1016/j.jbi.2007.10.004] [Medline: 18082468]
39. Sommervold MM, van der Velden M. Transition cards: designing a method with and for young patients. IADIS International Journal on Computer Science and Information System 2015;10(2):79-94.
40. Walker C. Recognising the changing boundaries of illness in defining terms of chronic illness: a prelude to understanding the changing needs of people with chronic illness. Aust Health Rev 2001;24(2):207-214. [Medline: 11496465]
41. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005 Aug;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
42. Park N, Peterson C, Seligman ME. Strengths of Character and Well-Being. J Soc Clin Psychol 2004;23(5):603-619. [doi: 10.1521/jscp.23.5.603.50748]
43. Kristjansdottir OB, Stenberg U, Mirkovic J, Krogseth T, Ljoså TM, Stange KC, et al. Personal strengths reported by people with chronic illness: a qualitative study. Health Expect 2018 Aug;21(4):787-795 [FREE Full text] [doi: 10.1111/hex.12674] [Medline: 29478260]
44. Harzer C. The Eudaimonics of Human Strengths: The Relations Between Character Strengths and Well-Being. In: Vittersø J, editor. Handbook of Eudaimonic Well-Being. International Handbooks of Quality-of-Life: Springer, Cham; 2016:307-322.
45. Mirkovic J, Jessen S, Kristjánsdóttir Ó, Krogseth T, Koricho AT, Ruland CM. Developing technology to mobilize personal strengths in people with chronic illness: positive codesign approach. JMIR Formativ Res 2018 Jun 05;2(1):e10774. [doi: 10.2196/10774]
46. Linn MC, Eylon BS. Science Learning and Instruction: Taking Advantage Of Technology To Promote Knowledge Integration. New York, NY, USA: Routledge; 2011.
47. de Jong T. Cognitive load theory, educational research, and instructional design: some food for thought. Instructional Science 2009 Aug 27;38(2):105-134. [doi: 10.1007/s11251-009-9110-0]
48. Cooper A, Reimann R, Cronin D, Noessel C. About Face: The Essentials of Interaction Design. 4th ed. Indianapolis, IN, USA: John Wiley & Sons, Inc; 2014.
49. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
50. NVivo 11 qualitative data analysis software. Melbourne, Australia: QSR International; 2015. URL: https://www.qsrinternational.com/nvivo/home [WebCite Cache ID 74EnHdyFJ]
51. Creswell JW. Qualitative Inquiry & Research Design: Choosing Among Five Approaches. 3rd ed. Thousand Oaks, CA: SAGE Publications Ltd; 2013.
52. Wölfel C, Merritt T. Method Card Design Dimensions: A Survey of Card-Based Design Tools. : Springer; 2013 Presented at: 14th International Conference on Human-Computer Interaction (INTERACT); 2-6 September 2013; Cape Town, South Africa p. 479-486. [doi: 10.1007/978-3-642-40483-2_34]

53. Jiménez AS. Game On Lab!. Game On! Toolkit Internet URL: http://www.gameonlab.com/toolkit/ [accessed 2018-07-12] [WebCite Cache ID 70r4mc03b]
54. Halskov K, Dalsgård P. Inspiration card workshops. In: DIS '06 Proceedings of the 6th conference on Designing Interactive systems. New York, NY, USA: ACM; 2006 Presented at: DIS 06 The 6th conference on Designing Interactive systems; 26-28 June 2006; University Park, PA, USA p. 2-11. [doi: 10.1145/1142405.1142409]
55. Bratteteig T, Wagner I. Design decisions and the sharing of power in PD. In: PDC '14 Proceedings of the 13th Participatory Design Conference: Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium papers, and Keynote abstracts - Volume 2. New York, NY, USA: ACM; 2014 Presented at: The 13th Participatory Design Conference on Participatory Design; 6-10 October 2014; Windhoek, Namibia p. 29-32. [doi: 10.1145/2662155.2662192]
56. Ludden GDS, Kelders SM, Snippert BHJ. This is your life! The design of a positive psychology intervention using metaphor to motivate. In: Persuasive Technology. PERSUASIVE 2014. Lecture Notes in Computer Science, vol 8462.: Springer, Cham; 2014 Presented at: 9th International Conference on Persuasive Technology (PERSUASIVE); 21-23 May 2014; Padua, Italy p. 179-190. [doi: 10.1007/978-3-319-07127-5_16]
57. Vorderstrasse A, Lewinski A, Melkus GD, Johnson C. Social support for diabetes self-management via eHealth interventions. Curr Diab Rep 2016 Dec;16(7):56. [Medline: 27155606]
58. Culén AL, van der Velden M. The digital life of vulnerable users: designing with children, patients, and elderly. In: 4th Scandinavian Conference on Information Systems, SCIS 2013, Oslo, Norway, August 11-14, 2013. Proceedings. The digital life of vulnerable users: Springer; 2013 Presented at: SCIS 2013: Scandinavian Conference on Information Systems; 11-14 August 2013; Oslo, Norway p. 161-176. [doi: 10.1007/978-3-642-32270-9]
59. Giunti G, Kool J, Rivera RO, Dorronzoro ZE. Exploring the specific needs of persons with multiple sclerosis for mHealth solutions for physical activity: mixed-methods study. JMIR Mhealth Uhealth 2018 Feb 09;6(2):e37 [FREE Full text] [doi: 10.2196/mhealth.8996] [Medline: 29426814]
60. Frost JH, Massagli MP. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. J Med Internet Res 2008;10(3):e15 [FREE Full text] [doi: 10.2196/jmir.1053] [Medline: 18504244]
61. Chen Y, Pu P. HealthyTogether: exploring social incentives for mobile fitness applications. In: Proceedings of the Second International Symposium of Chinese CHI on - Chinese CHI '14. New York, NY, USA: ACM; 2014 Presented at: Chinese CHI '14; 26-27 april 2014; Toronto, ON, Canada. [doi: 10.1145/2592235.2592240]
62. Hamari J. Aalto University. Aalto, Finland: Aalto University; 2015. Gamification - Motivations & Effects URL: https://aaltodoc.aalto.fi/handle/123456789/15037 [accessed 2018-11-27] [WebCite Cache ID 74EnnsYfx]
63. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]

Abbreviations

eHealth: electronic health
mHealth: mobile health

Edited by G Eysenbach; submitted 13.07.18; peer-reviewed by O Rivera, L Hightow-Weidman, A Lewinski, E Edwards, K Usop; comments to author 05.08.18; revised version received 08.10.18; accepted 29.10.18; published 14.12.18.

Please cite as:
Jessen S, Mirkovic J, Ruland CM
Creating Gameful Design in mHealth: A Participatory Co-Design Approach
JMIR Mhealth Uhealth 2018;6(12):e11579
URL: http://mhealth.jmir.org/2018/12/e11579/
doi:10.2196/11579
PMID:30552080

Original Paper

A Mobile App to Provide Evidence-Based Information About Crystal Methamphetamine (Ice) to the Community (Cracks in the Ice): Co-Design and Beta Testing

Louise Birrell[1], BSocSc, BPsych (Hons), PhD; Hannah Deen[1], BPsych (Hons); Katrina Elizabeth Champion[1,2], BAPsych (Hons), BHealth, PhD; Nicola C Newton[1], BPsych (Hons), PhD; Lexine A Stapinski[1], BPsych (Hons), MClinPsych, PhD; Frances Kay-Lambkin[1,3], BScPsych (Hons), PhD; Maree Teesson[1], BScPsych (Hons), PhD; Cath Chapman[1], BA (Hons), PhD

[1]National Health and Medical Research Council Centre of Research Excellence in Mental Health and Substance Use, National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia

[2]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Priority Research Centre for Brain and Mental Health, The University of Newcastle, Newcastle, Australia

Corresponding Author:
Hannah Deen, BPsych (Hons)
National Health and Medical Research Council Centre of Research Excellence in Mental Health and Substance Use
National Drug and Alcohol Research Centre
University of New South Wales
22 – 32 King Street
Randwick
Sydney, 2031
Australia
Phone: 61 2 8936 1217
Email: h.deen@unsw.edu.au

Abstract

Background: Despite evidence of increasing harms and community concern related to the drug crystal methamphetamine ("ice"), there is a lack of easily accessible, evidence-based information for community members affected by its use, and to date, no evidence-based mobile apps have specifically focused on crystal methamphetamine.

Objective: This study aims to describe the co-design and beta testing of a mobile app to provide evidence-based, up-to-date information about crystal methamphetamine to the general community.

Methods: A mobile app about crystal methamphetamine was developed in 2017. The development process involved multiple stakeholders (n=12), including technology and drug and alcohol experts, researchers, app developers, a consumer expert with lived experience, and community members. Beta testing was conducted with Australian general community members (n=34), largely recruited by the Web through Facebook advertising. Participants were invited to use a beta version of the app and provide feedback about the content, visual appeal, usability, engagement, features, and functions. In addition, participants were asked about their perceptions of the app's influence on awareness, understanding, and help-seeking behavior related to crystal methamphetamine, and about their knowledge about crystal methamphetamine before and after using the app.

Results: The vast majority of participants reported the app was likely to increase awareness and understanding and encourage help-seeking. The app received positive ratings overall and was well received. Specifically, participants responded positively to the high-quality information provided, usability, and visual appeal. Areas suggested for improvement included reducing the amount of text, increasing engagement, removing a profile picture, and improving navigation through the addition of a "back" button. Suggested improvements were incorporated prior to the app's public release. App use was associated with an increase in perceived knowledge about crystal methamphetamine; however, this result was not statistically significant.

Conclusions: The *Cracks in the Ice* mobile app provides evidence-based information about the drug crystal methamphetamine for the general community. The app is regularly updated, available via the Web and offline, and was developed in collaboration with experts and end users. Initial results indicate that it is easy to use and acceptable to the target group.

(JMIR Mhealth Uhealth 2018;6(12):e11107) doi:10.2196/11107

KEYWORDS

internet; methamphetamine; mobile apps; mobile phone; substance-related disorder

Introduction

In recent years, there has been widespread attention and concern across the globe as rates of methamphetamine production, consumption, and related-harms rise [1]. The crystalline form of methamphetamine, also known as crystal methamphetamine or "ice," is typically the strongest and purest form of the stimulant drug. Crystal methamphetamine is now the main form of methamphetamine consumed and the number one drug of concern in Australia [2]. While population prevalence rates have remained relatively low and stable, with 6.3% of Australians aged over 14 years ever reporting using any form of methamphetamine [2], there is evidence that harms related to the use of crystal methamphetamine are increasing [3,4]. Data also indicate that rates of use in rural and regional areas of Australia are higher than that in metropolitan areas [5].

In response to the increasing harms and community concern about crystal methamphetamine, the Web-based *Cracks in the Ice Community Toolkit* [6] was developed and launched in April 2017. *Cracks in the Ice* is a freely available website funded by the Australian Government Department of Health to provide trusted, evidence-based, and up-to-date information about crystal methamphetamine for the Australian community. The website was developed in collaboration with the community, leading experts in the field, and consumer experts, over an 18-month period from 2015 to 2016. The development was a broad-reaching and iterative process (for further details about the development process and beta testing see Ref. [7]). Its target audience includes people who use crystal methamphetamine, their friends and family, health care professionals, schools, and general community members with an interest or concern about the drug. The website includes resources that were developed by the research team in relation to the most up-to-date evidence and external resources. Prior to inclusion on the website, external resources (including fact sheets, guidelines, and Web-based programs) were independently reviewed by the *Cracks in the Ice* project team. Resources were assessed for eligibility for inclusion using an adapted version of the National Health and Medical Research Council (NHMRC) Body of Evidence Matrix (2009; see Multimedia Appendix 1) against the following criteria: evidence base, impact and utility, generalizability, applicability (applicable to an Australian context), recent (resource updated in the past 10 years), and duplication. Textbox 1 summarizes the website content. The website is regularly monitored, and website traffic tracked through Google Analytics. To ensure the information and resources remain up-to-date and include the latest evidence, the website content is reviewed on a regular basis with a systematic review of all content conducted once per year. Since launching in April 2017, the website has reached >79,000 unique users (as of May 2018). While the website usage has shown steady growth since launch, the website was primarily designed to be viewed on a desktop or laptop computer. The *Cracks in the Ice* app aims to bring together the best available evidence about crystal methamphetamine and improve access for the community to accurate information, including information about the effects of ice, where or how to seek help, and relevant support services. The provision of accurate information is an important part of community prevention strategies. While *Cracks in the Ice* is not a treatment intervention, it aims to promote help seeking by providing up-to-date, accurate information about treatment options, service contact details, and conversation starters. Furthermore, community consultation during the development of the Web-based toolkit indicated that Australian community members were seeking evidenced-based information about ice [7]. However, evidence indicates that the quality of information currently available on this topic in an app-based format is poor, and no existing apps focusing on crystal methamphetamine have undergone evaluation [8].

Recent data indicate that people are increasingly using mobile devices to access the internet, with mobile devices now the most frequently used device to access the internet in Australia [9]. Furthermore, consultation with community members, as part of the development process of the *Cracks in the Ice* website, indicated that nearly two-thirds (287/451, 63.6%) of participants said that they would use a mobile device to access an information website about crystal methamphetamine.

Over the past decade, there has been a proliferation of smartphone device ownership, particularly in high-income countries. The smartphone adaption in the United Kingdom rose from 52% of the population in 2012 to 85% in 2017 [10]. Australians have some of the highest smartphone ownership in the world, with 88% of the Australian population owning a smartphone in 2017, up from 84% in 2016 [11]. Mobile apps can extend the reach of public health information and offer offline capabilities, improving access for rural and regional communities, where access to the internet may be unreliable.

A recent review by our team of mobile apps containing information about methamphetamines, including crystal methamphetamine, identified a clear shortage of high-quality and engaging apps providing educational information [8]. To address this gap, a companion *Cracks in the Ice* mobile app was developed. The target audience and anticipated end users for the app mirrored that of the *Cracks in the Ice* website, including people who use crystal methamphetamine, their friends and family, health care professionals, schools, and general community members with an interest in, or concern about, the drug.

Textbox 1. Key content areas of the Cracks in the Ice website and mobile app.

Get the facts about ice

- What is ice
- How many people use ice
- What are the laws about ice

Staying safe

- When and where to get help (key support services in Australia)
- How to support a loved one
- Protecting yourself and others
- Support for Aboriginal and Torres Strait Islanders

What are the effects of ice?

- How ice works in the brain and body
- The mental health effects of ice
- Using ice with other drugs

Tailored resources for specific groups

- Community groups
- Families and friends of people who use ice
- Schools
- Health professionals

This paper describes the development and beta testing of the *Cracks in the Ice* mobile app, the first of its kind to provide evidence-based information about crystal methamphetamine in a mobile app format. The app aims to extend the reach and dissemination of high-quality information about crystal methamphetamine in Australia through an easily accessible and engaging mobile format. Specifically, the app has been designed to provide a condensed, offline-accessible version of the *Cracks in the Ice* Web-based toolkit [6], tailored to smartphone and tablet devices. These offline capabilities extend previous dissemination efforts by improving access for those without reliable internet access such as rural and regional residents. As with the website, the app has been designed to increase access to information about crystal methamphetamine, rather than actively facilitating behavioral change among end users (such as monitoring or decreasing drug use). The objective of the beta testing was to trial the *Cracks in the Ice* app with general community members, focusing on the usability, functionality, design, visual appeal, and app engagement. Furthermore, beta testing aimed to explore the impact of the app on perceived knowledge about crystal methamphetamine, attitudes toward crystal methamphetamine, and help-seeking intentions.

Methods

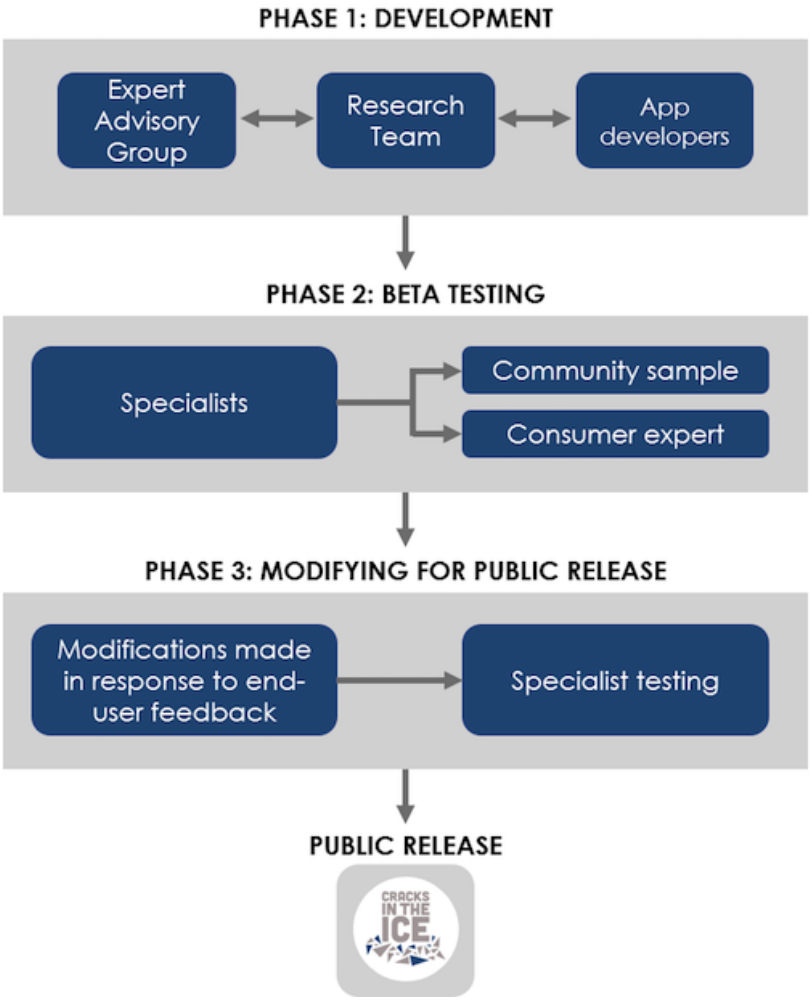
Overview of the Co-design Process

The *Cracks in the Ice* mobile app was developed over 4 months through a collaborative and iterative process using co-design. Importantly, co-design includes end users (people who will use the app after development) in the design process as experts on their experiences [12]. The inclusion of end users into the development of initial design concepts has been shown to result in outcomes with greater benefit to the user than ideas generated by in-house experts alone [13]. An established Expert Advisory Group (EAG), consisting of leading experts in drug and alcohol prevention and treatment, internet interventions, and mobile app development, provided guidance and recommendations throughout the development process. Multiple stakeholders were consulted throughout the process, including the core research team (that has expertise in addiction and mental health and was responsible for project management and oversight), app developers, a consumer expert with lived experience of addiction, and a sample of end users from the Australian community. Consumer participation was incorporated to ensure perspectives and needs of those directly impacted by the research were considered. Consumer participation is considered important for optimizing research outcomes across health fields, including mental health [14,15] and drug treatment [16].

The co-design process (Figure 1) consisted of 3 phases as follows:

1. Development of a beta version of the app, in collaboration with app developers and consultation with the EAG.
2. Beta testing among the EAG, app developers, and a community sample of end users (members of the Australian general population, including people who use crystal methamphetamine, their families and friends, health professionals, and concerned members of the community), and a consumer expert with lived experience of addiction.
3. Modifications to the app in response to end user feedback prior to public release.

Figure 1. The co-design process.



Phase 1: Development of the Beta Version of the *Cracks in the Ice* App

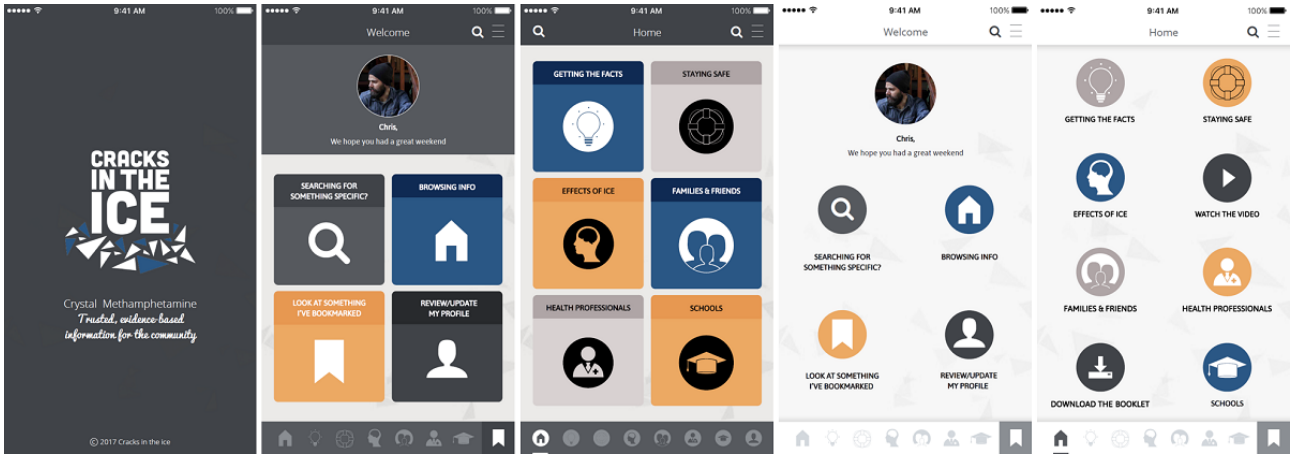
A beta version of the app was collaboratively developed by the research team, app developers, and EAG. To optimize reach, the app was designed to be compatible with the 2 most popular operating systems—iOS and Android. All app content was sourced from the existing Web-based toolkit [6] and included general information about crystal methamphetamine, its physical and mental health effects, and how to access support and treatment services, as well as more targeted resources for families and friends of individuals affected by crystal methamphetamine, health professionals, school, and community groups.

The app's structure and overall visual design (ie, branding, color scheme, and font) were designed to be consistent with the Web-based toolkit, which was informed by consultations and feedback with end users [7]. Figure 2 presents screenshots of the app's preliminary design options. Several functions from the toolkit were built into the app, including the following: (1) a search bar, allowing users to search for specific information by keyword(s); (2) a bookmarking function, allowing users to log in, save, and revisit specific sections or resources; and (3) a share button, allowing users to share content with others through email or social media. Two new functions included offline capabilities, allowing access to key information and resources without an internet connection, and "push notifications," allowing users to keep up-to-date with new resources, as they become available.

To ensure the toolkit's Web-based content was ready for transfer to mobile, all webpages from the toolkit were assessed for length and text density by 2 members of the research team. Pages considered too text heavy for mobile display were then condensed. To ensure no important information was lost during this process, each condensed page was then compared with the original Web-based version by another independent reviewer.

The content was then optimized for mobile display. Specifically, the Web-based content was consolidated into blocks of content known as snippets. Snippets are pieces of dynamic content (text, images, etc) that can be styled independently from the rest of the page and are often used to make large sections of text easier to read and navigate on mobile display. In addition, the Web content was packaged in the show/hide or "accordion" style sections in the app. Show/hide sections are responsive displays that expand and retract when the end user taps on them. When tapped by users, the section expands to "show" the full piece of content (eg, the full paragraph of text). Snippets and show/hide sections were designed to work together to enhance clarity and lessen the amount of scrolling for end users.

Figure 2. Example preliminary design options.



Phase 2: Beta Testing

At the completion of phase 1, the first beta version of the app was scripted in both iOS and Android operating systems for testing. As the beta testing study was a pilot, with the purpose of improving the app prior to release, it focused on the acceptability, usability, perceived knowledge, and attitudes after using the app. Beta testing consisted of 2 phases as follows: (1) specialist testing; and (2) end user testing. Specialist testing involved 4 experienced app developers and 2 members of the research team running initial tests to check for software bugs, errors, and crashes. A revised version of the app was then sent out for end user testing among a community sample and an expert consultant with lived experience of addiction.

Design and Procedure

End user beta testing took place over a 1-week period during September 2017. All aspects of this study were approved by the University of New South Wales Human Research Ethics Committee (HC15732). Testing was conducted through an anonymous internet survey (of approximately 30 minutes; a full copy of the survey is available on request). Participants were recruited through paid Facebook advertising, as well as through electronic notifications sent to the existing *Cracks in the Ice* website subscriber list and posted on the *Cracks in the Ice* Facebook and Twitter pages. Paid Facebook advertising was broadly targeted at Australian community members over the age of 16 years. Facebook recruitment for health research has been shown to result in samples that are generally representative of the total population and is particularly useful in engaging hard-to-reach populations, such as people who use drugs, while traditional methods tend to underrepresent these groups [17].

The survey was open to Australian residents aged ≥16 years who had access to an iOS or Android device capable of downloading and running mobile apps. All respondents were required to provide informed consent and were given the opportunity to enter a draw to win an iPad at the completion of the survey as reimbursement for their time.

On starting the survey, all participants were asked to download and preview the beta version of the app for 5-10 minutes before answering questions about its functionality, visual appeal, usability, and engagement. To ensure respondents used the app for a minimum of 5 minutes before providing feedback, a timer was incorporated into the survey preventing respondents from completing the evaluation questions until at least 5 minutes had passed.

Measures

The demographic data collected included gender, age, state or territory of residence, Aboriginal or Torres Strait Islander heritage, occupation, and number of children. In addition, participants were asked whether they had ever used the drug ice (yes or no), their frequency of use in the past year (ranging from “never” to “once a week”), whether they knew someone who uses the drug ice (yes or no), and whether they knew a friend or family member who uses ice (yes or no). Furthermore, participants were asked to report if they used a mobile or tablet device to access the app, how much prior experience they had of using mobile apps, whether they were aware of the *Cracks in the Ice* Web-based toolkit (yes or no), and how familiar they were with the Web-based toolkit prior to using the app.

Respondents were asked to rate the overall appeal of the app, layout, visual design, ease of use, features, and functions of the app through questions such as, “what do you think about the overall visual design of the app?” (rated from 1 “strongly dislike” to 5 “strongly like”). Several of these questions were adapted from the Mobile Application Rating Scale (MARS) [18], including questions measuring the likelihood to recommend the app to others, or use the app in future, as well as 2 questions assessing how easy and interesting the app was to use. The MARS is a well-established rating scale used by both professionals and end users to assess the quality of mobile apps. In addition, respondents were given the opportunity through open-ended questions to provide any suggestions for improvements or suggest additional content or features that they thought should be included.

To assess whether the app had any impact on an end user’s self-reported knowledge of crystal methamphetamine, respondents’ perceived level of knowledge about crystal methamphetamine was assessed before and after using the app. Specifically, participants were asked to rate their perceived knowledge of the drug crystal methamphetamine on a 4-item Likert scale ranging from 0 (“I have no knowledge of the drug ice”) to 4 (“I am very knowledgeable about the drug ice”). Additional questions, adapted from the MARS, assessed whether

respondents believed the app would have any impact on awareness, knowledge, and understanding of crystal methamphetamine or crystal methamphetamine prevention messages, attitudes toward crystal methamphetamine use, actual crystal methamphetamine use, and help-seeking behaviors among others who use the app.

Data Analysis

Data analyses were conducted in IBM SPSS Statistics 24 [19]; this included descriptive statistics and a Wilcoxon signed-rank test to investigate a change in participants' perceived level of knowledge about crystal methamphetamine after they had used the app.

Phase 3: Modifying the App for Public Release

Revisions were made in response to end user feedback to improve the app before its public release. Another round of testing by the research team ensured all changes were implemented correctly.

Results

Phase 1: Development of the Beta Version

Figure 3 summarizes key design elements and features of the final beta version of the *Cracks in the Ice* mobile app.

Phase 2: Beta Testing

End User Survey

Participants

A total of 34 participants completed the survey [age range, 21-60 years; mean, 37.2 (SD 10.1) years]. Table 1 summarizes further descriptive statistics.

Nearly all participants (33/34, 97%) were aware of the drug crystal methamphetamine ("ice"), with 82% (28/34) reporting they had some knowledge of the drug or were very knowledgeable about the drug. Furthermore, the majority of participants (24/34, 71%) were aware of the *Cracks in the Ice* Web-based toolkit.

Overall Response to the App

Participants' overall response to the app was positive. The app received a high average star rating of 3.8 out of 5, where a score of 1 corresponds to "one of the worst apps I've used" and 5 corresponds to "one of the best apps I've used." Most participants (28/34, 82%) "liked" or "strongly liked" the app overall. In addition, 94% (32/34) said that they would use the app again in the next 12 months if it was relevant to them, and 65% (22/34) said that they would recommend the app to "many people" or "everyone" who might benefit from using it. Table 2 provides examples of qualitative feedback.

Figure 3. Key design elements and features of the beta version of the *Cracks in the Ice* mobile app.

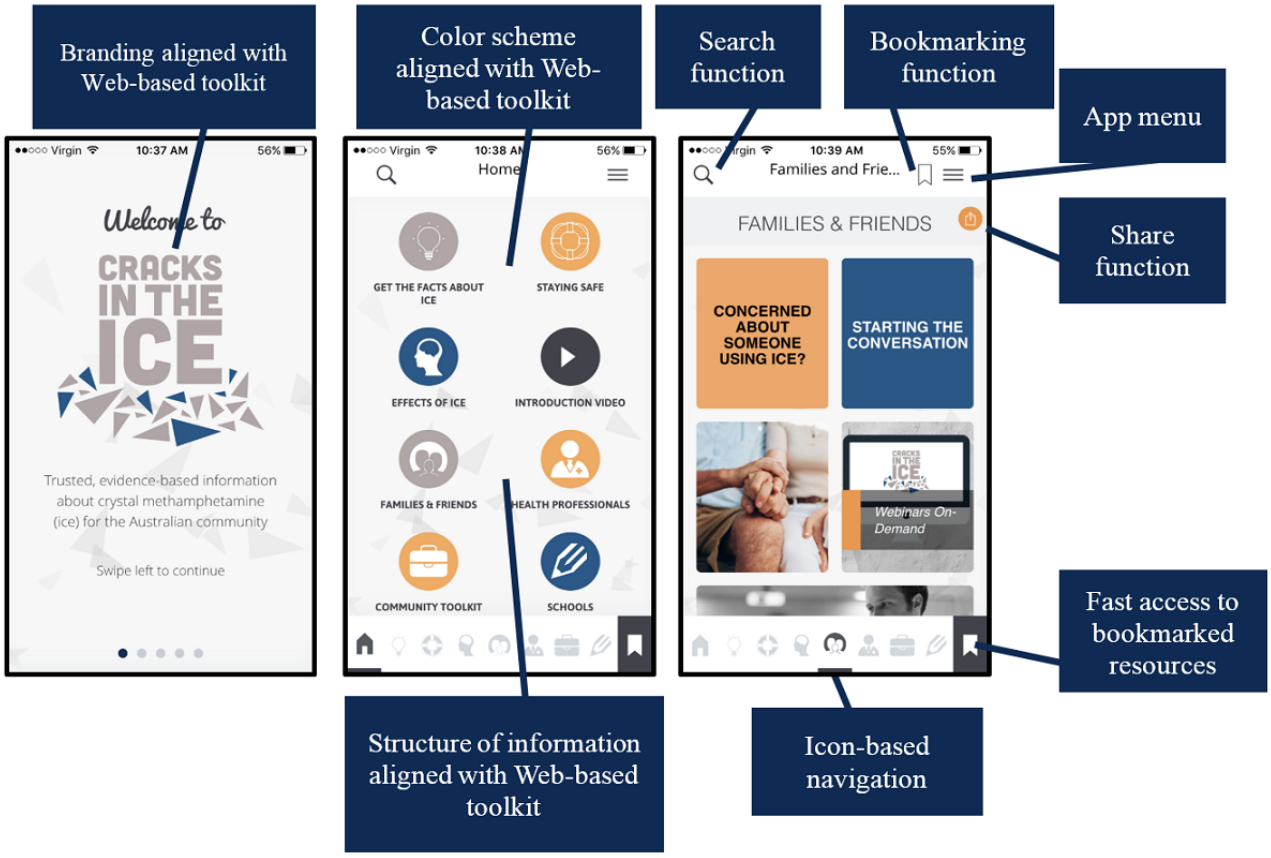


Table 1. Sample characteristics of survey respondents.

Characteristics	n (%)
Sex	
Male	7 (21)
Female	27 (79)
Aboriginal or Torres Strait Islander	
Yes	3 (9)
No	31 (91)
State of residence	
New South Wales	16 (47)
Queensland	6 (18)
South Australia	5 (15)
Victoria	2 (6)
Western Australia	2 (6)
Northern Territory	1 (3)
Tasmania	1 (3)
Australian Capital Territory	1 (3)
Area	
Metropolitan area	19 (56)
Regional area	9 (26)
Rural area	6 (18)
Employment	
Employed	32 (94)
Unemployed	2 (6)
Employment type	
Health Professional	13 (38)
Professional worker	9 (26)
Student	6 (18)
Other	6 (18)
Ever used the drug ice	
Yes	6 (18)
No	28 (82)
Used ice in the past year	
Yes	1 (3)
No	33 (97)
Know a friend who uses ice	
Yes	14 (58)
No	10 (42)
Have a family member who uses ice	
Yes	8 (33)
No	16 (67)
Mobile platform used to test app	
iPhone	20 (59)
Android	14 (41)

Characteristics	n (%)
Previously downloaded an app	
Yes	33 (97)
No	1 (3)
Use apps several times a week	
Yes	34 (100)
No	0 (0)
Downloaded an app in the last week	
Yes	24 (71)
No	10 (29)
Typically use apps several times a day	
Yes	24 (71)
No	10 (29)

Table 2. Example feedback from end users about the beta version of the app.

Type of feedback and app feature	Example feedback
Positive feedback	
Information and resources	• "I like the fact that there are resources on the app that are proven and reliable, and the way the information is presented is attractive and interesting" • "I like the app because it is easy to use, provides evidence-based information. Also, the fact that there is up to date information for health professionals as well." • "Caters to different groups/stakeholders"
Ease of use and navigation	• "It's well designed and easy to use" • "I love that it is almost a mirror of the website, so it will make it easier to use in my work." • "Smooth UI/UX[a]"
Visual design	• "Beautiful design" • "I love the colour scheme of the app. It's warm and inviting without being over the top which may be distracting." • "I really love the small icons at the bottom of screen."
Negative feedback	
Too much information	• "Too much information—overwhelmed me" • "Lot of information"
Text heavy	• "It's not bad but very wordy." • "I thought it was kind of wordy in places."
Low engagement	• "Bit boring"
Too similar to the website	• "It also comes off as a pure copy of the website rather than something new." • "Seemed very similar to the website"

[a]UI/UX: user interface/user experience.

Table 3. The frequency and proportion of respondents endorsing knowledge items before and after using app (n=34).

Response items	Before using app, n (%)	After using app, n (%)
I have no knowledge about the drug ice	2 (6)	0 (0)
I know very little about the drug ice	4 (12)	0 (0)
I have some knowledge about the drug ice	10 (29)	15 (44)
I am very knowledgeable about the drug ice	18 (53)	19 (56)

Table 4. Agreement with statements regarding app's impact on awareness, knowledge, understanding, attitudes, motivation, and behavior.

Statement	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
This app is likely to increase awareness of ice/ice prevention messages	0 (0)	5 (15)	4 (12)	17 (50)	8 (24)
This app is likely to increase knowledge and understanding of ice/ice prevention messages	0 (0)	2 (6)	6 (18)	18 (53)	8 (24)
This app is likely to change attitudes toward ice use	0 (0)	8 (24)	9 (26)	15 (44)	2 (6)
This app is likely to increase motivation to reduce ice use	2 (6)	6 (18)	17 (50)	7 (21)	2 (6)
Use of this app is likely to encourage further help seeking for ice use (if it's required)	1 (3)	1 (3)	10 (29)	19 (56)	3 (9)

Feedback on Specific App Features

Most participants (28/34, 82%) reported that they "liked" or "strongly liked" the overall visual design and layout of the app, with the majority (28/34, 82%) reporting the images and infographics were engaging. Most participants (29/34, 85%) found the app to be "moderately" or "very" interesting to use. The usability rated highly. Almost half the sample (15/34, 44%) reported being "able to use the app immediately," and over a third (12/34, 35%) found the app "easy to learn how to use" or agreed it "had clear instructions." Almost all participants agreed that offline functionality (32/34, 94%) and automatic information updates (33/34, 97%) would be useful.

Perceived Change in Personal Knowledge of Crystal Methamphetamine

A high number of participants indicated that they had a moderate-to-high level of knowledge about crystal methamphetamine after using the app, compared with prior to using the app (see Table 3). While we observed a trend toward increased perceived knowledge about crystal methamphetamine after using the app, this difference was not statistically significant (z=−1.90, P=.058).

Perceived Impact on Others' Awareness, Knowledge, Understanding, Attitudes, Motivation, and Behavior

Table 4 presents descriptive statistics in relation to the apps perceived impact on others' knowledge, attitudes, and behavior. The majority of participants "agreed" or "strongly agreed" that the app would likely increase awareness, knowledge, and understanding of crystal methamphetamine and associated prevention messages, as well as encourage further help seeking. Half "agreed" or "strongly agreed" the app would likely change attitudes toward crystal methamphetamine use. Comparatively fewer participants agreed the app would increase motivation to reduce crystal methamphetamine use.

Suggested Improvements

When asked about potential improvements to the app, almost all participants (32/34, 94%) agreed that they would like to be able to control how frequently the app sends notifications. Options to personalize other app features, such as the color scheme, font size, font style, welcome message, and password log-in, were less popular, with <30% endorsing each of these suggested changes. Almost two-thirds (22/34, 65%) wanted to see more infographics and images incorporated.

Suggestions for improvement were submitted through open-ended feedback and the most commonly cited included adding more information for people who use crystal methamphetamine, incorporating more information about national and local support services, and improving navigation. Other suggestions included making the app more interactive, incorporating stories of lived experience, and tailoring the information to be more mobile friendly. Multimedia Appendix 2 provides example qualitative feedback. When participants were asked how important it was for the app developers to action their suggested improvements, 35% (12/34) classified their suggested changes as "important," with the remaining participants being "unsure" (10/34, 29%) or classifying their suggestions as "not necessary/just a suggestion (the app would work well or very well without these improvements)" (12/34, 35%).

Consumer Expert Feedback

Feedback from the consumer expert was largely in line with the community sample. The expert "liked" the app overall, rating it 4 out of 5 stars, reporting they would recommend it to "many people." It was noted that the app was easy to navigate through its icon-based design. As with the community sample, the overall visual design, layout, engagement, functionality, and usability of the app was rated highly, and offline functionality and updates were endorsed as useful features.

The expert "strongly agreed" the app is likely to increase people's awareness, knowledge, and understanding of crystal methamphetamine and crystal methamphetamine prevention messages, and encourage further help seeking. They "agreed" it would likely change attitudes toward crystal methamphetamine use and increase motivation to reduce use. It was suggested that more images and infographics would improve the app, as well as more quotes from people with lived experience. In addition, it was suggested that the option for users to upload a picture to their profile should be removed, to eliminate the possibility that app users' identities could be exposed when viewing and bookmarking sensitive information about drug use.

Phase 3: Modifying the App for Public Release

Summary of Modifications

Revisions were made in response to end user and expert feedback, and these are summarized in Table 5. Figure 4 provides screenshots of the final app released to the public.

Usage Statistics

The *Cracks in the Ice* mobile app was officially launched on January 9, 2018. Approximately 4 months following launch, 797 unique users had downloaded the app (measured by "installs by user" on Google Play and "app units" on iTunes). Downloads have primarily been to devices within Australia (571 downloads); however, there has also been some interest from the United States (132 downloads). Uninstalls monitored through the Google Play store indicated a substantial proportion of Android app users (57.8%) have uninstalled the app after download. To investigate potential reasons for uninstalls, an optional open feedback form has since been incorporated in both the Android and iTunes app. Users are given the option of providing their feedback through this form within the first 10 minutes of app use and again one week later.

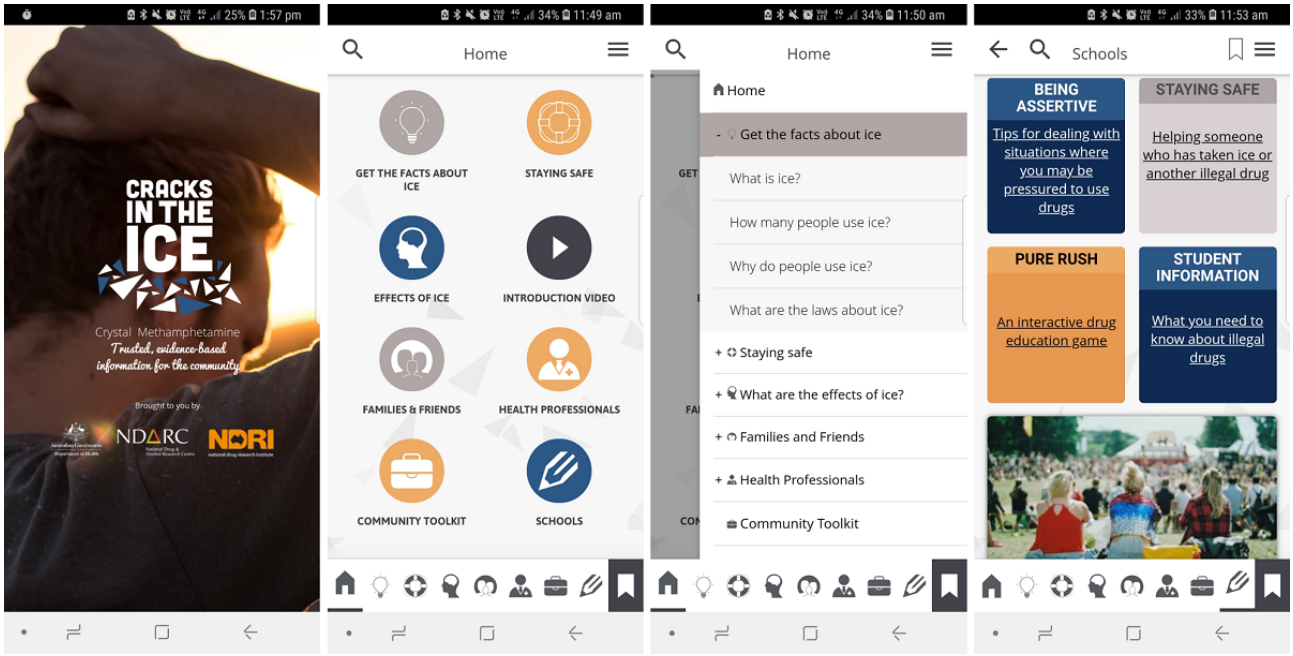
Table 5. Summary of feedback and modifications made to the app.

Suggestions from feedback	Modification(s) made
Include more images and infographics[a]	More show/hide displays, images and infographics were incorporated to make the app's content more mobile friendly. The majority of existing infographics were also optimized for mobile display.
Improve navigation[a]	To improve navigation, a back button was added to each of the 6 information tabs. This new feature allows users to easily navigate back to pages recently visited and return to the home page.
Improve navigation[a]	To further assist end users using the app for the first time, more references to the app's icons (corresponding with the app's 6 information tabs) were incorporated to ease navigation.
Add more information for people who use crystal methamphetamine, incorporating more national and local support services[b]	The listing of National Support Services was reviewed and considered sufficient. A local support list and inclusion of more information for people who use crystal methamphetamine was not implemented as it was considered outside of the scope of the initial release. The suggestion will be considered in later releases.
Potential for people's identities to be exposed when viewing and bookmarking sensitive information about drug use[a]	In accordance with the consumer expert's feedback, the profile picture function was disabled in the final version of the app.
Technical bugs[a]	All technical bugs were resolved and retested.

[a]Feedback actioned prior to the final release.

[b]Feedback not actioned

Figure 4. Screenshots of Cracks in the Ice mobile app when released to the public January 2018. Source: Cracks in the Ice App Version 1. Developers: National Health and Medical Research Council Centre of Research Excellence in Mental Health and Substance Use and Netfront Pty Ltd. Licensed under fair use.



Discussion

Principal Results and Comparison With Prior Work

To date, there has been a lack of easily accessible, up-to-date, evidence-based information about the drug crystal methamphetamine ("ice") for the general community. This is of particular importance in Australia, where community concern about crystal methamphetamine is at an all-time high [2], and harms related to crystal methamphetamine are rising [3,4]. This paper described the co-design and beta testing of the first mobile app, developed in collaboration with end users (Australian general community members) and leading experts, to provide easy access to evidence-based information and support options for crystal methamphetamine for the general community. The app was based on the Web-based *Cracks in the Ice Community Toolkit* [6] and provides an alternative method for community members to access information about crystal methamphetamine, in a convenient and engaging format.

Results indicated that the app was very well received and strongly liked. In particular, participants responded positively to the evidence-based information provided, usability, and visual appeal. Previous reviews have found that available drug-related mobile apps are rarely supported or developed with research evidence [20,21], with some apps even promoting illicit drug use through simulated drug taking and dealing [22]. Within this landscape, there is a clear need for accurate public health information and it is encouraging that the evidence base of the information was very positively rated by end users (Australian general community members). Yet, it is equally important that mobile apps are easy to navigate and engaging for users. Areas suggested for improvement in the current app included reducing the amount of text, increasing engagement, removing a profile picture, and improving navigation through the addition of a "back" button; all these suggestions were incorporated in the final version prior to public release, and it is likely that improvements to app engagement will increase usage and enhance recall of information. The revisions made to increase engagement and visual appeal make the *Cracks in the Ice* app the first of its kind to present evidence-based information in an engaging and appealing way to end users. Removal of the option to add a profile picture is in line with recommendations from a review of the potential of smartphones in addiction research and treatment, which outlines that protecting user privacy is a key ethical consideration when developing apps [23]. Although the current app does not actively collect information about drug taking or illegal behavior, it does focus on an illicit drug and users may prefer to remain anonymous.

While a higher number of participants agreed they had some knowledge of crystal methamphetamine or were very knowledgeable about the drug after using the app (compared with prior to using the app), this difference did not reach statistical significance; this is most likely because of the high baseline knowledge of the sample, with 82% (28/34) reporting they had at least some knowledge about crystal methamphetamine prior to using the app, therefore making it harder to see knowledge changes because of ceiling effects. While marked knowledge effects were not found in this study, participants did agree the app was likely to increase awareness, knowledge, and understanding about crystal methamphetamine. It is promising that most participants agreed the app would encourage help seeking; however, only one-quarter agreed the app would increase motivations to decrease actual crystal methamphetamine use. This is consistent with the app's goal, to provide high-quality, evidence-based information, rather than affect behavioral change.

While some end users offered suggestions to include more information about national and local support services, such directories require regular maintenance to avoid becoming out of date. In reviewing this request, it was decided the existing list of National Support Services, including >12 national directories and specific services for Aboriginal and Torres Strait Islander people, was sufficient. Listings of local services were considered outside of the scope of the current project to maintain. In addition, suggestions to add more information targeted at people who use crystal methamphetamine and include more stories of lived experience were considered valuable but also outside of the scope of modifications for the first release of the app; these suggestions will be considered for future releases.

Several findings merit further discussion. Notably, nearly all participants rated the app's ability to function while offline (once downloaded) and provide automatic updates to content, as useful. These 2 functions are relatively unique, setting the current app apart from others in the health information landscape. Offline availability is emerging as an important function for apps delivering time-critical information, for example, in clinical and disaster information, in which intermittent internet access can have detrimental effects [24,25]. Offline capabilities were built in the current app as a core function to enable access to key information and resources for people in areas where internet access may be unreliable (ie, rural or regional areas). This is particularly valuable in Australia, where there is evidence that the use of crystal methamphetamine is higher in remote rural areas [5], where internet connections are typically less reliable than that in urban areas. In addition, the app was built to automatically update content when updates are made to the companion website. The *Cracks in the Ice* website content is systematically reviewed once a year, with ad-hoc updates made as needed, to ensure new evidence and resources are included as they become available. Moreover, this process includes the removal of inactive or out-of-date links and resources as necessary. Such systematic maintenance is a unique feature in the quickly changing landscape of mobile apps and health information, and ensures the *Cracks in the Ice* app will remain current and based on the latest scientific evidence.

Strengths and Limitations

The key limitation of this study was that the beta testing was a pilot study of a small group of community members exploring the acceptability, usability, attitudes, and knowledge pre- and postapp usage. The small number of participants may not reflect the diverse Australian population. Females were overrepresented in the sample (27/34, 79% females), and a significant number of participants were based in one Australian state (16/34, 47%); however, the study did include at least 1 participant from each

state and territory in Australia. It should be noted that a relatively high number of participants (24/34, 71%) were aware of the *Cracks in the Ice* Web-based toolkit; this is likely attributed to the promotion of the beta testing survey through existing networks associated with *Cracks in the Ice*, including an associated Facebook page and Twitter account. While this group represents a key target audience, a future challenge may be how to engage and attract others, who are unfamiliar with the Web-based toolkit, to use the app. As this was a feasibility and acceptability study, it is not possible to make conclusions about the app's effects on knowledge, drug usage, and help-seeking behavior. A larger randomized controlled design with larger sample size is needed to test the app's effects on these constructs. A further limitation of this study is the short length of time participants were asked to use the app during beta testing. While the beta testing period was short, the development and expert testing by researchers and Web developers involved an in-depth examination of all content and design features over a period of months, so that the entire content was examined in-depth during the design process overall. While valuable information would have been gained by asking participants to go through the app in-depth over a period of several weeks, this would not have represented how most people use mobile apps in a real-world context. Furthermore, the aim of beta testing was to gain feedback on the feasibility and initial acceptability of the app from people who are likely to use the app after public release. It is the standard practice in the field to gain app evaluations after using an app for a short period (see MARS recommendations to use an app for a minimum of 10 minutes; Stoyanov et al [18]).

Key strengths of this study include the co-design approach, automated updating of the information on the app, the ability to access the app on both the iPhone and Android operating systems, and the availability of information offline once the app is downloaded. Another key strength was the inclusion of people with lived experience of the drug crystal methamphetamine. A substantial portion of the sample reported they had used crystal methamphetamine (6/34, 18%), and over half reported they had a friend who used crystal methamphetamine, with one-third reporting they had a family member who used crystal methamphetamine. These figures are much higher than national averages and give confidence that the target audience (people with experience or interest in the drug crystal methamphetamine) were included in the app design and testing. While the focus of this study was on the development and beta testing of the app, it will be of interest to track the app usage and explore how different end users engage with the app. It is envisaged that people will use the app in different ways depending on their circumstances, for example, rural health professionals interacting with clients who use drugs may utilize the offline capability repeatedly, while a family member wanting support services for their loved one who is at the point of crisis may find the resources they need in a single visit.

Although mobile apps with some focus on crystal methamphetamine do exist, for example, drug handbooks, games, and apps tracking drug use, very few have been developed through a scientific process or have involved co-designing with key stakeholders [8]. The development of the *Cracks in the Ice* mobile app used a co-design process, involving experts, researchers, app developers, a consumer expert with lived experience, and beta testing among a sample of app end users from the Australian community; this resulted in a wide range of perspectives being incorporated into the app, ensuring the final product was grounded in scientific evidence, and useful, relevant, and engaging for end users.

Conclusions

This is the first study to describe the co-design and beta testing of a mobile app to disseminate evidence-based information about the drug crystal methamphetamine or "ice." The design was an iterative process, incorporating a number of key stakeholders, including public health experts, app developers, end users, and a consumer expert. Initial findings show the app was well received and rated highly in terms of the usability, design, and the provision of high-quality information. Key improvements to the app included the addition of more infographics and images, show/hide test displays, and the addition of a back button to assist with app navigation. In addition, initial results indicate the *Cracks in the Ice* mobile app is easy to use, engaging, and acceptable to the target group. Translation of evidence-based information into a mobile app format, accessible offline, has the potential to increase the reach and impact of information and support services for people impacted by crystal methamphetamine.

Acknowledgments

We acknowledge the work of the app developers Netfront Pty Ltd, and the participants and other experts who provided invaluable feedback during the development, testing, and refinement of this app.

Cracks in the Ice is funded by the Australian Government Department of Health. This study was also supported by funding from the National Health and Medical Research Council (NHMRC) through an NHMRC Centre of Research Excellence (APP1041129).

Authors' Contributions

LB, HD, and CC conceptualized the paper. LB led the manuscript drafting and submission. LB, CC, and MT oversaw app development and beta testing. HD collected data for beta testing and worked with Web developers on refining the app. KEC, NCN, LAS, FKL, MT, and CC provided expert guidance of the development of the app, and all authors reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Adapted version of the National Health and Medical Research Council Body of Evidence Matrix.

[PDF File (Adobe PDF File), 47KB - mhealth_v6i12e11107_app1.pdf]

Multimedia Appendix 2

Example suggestions for improvement from end users about the app beta version.

[PDF File (Adobe PDF File), 22KB - mhealth_v6i12e11107_app2.pdf]

References

1. United Nations Office on Drugs and Crime. World Drug Report 2017. Vienna: United Nations; 2017. URL: https://www.unodc.org/wdr2017/index.html [accessed 2018-11-01] [WebCite Cache ID 73cSkX20t]
2. Australian Institute of Health and Welfare. National Drug Strategy Household Survey 2016: detailed findings. Canberra: Australian Institute of Health and Welfare; 2017. URL: https://www.aihw.gov.au/getmedia/15db8c15-7062-4cde-bfa4-3c2079f30af3/21028a.pdf.aspx?inline=true [accessed 2018-11-01] [WebCite Cache ID 73cSUhfYF]
3. Darke S, Kaye S, Duflou J. Rates, characteristics and circumstances of methamphetamine-related death in Australia: a national 7-year study. Addiction 2017 Dec;112(12):2191-2201. [doi: 10.1111/add.13897] [Medline: 28603836]
4. Degenhardt L, Sara G, McKetin R, Roxburgh A, Dobbins T, Farrell M, et al. Crystalline methamphetamine use and methamphetamine-related harms in Australia. Drug Alcohol Rev 2017 Mar;36(2):160-170. [doi: 10.1111/dar.12426] [Medline: 27286742]
5. Roche A, McEntee A. Ice and the outback: Patterns and prevalence of methamphetamine use in rural Australia. Aust J Rural Health 2017 Aug;25(4):200-209. [doi: 10.1111/ajr.12331] [Medline: 27868256]
6. Cracks in the Ice Community Toolkit. URL: https://cracksintheice.org.au/ [accessed 2018-03-14] [WebCite Cache ID 6xtqdx2P9]
7. Champion KE, Chapman C, Newton NC, Brierley M, Stapinski L, Kay-Lambkin F, et al. A Web-Based Toolkit to Provide Evidence-Based Resources About Crystal Methamphetamine for the Australian Community: Collaborative Development of Cracks in the Ice. JMIR Ment Health 2018 Mar 20;5(1):e21 [FREE Full text] [doi: 10.2196/mental.8891] [Medline: 29559427]
8. Chapman C, Champion KE, Birrell L, Deen H, Brierley ME, Stapinski LA, et al. Smartphone Apps About Crystal Methamphetamine ("Ice"): Systematic Search in App Stores and Assessment of Composition and Quality. JMIR Mhealth Uhealth 2018;6(11):e10442. [doi: 10.2196/10442]
9. Australian Communications and Media Authority. ACMA Communications Report 2016/17. 2017. URL: https://www.acma.gov.au/-/media/Research-and-Analysis/Report/pdf/Communications-report-2016-17-pdf.pdf?la=en [accessed 2018-11-02] [WebCite Cache ID 73cTiLBxT]
10. Deloitte. Global Mobile Consumer Survey 2017: UK Cut. 2017. URL: http://www.deloitte.co.uk/mobileuk/ [accessed 2018-11-02] [WebCite Cache ID 73cTqnAms]
11. Deloitte. Mobile Consumer Survey: The Australian Cut. 2017. URL: https://www2.deloitte.com/au/mobile-consumer-survey [accessed 2018-11-02] [WebCite Cache ID 73cTrLZ4y]
12. Visser FS, Stappers PJ, van der Lugt R, Sanders EB. Contextmapping: experiences from practice. CoDesign 2005 Apr;1(2):119-149. [doi: 10.1080/15710880500135987]
13. Trischler J, Pervan SJ, Kelly SJ, Scott DR. The Value of Codesign: The Effect of Customer Involvement in Service Design Teams. Journal of Service Research 2017 Jul 10;21(1):75-100. [doi: 10.1177/1094670517714060]
14. Happell B, Ewart SB, Platania-Phung C, Stanton R. Participative mental health consumer research for improving physical health care: An integrative review. Int J Ment Health Nurs 2016 Oct;25(5):399-408. [doi: 10.1111/inm.12226] [Medline: 27159221]
15. Linhorst DM, Eckert A. Involving people with severe mental illness in evaluation and performance improvement. Eval Health Prof 2002 Sep;25(3):284-301. [doi: 10.1177/0163278702025003003] [Medline: 12229070]
16. Brener L, Resnick I, Ellard J, Treloar C, Bryant J. Exploring the role of consumer participation in drug treatment. Drug Alcohol Depend 2009 Nov 01;105(1-2):172-175. [doi: 10.1016/j.drugalcdep.2009.06.016] [Medline: 19625143]
17. Whitaker C, Stevelink S, Fear N. The Use of Facebook in Recruiting Participants for Health Research Purposes: A Systematic Review. J Med Internet Res 2017 Aug 28;19(8):e290 [FREE Full text] [doi: 10.2196/jmir.7071] [Medline: 28851679]
18. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
19. IBM Corp. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp; 2016.

20. 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]
21. Photiou AJ. Smartphone Applications for Use in Drug Overdose: A review. J Clin Toxicol 2015;05(04). [doi: 10.4172/2161-0495.1000262]
22. Bindhim NF, Naicker S, Freeman B, Mcgeechan K, Trevena L. Apps Promoting Illicit Drugs—A Need for Tighter Regulation? Journal of Consumer Health On the Internet 2014 Feb 24;18(1):31-43. [doi: 10.1080/15398285.2014.869166]
23. Capon H, Hall W, Fry C, Carter A. Realising the technological promise of smartphones in addiction research and treatment: An ethical review. Int J Drug Policy 2016 Oct;36:47-57. [doi: 10.1016/j.drugpo.2016.05.013] [Medline: 27455467]
24. Herron J. Disaster Resources: Going Offline. Journal of Electronic Resources in Medical Libraries 2018 Feb 21;15(1):49-52. [doi: 10.1080/15424065.2018.1428136]
25. Bullock A, Dimond R, Webb K, Lovatt J, Hardyman W, Stacey M. How a mobile app supports the learning and practice of newly qualified doctors in the UK: an intervention study. BMC Med Educ 2015 Apr 08;15:71 [FREE Full text] [doi: 10.1186/s12909-015-0356-8] [Medline: 25889996]

Abbreviations

EAG: Expert Advisory Group
MARS: Mobile Application Rating Scale
NHMRC: National Health and Medical Research Council

Edited by G Eysenbach; submitted 21.05.18; peer-reviewed by M Deady, H Jin; comments to author 12.07.18; revised version received 03.08.18; accepted 08.09.18; published 20.12.18.

Please cite as:
Birrell L, Deen H, Champion KE, Newton NC, Stapinski LA, Kay-Lambkin F, Teesson M, Chapman C
A Mobile App to Provide Evidence-Based Information About Crystal Methamphetamine (Ice) to the Community (Cracks in the Ice): Co-Design and Beta Testing
JMIR Mhealth Uhealth 2018;6(12):e11107
URL: https://mhealth.jmir.org/2018/12/e11107/
doi:10.2196/11107
PMID:30573443

Original Paper

A Mobile App for Assisting Users to Make Informed Selections in Security Settings for Protecting Personal Health Data: Development and Feasibility Study

Leming Zhou[1], DSc, PhD; Bambang Parmanto[1], PhD; Zakiy Alfikri[1], MSc; Jie Bao[1], MSc

Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Leming Zhou, DSc, PhD
Department of Health Information Management
University of Pittsburgh
6021 Forbes Tower
3600 Forbes Avenue
Pittsburgh, PA, 15260
United States
Phone: 1 412 383 6653
Fax: 1 412 383 6655
Email: lmzhou@gmail.com

Abstract

Background: On many websites and mobile apps for personal health data collection and management, there are security features and privacy policies available for users. Users sometimes are given an opportunity to make selections in a security setting page; however, it is challenging to make informed selections in these settings for users who do not have much education in information security as they may not precisely know the meaning of certain terms mentioned in the privacy policy or understand the consequences of their selections in the security and privacy settings.

Objective: The aim of this study was to demonstrate several commonly used security features such as encryption, user authentication, and access control in a mobile app and to determine whether this brief security education is effective in encouraging users to choose stronger security measures to protect their personal health data.

Methods: A mobile app named *SecSim* (Security Simulator) was created to demonstrate the consequences of choosing different options in security settings. A group of study participants was recruited to conduct the study. These participants were asked to make selections in the security settings before and after they viewed the consequences of security features. At the end of the study, a brief interview was conducted to determine the reason for their selections in the security settings. Their selections before and after the security education were compared in order to determine the effectiveness of the security education. The usability of the app was also evaluated.

Results: In total, 66 participants finished the study and provided their answers in the app and during a brief interview. The comparison between the pre- and postsecurity education selection in security settings indicated that 21% (14/66) to 32% (21/66) participants chose a stronger security measure in text encryption, access control, and image encryption; 0% (0/66) to 2% (1/66) participants chose a weaker measure in these 3 security features; and the remainder kept their original selections. Several demographic characteristics such as marital status, years of experience using mobile devices, income, employment, and health status showed an impact on the setting changes. The usability of the app was good.

Conclusions: The study results indicate that a significant percentage of users (21%-32%) need guidance to make informed selection in security settings. If websites and mobile apps can provide embedded security education for users to understand the consequences of their security feature selection and the meaning of commonly used security features, it may help users to make the best choices in terms of security settings. Our mobile app, SecSim, offers a unique approach for mobile app users to understand commonly used security features. This app may be incorporated into other apps or be used before users make selections in their security settings.

(JMIR Mhealth Uhealth 2018;6(12):e11210) doi:10.2196/11210

KEYWORDS

data security; mobile app; education; feasibility studies

Introduction

Background

In recent years, health data breaches have begun to occur more frequently, impacting a growing number of people. From February 10, 2016, to February 6, 2018, there were 2201 reported Protected Health Information (PHI) breaches in the United States and each affected 500 or more individuals; in total, more than 177 million Americans (54.1% of the US population) across the nation were affected by these PHI breaches [1].

PHI breaches are costly to industries. For example, according to the 2017 Cost of Data Breach Study released by IBM Security and the Ponemon Institute, the average global cost of a health data breach per lost or stolen record was US $380 [2]. Overall, the US health care industry spent approximately US $67 billion dealing with issues triggered by PHI breaches on activities such as conducting investigations, notifying customers, recovering data, subscribing to credit monitoring services for customers, hiring knowledgeable security personnel, and strengthening the security measures of information technology (IT) systems.

There are many good approaches to reduce the number of PHI breaches such as using a highly qualified security incident response team, extensively using encryption in the IT system, and providing security training to health IT system users [2]. Compared with the first approach (using highly qualified security experts), providing security training to health IT system users tends to be a very cost-efficient and effective approach. After all, among these PHI breaches in the last 2 years, only 19% were because of hacking or IT incidents, which are handled by security incident response teams. The other, more than 80%, were because of issues on the user end such as improper disposal of PHI, theft or loss of devices, and incidents of unauthorized access or disclosure [1].

The security education of users is particularly important as smartphones and tablets are widely used in the health care industry for PHI access. By the end of 2017, 77% of Americans owned a smartphone and 53% of Americans owned a tablet computer, compared with those in 2011, when ownership of these 2 mobile computing devices was just 35% and 8%, respectively [3]. As the mobile user population has grown, smartphones and tablets have become popular within the health care domain for both providers and patients. According to a recent survey study of 3800 physicians, 83% owned at least one mobile device and 25% of these physicians used both smartphones and tablets within their clinical practice [4]. Similarly, many patients use their mobile devices to receive health care services [5]. As the health care–related uses increase and more sensitive information is accessed via mobile devices, there is a growing need for users (both health care providers and patients) to be conscious of information security.

Previous studies have indicated that mobile health (mHealth) app users, especially patients, are concerned about their health data security and their individual privacy, and some users choose not to use mHealth apps because of this concern [5-8]. mHealth app users' perception of security and privacy are highly contextual and are related to multiple demographics such as age, gender, income, race, health status, and education [9-11].

On the technical end, mHealth apps and mobile operating systems offer various security features such as passcodes, usernames and passwords, data encryption, and remote wiping. Researchers have also provided detailed security recommendations for mHealth app development in particular [12]. However, many smart device users did not use even the most basic authentication features (such as a passcode) to prevent the access of private data on their mobile devices [13-15]. In other words, security features are available to mHealth app users, but the problem is whether these users are capable of using these security features to protect the PHI.

On today's websites and mobile apps, a security setting page and privacy policy are often provided to users. Examples of the security setting page are the "Touch ID & Passcode" page in iOS and the "Sign-in & Security" page in Gmail.

Privacy policies detail a website's or mobile app's specific practices with regards to data collection, storage, and use. It is assumed that users of the website or mobile app would be able to understand the content of these privacy policies. These privacy policies can be very useful for people who can understand the security terms and technologies such as encryption, access control, and security protocol names. However, for people who have not had a chance to receive formal education in information security (a majority of people), it is fairly challenging to fully understand the content of the privacy policy, let alone make an informed selection in the security settings. A specific example from Apple demonstrates this. In Apple's privacy policy updated on May 22, 2018, it stated that "Apple online services such as the Apple Online Store and iTunes Store protect your personal information during transit using encryption such as Transport Layer Security (TLS). When your personal data is stored by Apple, we use computer systems with limited access housed in facilities using physical security measures. With the exception of iCloud Mail, iCloud data is stored in encrypted form including when we utilize third-party storage." [16] Even if ignoring the specific name of the protocol (eg, TLS), most people will still wonder what the terms *encryption*, *encryption during transit*, and *limited access* mean.

The other assumption on privacy policy is that users will read the privacy policy carefully and make the proper selections for their PHI on the security setting page according to their security and privacy needs. However, an earlier study also indicated that participants often install mobile apps from unfamiliar vendors without reading the app's privacy policies [17]. The existence of such a high number of PHI breaches also indicates that this assumption is not true.

People can make proper selection in security settings only if they have sufficient knowledge on this topic. However, most people do not receive formal, intensive security education. Even health care providers, who typically receive training on the Health Insurance Portability and Accountability Act regulations and the specific policies of their organization, do not receive much training on information security itself. In other words, many people do have the general idea that there is a potential for security risks. If they do not have a clear idea about how to use the protection provided by specific security features, they may take risky actions or intentionally sacrifice their information security and privacy for convenience or a small amount of financial benefits [18].

Objectives

The purpose of this study was to determine whether a brief security education offered in a mHealth app can change users' behavior in choosing security settings to improve the current situation. Our hypothesis is that mHealth app users can benefit from this brief and informal security education, and once they receive such education, many of them will choose a stronger security measure if they did not do so initially. Here *users* can be both health care providers and patients.

Methods

Features of the Mobile App

In this study, we chose a few commonly used security features in health IT systems and implemented a simulation or demonstration of these features in a mobile app named SecSim (Security Simulator). The chosen security features were as follows: (1) data encryption (a process of converting plain text into something that appears to be random and meaningless), (2) user authentication (a process that allows an entity, such as a Web server, to verify the identity of someone), (3) access control (the selective restriction of access to information or other resources), and (4) image encryption (a process of hiding the meaning of a private or sensitive image). We also elicited study participants' opinions on password update frequency and preferred data storage and backup locations.

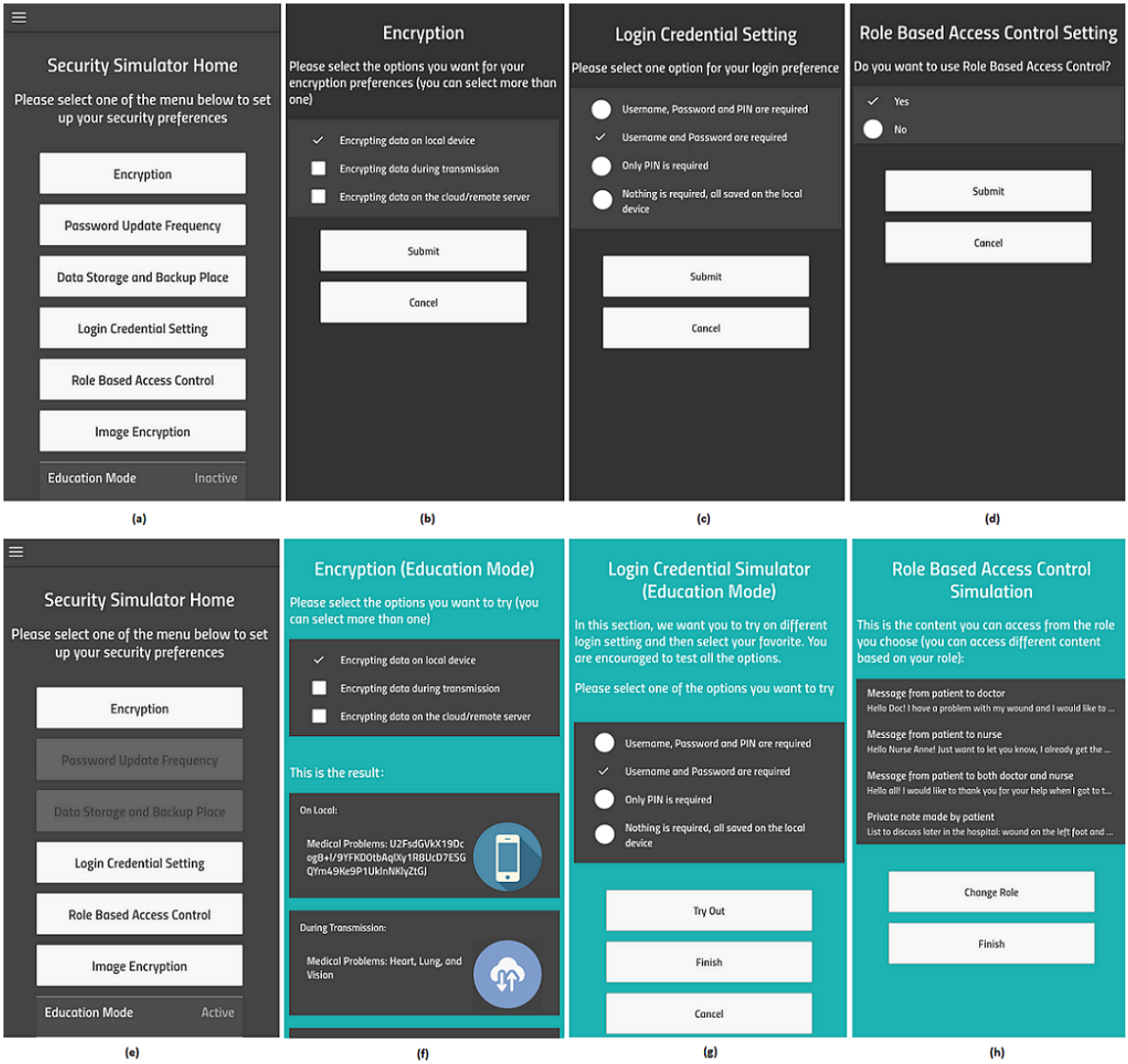
This mobile app, SecSim, has several major components, including (1) pages for registering the users and collecting the individual's personal health data, which is the information users wants to protect in the app; (2) pages for simulating or demonstrating the chosen security features; (3) pages offering the user security setting options; (4) a log-in page; and (5) a page with a summary of security settings chosen. The order of running these components was 1-4-3-2-3-5. In other words, first the user registers a unique account and enters some personal health information. Then, the user can log in and make selections in security settings, the same as he or she does when using other mobile apps or websites. These actions take place before the security education (or where the education mode is not activated). In the next step, the user goes through each of the implemented security features and views the consequences of their selection. This is the mode where the security education takes place (or where the education mode is activated). At the end of the demonstration, the user is given a chance to make another round of selections in the security settings and view the summary of his or her selections.

Figure 1 shows some screenshots of the app's components. The top 4 screenshots show options available in (a) the main page of the app, (b) the encryption page, (c) the log-in credentials page, and (d) the role-based access control (RBAC) page when the education model is not activated. In this mode, users simply make their selections according to their own understanding. Once the security education mode is activated, the corresponding pages are updated. The bottom 4 screenshots show the following updated contents: (e) the options available on the main page, (f) the simulation for when different encryption options are selected, (g) the simulation for when different combinations of log-in credential are selected, and (h) a page of the RBAC simulation. On the encryption page, when no option is selected, all the contents are shown in clear text (not shown). If one of the options in (f) is selected, the contents at this location are shown as cipher texts and others are still shown as clear text (the content on the remote server is at the lower part of the screen, not shown in the screenshot). On the log-in credential simulator page (g), after one option is selected, the corresponding log-in page is shown, and the user is required to enter the log-in credentials accordingly to enter the system. On the RBAC simulation page (h), the user can choose one of the 3 roles: patient, physician, or nurse, and the corresponding content is shown in a new page (h). Not included in the figure are the pages for log-in, user registration or data collection, image encryption, the details of the log-in simulation and RBAC simulation, and the security settings summary page. Below is a detailed description of the options available corresponding to the implemented security features in the SecSim app.

First, there are 5 options for password update frequency: (1) once a month, (2) once every 3 months, (3) once every 6 months, (4) once a year, and (5) never. There are 3 options for data storage and backup location: (1) on local drive only, (2) on remote server only, and (3) both on local drive and remote drive.

On the encryption page, there are 3 options for data encryption: (1) applying encryption on local device (mobile device), (2) applying encryption during the transmission, and (3) applying encryption on the cloud server. Users could choose none, one, or more than one of these 3 options. For this specific question, the more options selected, the stronger the security. In other words, encrypting data at all three places is the strongest, encryption data at any two places is not as strong, and only encrypting data at one place is even weaker, with the weakest being when no encryption is applied on data. To compare the selections made by participants on data encryption, we introduced a concept named protection level (PL) as a quantifiable measure. If they chose to encrypt data at one place only, the PL was 1; if they applied encryption on data at two locations, the PL value was 2; and if they chose to encrypt the data at all three locations, the PL value was 3.

Figure 1. Screenshots for the list of security features and the options and simulation of 3 security features.



On the log-in credential page, there are 4 options for log-in credentials: (1) users are required to first provide a username and the corresponding password and then a randomly generated personal identification number (PIN) for 2-factor authentication, (2) users are required to enter their username and password only, (3) users only need to enter a PIN, and (4) users are not required to enter anything to log into the system as the log-in credentials are stored on the local device. Obviously, the security strength decreases as the number of items in the log-in credentials needed for accessing the system decreases. Among these log-in options, the strongest is the 2-factor authentication (username, password, and PIN); however, this choice also requires the largest number of steps, which makes it the least convenient or with the lowest usability from the user's perspective.

On the RBAC page, there are 2 options for RBAC: (1) not using RBAC and (2) using RBAC. Clearly, option 2 is stronger than option 1. Similarly, on the image encryption page, there are also 2 options for image encryption: (1) not using imaging encryption and (2) using image encryption. Again, option 2 is stronger than option 1.

When we designed the SecSim app, we intentionally did not design an elaborate user interface but instead used a very clear and simple design to explain the meaning of each security feature and the consequences of each option for each security feature. After all, the purpose of this app is to provide a brief security education before the user makes a decision in their security settings. An interface with a lot of details and fancy colors may actually distract users from the core content of the security education.

Study Participant Recruitment

The participants were recruited through flyers distributed at public places in the Greater Pittsburgh area and on the Pitt+Me website at the University of Pittsburgh. Participants were screened using the following selection criteria: native English speaker, high school or higher education, aged between 18 and 65 years, capable of communicating with others verbally and

in writing, and having at least a few years of experience using smart devices such as a smartphone, tablet, or smart watch.

Study Procedure

Before the study, all study participants were required to read a consent form and sign the consent form if they were willing to participate in the study. The study participation was completely voluntary, and the participant could leave the study at any time. At the beginning of the study, a general introduction to the purpose of the study and the mobile app (SecSim) was provided, along with a brief demo of the app. All the participants were then asked to use the SecSim mobile app on an Android tablet (Samsung Galaxy Tab 4 10.1 inches, 16 GB, white tablet, Android version 4.4.2) to finish the tasks described earlier, such as registering an account, making selections in security settings, and receiving the brief security education. At the end of the study, the study participants were asked to provide responses to a usability questionnaire (IBM Post-Study System Usability Questionnaire [PSSUQ]) [19] and answer a few open-ended interview questions. On the PSSUQ, study participants were asked to respond to the 19 statements, with a scale ranging from 1, meaning strongly agree, to 7, meaning strongly disagree. The study participants were asked to fill out the usability questionnaire via the Web-based Qualtrics system. The open-ended questions were used to obtain study participants' comments and suggestions on this study, the SecSim mobile app itself, the implemented security features on the mobile app, their ideas about information security in general, and their source of security knowledge (eg, classes, friends and family) if any.

Study participants' demographics and their responses to the PSSUQ usability questionnaire were exported into a SPSS data file. The study participants' selections in the security settings (both before and after the security education) were also downloaded from the SecSim app. IBM SPSS version 24 was used to perform the data analysis. Mean and SD were reported for the usability study. One-way analysis of variance (ANOVA) was used to find the setting differences among the various demographic groups.

The selections made by the study participants before and after the security feature demonstration (the brief security education) were compared and assigned to 3 categories: weaker, no change, and stronger. The number of study participants in each category for each security feature was calculated. One-way ANOVA was also used to determine the setting change behavior among the various demographic groups.

Results

Demographics

The study was conducted from May 2017 to September 2017 in the Greater Pittsburgh area. In total, 66 participants were recruited to undertake the study. The mean age of participants was 31.1 years (SD=13.42). More specifically, there were 40 participants (40/66, 61%) aged 18 to 28 years, 16 participants (16/66, 24%) aged 29 to 50 years, and 10 participants (10/66, 15%) aged 51 to 65 years. The gender of participants was balanced. There were 31 males (31/66, 47%) and 35 females (35/66, 53%). There were 11 African Americans (11/66, 17%), 38 white Americans (38/66, 58%), and 17 Asian Americans (17/66, 26%). Furthermore, 25 participants (25/66, 38%) had received an associate's degree or lower education, 17 (17/66, 26%) had a Bachelor's degree, and 24 (24/66, 36%) had a graduate degree. A total of 51 (51/66, 77%) participants were single, 13 (13/66, 20%) were married or in a long-term committed relationship, and 2 (2/66, 3%) were divorced or separated. Overall, 48 (48/66, 73%) participants lived in an urban area, 16 (16/66, 24%) lived in a suburban area, and 2 (2/66, 3%) lived in a rural area. Most of these study participants (45/66, 68%) had a part-time or full-time job, 17 (17/66, 26%) were not employed, and the other 4 participants (4/66, 6%) were retired or disabled. These participants had diverse occupations, including student, researcher, administrative personnel, and customer services personnel such as chef, bartender, other restaurant service person, as well as teacher, professor, attorney, and census field representative.

The participants were asked to perform a self-assessment on their own health status, rating it as excellent, very good, good, fair, or poor. None chose poor. In total, 3 (5%) participants chose fair, 19 (19/66, 29%) selected good, 24 chose very good (24/66, 36%), and the rest (20/66, 30%) claimed their health was excellent. Overall, 19 (19/66, 29%) of these participants used Android-based smartphones or tablets, 44 of them (44/66, 67%) used iOS-based mobile devices, and the other 3 participants used different mobile operating systems. The average number of years of experience using smart devices was 6.0 (SD=2.59). More than half of these participants (38/66, 58%) had used mHealth apps such as Apple Health, MyFitnessPal, MyChart, Fitbit app, Pink Pad, Clue, SnoreLab, 10% Happier, 7 Minute Workout, Garmin Connect, and Samsung Health. The household income of the study participants fit into 6 categories: less than US $10,000 (13/66, 20%), between US $10,001 and US $25,000 (14/66, 21%), between US $25,000 and US $50,000 (18/66, 27%), between US $50,000 and US $100,000 (7/66, 11%), greater than US $100,000 (9/66, 14%), and decline to answer (5/66, 8%). The demographic information is summarized in Table 1.

Security Settings Before Education

As there was no simulation on password update frequency or data storage and backup locations, study participants were simply asked to make selections for the given options. As it turned out, 11 participants (11/66, 17%) chose to update their password once a month, 15 participants (15/66, 23%) chose once every 3 months, 19 participants (19/66, 29%) chose once every 6 months, 6 participants (6/66, 9%) chose once a year, 13 participants (13/66, 20%) chose never, and the remaining 2 participants (2/66, 3%) did not make any selection for this question. Furthermore, 34 participants (34/66, 52%) chose local device only for data storage and backup, 5 participants (5/66, 8%) chose remote server only, and 25 participants (25/66, 38%) chose to use both local device and remote server for data storage and backup. There were 2 participants (2/66, 3%) who did not indicate a preference on data storage and backup.

Table 1. Demographic characteristics of the study participants (N=66).

Demographic characteristic	Value
Age in years, mean (SD)	31.1 (13.42)
Years of using smart mobile devices (1-10), mean (SD)	6.0 (2.59)
Age in years, n (%)	
18-28	40 (61)
29-50	16 (24)
51-65	10 (15)
Gender, n (%)	
Male	31 (47)
Female	35 (53)
Race, n (%)	
Black	11 (17)
White	38 (58)
Asian	17 (26)
Education, n (%)	
High school or lower	2 (3)
Some college, no Bachelor's degree	23 (24)
Bachelor's	17 (26)
Graduate	24 (36)
Marital status, n (%)	
Single	51 (77)
Married or in a long-term committed relationship	13 (20)
Divorced or separated	2 (3)
Living place, n (%)	
Urban	48 (73)
Suburban	16 (24)
Rural	2 (3)
Employment, n (%)	
Employed, working 1 to 20 hours per week	14 (21)
Employed, working 21 to 40 hours per week	22 (33)
Employed, working more than 40 hours per week	9 (14)
Not employed, looking for a job	9 (14)
Not employed, not looking for a job	8 (12)
Retired or disabled	4 (6)
Occupation, n (%)	
Student	24 (36)
Researcher	10 (15)
Administrative personnel	6 (9)
Customer service	5 (8)
Retired, disabled, unemployed	4 (6)
Other	14 (21)
No answer	3 (5)
Self-assessed health status, n (%)	

Demographic characteristic	Value
Excellent	20 (30)
Very good	24 (36)
Good	19 (29)
Fair	3 (5)
Mobile OS, n (%)	
Android	19 (29)
iOS	44 (67)
Other	3 (5)
Used mobile health apps?, n (%)	
Yes	38 (58)
No	28 (42)
Household income, n (%)	
≤US $10,000	13 (20)
US $10,001–US $25,000	14 (21)
US $25,001–US $50,000	18 (27)
US $50,001–US $100,000	7 (11)
>US $100,000	9 (14)
Decline to answer	5 (8)

Table 2 lists the number and percentage of study participants who chose the specific options in those security features before and after they received the security training in the SecSim app. One-way ANOVA was used to determine whether their initial security settings were significantly different for participants having different demographic characteristics. The results indicated that responses to security settings did not differ significantly (P>.05) for participants of different ages, gender, race, marital status, living environment, employment status, household income, mobile operating system, occupation, or education.

However, password update frequency did differ significantly ($F_{3,60}$=5.208, P=.003) for participants in different self-assessed health status categories. More specifically, participants who reported being in *very good* health did not want to change their password as frequently as those who reported being in *good* health. One possible reason is that participants with *very good* health believed that they did not have much highly sensitive health information to protect. This was confirmed by the responses from participants who claimed to have *excellent* health, which are similar to the ones from participants with *very good* health (P=.98). Therefore, the difference in the password update frequency between participants with *excellent* and *good* health was also large, although it was not statistically significant (P=.08). Here, we ignored the responses from participants with *fair* health because of the small number of participants in that category (n<5). This rule (not including categories with less than 5 participants) applies for all of the other one-way ANOVA analysis results below as well.

Security Setting Changes After Education

All the study participants were able to finish all the assigned tasks easily in approximately 10 min. None of them had any significant difficulties using the mobile app during the study. After the brief security education, study participants performed another round of security preference selection (see Table 2, last column). These selections made in this round were compared with the initial selections made before the security education, and the results were arranged into 3 categories: stronger, weaker, and no change, as shown in Table 3.

In terms of encryption, 21 participants (21/66, 32%) chose to use a stronger measure after education; for instance, instead of only encrypting the data on the local device or remote server, they chose to encrypt data at both locations or at all 3 places. No one chose to use a weaker security measure after the security education.

There were 33 (33/66, 50%) study participants who wanted to use RBAC before the security education; after the education, that number increased to 53 (53/66, 80%), that is, 20 (20/66, 30%) participants chose to use a stronger security measure in access control. Only 1 (1.5%) participant chose to use a weaker security measure.

There were 44 participants (44/66, 67%) who wanted to use image encryption in the initial selection; after the security education, 58 participants (58/66, 88%) wanted to use image encryption. In other words, 14 (14/66, 21%) more participants chose to use a stronger security measure for image protection. Only 1 participant (1/66, 2%) chose a weaker security measure for image protection.

Table 2. Six security features implemented in the SecSim app, their options, and the selections made by 66 study participants before and after the security education.

Feature label	Feature description	Before, n (%)	After, n (%)
Encryption			
1	Encrypting data on local device (PL[a]=1)	28 (42)	18 (27)
2	Encrypting data when transmission (PL=1)	6 (9)	5 (8)
3	Encrypting data on the remote server (PL=1)	13 (20)	7 (11)
1,2	Encrypting data on local device and during transmission (PL=2)	0 (0)	2 (3)
1,3	Encrypting data on local device and remote server (PL=2)	1 (2)	0 (0)
2,3	Encrypting data during transmission and on remote server (PL=2)	3 (5)	5 (8)
1,2,3	Encrypting data on local, remote device and during transmission (PL=3)	14 (21)	29 (44)
	No answer	1 (2)	0 (0)
Password update frequency			
1	Once a month	11 (17)	—[b]
2	Once every 3 months	15 (23)	—
3	Once every 6 months	19 (29)	—
4	Once a year	6 (9)	—
5	Never	13 (20)	—
	No answer	2 (3)	—
Data storage and backup location			
1	On local device only	34 (52)	—
2	On remote server only	5 (8)	—
3	Both on local device and remote server	25 (38)	—
	No answer	2 (3)	—
Log-in credential			
1	Username, password, and PIN are required	5 (8)	10 (15)
2	Username and password are required	24 (36)	18 (27)
3	Only PIN is required	16 (24)	16 (24)
4	Nothing is required, all saved on the local device	20 (30)	21 (32)
	No answer	1 (2)	1 (2)
RBAC[c]			
1	Not using RBAC	31 (47)	13 (20)
2	Using RBAC	33 (50)	53 (80)
	No answer	2 (3)	0 (0)
Image encryption			
1	Not using image encryption	20 (30)	8 (12)
2	Using image encryption	44 (67)	58 (88)
	No answer	2 (3)	0 (0)

[a]PL: protection level.

[b]Not applicable.

[c]RBAC: role-based access control.

Table 3. A summary of the changes in security option selection after security education (N=66).

Security features	Stronger, n (%)	Weaker, n (%)	No change, n (%)
Encryption (local, remote, and transmission)	21 (32)	0 (0)	45 (68)
Log-in credentials	8 (12)	7 (11)	51 (77)
Role-based access control	20 (30)	1 (2)	45 (68)
Image encryption	14 (21)	1 (2)	51 (77)

In these 3 categories (encryption, RBAC, and image encryption), the number of participants who chose a stronger security measure is much larger than the ones who chose a weaker security measure after the education.

However, the change for the log-in setting was quite different. Only 8 participants (8/66, 12%) chose to use a stronger measure during log-in, whereas 7 participants (7/66, 11%) actually chose to use a weaker measure during log-in. The vast majority of them (51/66, 77%) chose not to change their original selection. This is expected as the participants *experienced* all 4 different log-in procedures, which required different number of steps in the app, and they wanted to balance convenience and security. Their reasoning was also confirmed by the answers to the brief interview questions at the end of the study (described in a later section).

One-way ANOVA was used to determine whether the changes in settings were significantly associated with any demographic characteristics. Participants in different age, gender, race, living place, mobile operating system, occupation, and education groups did not show a statistically significant difference in their security setting behavior after the security education. On the other hand, people in different marital status, years of experience using mobile devices, household income, employment status, and health status groups showed significantly different security setting behavior after the education.

Marital Status

Participants in different marital status groups had different setting behavior for image encryption ($F_{2,63}$=3.373, P=.04). For the single participants, only 14% (7/51) switched to a stronger protection (using image encryption) and 2% (1/51) switched to a weaker security (not using image encryption), whereas almost half of the married participants 46% (6/13) switched to a stronger security and none of them switched to a weaker security after the security education.

Years of Using Mobile Devices

Participants having different amounts of experience of using mobile devices showed a statistically significant difference in choosing options for image encryption ($F_{2,63}$=3.870, P=.03). More specifically, for the participants with 3 to 5 years' experience using mobile devices, only 13% (3/24) changed to a stronger security measure and 4% (1/24) changed to a weaker security measure. For the participants with more than 5 years of experience using mobile devices, 35% (12/34) switched to a stronger security protection and none of them switched to a weaker security.

Income

Participants in different income groups showed a statistically significant difference in setting change behavior for RBAC ($F_{5,60}$=3.846, P=.004). For the participants with income between US $10,001 and US $25,000, only 7% (1/14) switched to a stronger protection (using RBAC) and 7% (1/14) switched to a weaker security (not using RBAC). For the participants with an income greater than US $100,000, 78% (7/9) changed to a stronger security measure and none of them switched to a weaker security measure after the security education.

Employment and Health Status

Participants in different employment groups had different setting change behavior for encryption ($F_{5,60}$=2.807, P=.02); however, none of the compared pairs of groups showed a statistically significant difference. Similarly, participants in different health status groups had different setting change behavior for log-in credential selection ($F_{3,62}$=2.816, P=.046); however, none of those compared pairs of groups had a statistically significant difference.

Usability Study Results

As mentioned above, the PSSUQ usability questionnaire contained 19 statements for which study participants were required to choose answers on a scale of 1 to 7, where 1 meant strongly agree and 7 meant strongly disagree. Table 4 shows the average and SD of the 66 study participants' responses to each statement.

It is clear that most of the average values were around 2 out of 7; in other words, these study participants agreed with almost all the statements, indicating good usability. The exception was the score for statement 9: *the system gave error messages that clearly told me how to fix the problems*. In most cases, if the study participants paid attention during the demo session at the beginning of the study and strictly followed the instructions given by the investigator, the app would not generate any error messages as everything they did was correct. Only if the study participant did not follow the instructions or did not enter the correct information at the right place, would the error message pop up. Therefore, a large portion of these study participants finished the entire study without any problem and did not see any error message, and therefore, they were not sure how to respond to the statement about error message. This issue is quite common in many other usability studies using PSSUQ, and a higher value in this statement does not indicate a poor usability [19].

Table 4. A summary of usability study results.

Post-Study System Usability Questionnaire	Mean (SD)
1. Overall, I am satisfied with how easy it is to use this system	1.86 (0.892)
2. It was simple to use this system	1.97 (1.067)
3. I could effectively complete the tasks and scenarios using this system	1.95 (1.101)
4. I was able to complete the tasks and scenarios quickly using this system	1.97 (1.109)
5. I was able to efficiently complete the tasks and scenarios using the system	1.95 (1.044)
6. I felt comfortable using this system	2.03 (1.136)
7. It was easy to learn to use this system	1.89 (1.125)
8. I believe I could become productive quickly using this system	2.02 (1.130)
9. The system gave error messages that clearly told me how to fix the problems	3.18 (1.300)
10. Whenever I made a mistake using the system, I could recover easily and quickly	2.47 (1.205)
11. The information (such as on-line help, on-screen messages and other documentation) provided with this system was clear	2.30 (1.277)
12. It was easy to find the information I needed	2.26 (1.256)
13. The information provided for the system was easy to understand	2.23 (1.225)
14. The information was effective in helping me complete the tasks and scenarios	2.08 (1.042)
15. The organization of information on the system screens was clear	2.15 (1.218)
16. The interface of this system was pleasant	2.85 (1.765)
17. I liked using the interface of this system	2.68 (1.561)
18. This system has all the functions and capabilities I expect it to have	2.11 (1.083)
19. Overall, I am satisfied with this system	2.15 (1.167)

Interview Results

At the end of the study, the study participants were asked a few open-ended questions to elicit comments on and suggestions for this study, the mobile app itself, security features on mobile apps, their ideas about information security in general, and their source of security knowledge. Their answers are summarized briefly below.

All the study participants welcomed this type of study and expressed an interest in knowing more about security features. They believed the app was very easy to use and the security simulations were easy to understand. Many study participants mentioned that they wished to have a more colorful and graphical user interface. Some older study participants also mentioned the font size, expressing a desire to be able to adjust the font size to meet their needs. Below are some specific comments from the study participants. Please note, as study participants were randomly selected among 238 candidates, the study participants' IDs were numbered 1 to 238:

> *I like that I can see the consequence of those options visually in the encryption part.* [Participant #66]
>
> *Security education is very useful. If I can download the app I will use it again.* [Participant #162]
>
> *Security education, especially the role-based access control is new to me. The app can be even better if it is fancier on interface.* [Participant #36]
>
> *I hope to see more apps like this. It makes me more confident when I am asked to make selections.* [Participant #225]
>
> *The size of the texts is small. I want to be able to control the size of the texts. Button size can be larger as well.* [Participant #38]

All participants said that they knew about the basic security features offered by smart devices, such as a passcode to access a locked screen, before the study. They reported that they expected mHealth apps to protect their PHI, but at the same time, they did not like to enter log-in credentials every time, and only a small number of them were willing to use strong authentication methods. They liked to see the difference between encrypted and not encrypted data, the RBAC, and the outcome of image encryption. The latter two were new for most study participants.

All these study participants knew that information security was a big challenge, and they hoped mobile app developers could offer strong but convenient security protection to protect their PHI. The frequently mentioned sources of security knowledge were family members and friends. All these study participants had discussed information security issues with their family members or friends at different levels. Some were brief chats because of a recent news report, and some were in-depth conversations. Only a few study participants reported having taken security-related classes before this study.

Discussion

Principal Findings

The goal of this project is to demonstrate several commonly used security features, such as encryption, user authentication, and access control, in a mobile app and to determine whether this brief and informal security education is effective in encouraging users to choose stronger security measures to protect their personal health data.

In this study, a number of demographic characteristics were collected; however, almost none of these demographic characteristics made a significant difference on the initial security settings before the security education. In other words, the general population's understanding of information security and their preferences were highly similar before they had security education. Results did show that health status may have an impact on the selection of password update frequency. This could be because people with excellent or very good health status may not have much health information to protect and, therefore, may not feel it is necessary for them to update the password frequently.

Although it is known that providing security education can be useful for mHealth users to choose a stronger security approach to protecting personal health data, it is not feasible to give lectures and classes to every mHealth app user. Therefore, the challenge is to determine an effective and also cost-efficient approach to provide the desired security education. In this study, instead of providing formal security tutorials, we implemented those commonly used security features into the SecSim app and guided the app users to *experience* the difference after they choose different security options. This is the so-called *learning by doing* education approach ("For the things we have to learn before we can do them, we learn by doing them." by Aristotle, The Nicomachean Ethics).

After this brief and informal security education, participants showed significant changes in settings for encryption, RBAC, and image encryption. A significant percentage of study participants chose to use a stronger security measure after the education. In certain cases, demographic characteristics contributed to these changes; for instance, participants with different marital status or years of experience using mobile devices made statistically significant different decisions in using image encryption. A significantly larger percentage of married participants chose to use image encryption after the security education. Participants with more years of experience using mobile devices also tended to use image encryption. Similarly, a larger percentage of people with higher income (>US $100,000) chose to use RBAC. Participants in different employment groups and with different health status also showed statistically significant difference in setting change behavior on encryption after the security education. In other words, this brief and informal security education was effective in encouraging users to choose stronger security protection, which, in turn, may help them to better protect their personal health data, including PHI.

On the other hand, the security education did not produce a similar significant level of change in the log-in credential setting. The difference is that the other security features (eg, encryption and RBAC) are handled by the information system itself; therefore, these features may only have a slight impact on the performance of the information system and do not require the user to do any extra work. However, different selection on the log-in credentials' requirement can dramatically impact a user's experience in the information system. For instance, if they choose to use the 2-factor authentication (username, password, and a PIN), they need to first enter the username and password and manually retrieve a PIN before they can log into the system; however, if everything in the log-in credentials is stored in the system, all they need to do is to click on the icon of the app or URL of a Web portal and they immediately have access to the content of the system. In other words, although the log-in simulation can show them the differences in terms of security when using different authentication approaches, that knowledge may not be able to change their behavior as they need to have a balance between information security and *usability* of a system. If an app is accessed frequently, it becomes tedious and even annoying if complex user authentication is required for each access [20]. Therefore, although mobile app developers may be able to offer highly secure solutions in their apps, users may be reluctant to use the security features or the apps if they have poor usability [15]. Hence, mobile app developers need to use creative approaches to implement highly secure and also highly user-friendly apps.

One may notice that a large percentage of study participants did not make changes to their security settings after the security education (see Table 3). This is not necessarily a bad thing as some study participants had already obtained security knowledge from other sources (such as family members, friends, job training, and classes) as indicated in the summary of the brief interview and, therefore, chose a strong security protection from the beginning (eg, 67% participants wanted image encryption and 50% participants wanted RBAC before the security education). After the security education, 88% participants chose to use image encryption and 80% participants chose to use RBAC (see Table 2). In other words, after the security education, only a small percentage of study participants still wanted to use weaker security protection. For these people, further work is needed, such as a different type of security education.

The results of this study indicate that even a brief and informal security education as shown in this study can be effective and cost-efficient in providing the desired education to mobile app users. The SecSim app itself is small and may be embedded into other mobile apps. Before the app users are required to make decision on their security settings, they can choose to go through the security training offered by SecSim, which takes approximately 10 min. This brief and informal security training may encourage the app users to choose stronger security measures, even if sometimes the app users may have to sacrifice some convenience. If a large number of mHealth apps choose to adopt this embedded training approach (not necessarily using this SecSim app), the number of PHI breaches because of users' (patients and health care providers) reason could be reduced. For instance, if the PHI on a mobile device is encrypted with a

strong encryption algorithm such as Advanced Encryption Standard, even if the mobile device is lost or stolen, the PHI is still protected by the encryption and a data breach will not occur.

Comparison With Prior Studies

As mentioned earlier in the Introduction section, mHealth app users have security and privacy concerns when they use mHealth apps to access or manage their personal health data [5,8,11,21]. However, many of them do not have a clear understanding of the security features offered by mobile operating systems and mobile apps, and therefore, many of them either do not use any security protection methods or do not know which settings are stronger [14,22,23]. As improving security awareness among IT system end users is considered the most cost-effective security control, a number of methods have been created to facilitate such security awareness [25]. However, all these methods, such as posters; newsletter; lectures; Web-based training; and game-based, video-based, and simulation-based training, have mainly been designed for employees in an organization, for instance, health care providers in the health care domain, not the everyday users in the general population [26].

These in-job security training are not sufficient as most mobile app users do not have access to them. The security education described in this study is more generic, and it can be used by both patients and health care providers. This study described the use of an informal, brief, but effective security and privacy training by general users. In terms of the information delivery method, the approach described in this study can be categorized into simulation-based training [26,27].

Limitations

In this study, a large percentage of the study participants were young and highly educated (Bachelor's degree or higher). However, our results showed that the level of education and age did not significantly affect participants' initial selection of security options or their behavior after the brief security education. In other words, a higher education level or younger age does not necessarily mean better understanding of information security. Therefore, the study results may not change dramatically if it is done in study participants with lower education levels or older ages.

In our sample, there were a small number of people with *fair* health condition and no participants with *poor* health. Therefore, the study results may not be applicable to people with poor health conditions. According to the results from other studies, patients with severe diseases, such as heart failure and kidney transplant, do not pay much attention to their privacy but instead pay attention to receiving health care services they need [28,29].

In this study, there were 2 study participants (2/66, 3%) from the rural area. Therefore, the conclusion may not be applicable to people in the rural area. To make conclusions about the populations in the rural areas, a study with more participants from the rural areas is needed.

Although the number of participants in different household income categories was sufficient for the ANOVA analysis, the number of participants in the category with high household income (>US $100,000) was relatively small (n=9). Since the behavior of participants was dramatically different between the 2 income groups (7% switched to stronger protection among 14 participants with incomes between US $10,001 and US $25,000 vs 78% switched to stronger protection among 9 participants with income greater than US $100,000), we do not believe a bigger sample will change the conclusion. On the other hand, a bigger sample would surely make the result more convincing.

When the study participants were recruited for this study, there was no differentiation with respect to occupation. We accepted any mobile app users, whether health care providers or patients. However, as the focus of this study was personal health data protection using security features, the role of these study participants was closer to that of patients, even though the same security knowledge received by health care providers can be applied to patient data protection. A future area of study could be to conduct a similar study but making the settings closer to health care providers' work environment and then recruiting health care providers to be participants.

Conclusions

In this study, a brief and informal security education was delivered to the mobile app users, and their changes in behavior were observed. The results indicated that this simulation-based security education could be helpful for encouraging users to choose stronger security measures to protect their personal health information. In the future, this type of education may be integrated into websites and mobile apps for users to view before they make selections in security settings, which may eventually improve users' information security awareness and reduce the number of PHI data breaches.

Acknowledgments

The authors would like to thank Weiyi He and I Made Agus Setiawan for implementing the first version of the SecSim mobile app and Zara Ambadar for helping to prepare the institutional review board protocol. This study is funded in part by grants from the National Science Foundation (NSF) DGE1438809; the National Institutes of Health (NIH) 1 R43 DC04246-01 and 1 R43 DC04472-01; and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) 90RE5018, 90DP0064, and 90DPGE0002. The contents of this study do not represent the views of the NSF, NIH, or NIDILRR or the US Government.

Conflicts of Interest

None declared.

References

1. US Department of Health Human Services-Office for Civil Rights. Breach portal: Notice to the Secretary of HHS Breach of Unsecured Protected Health Information. 2018. URL: https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf [WebCite Cache ID 6x0wfMpd1]
2. Ponemon Institute. 2017 Cost of Data Breach Study. Armonk, NY: IBM Security; 2017. URL: https://info.resilientsystems.com/hubfs/IBM_Resilient_Branded_Content/White_Papers/2017_Global_CODB_Report_Final.pdf [WebCite Cache ID 73hGvtxfU]
3. PEW Research Center. Mobile Fact Sheet. 2018. URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 6xDIpUN2z]
4. Modahl M. Tablets set to change medical practice executive summary. 2011. URL: https://www.quantiamd.com/q-qcp/QuantiaMD_Research_TabletsSetToChangeMedicalPractice.pdf [WebCite Cache ID 73cIgyDm4]
5. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
6. Atienza AA, Zarcadoolas C, Vaughon W, Hughes P, Patel V, Chou WS, et al. Consumer attitudes and perceptions on mHealth privacy and security: findings from a mixed-methods study. J Health Commun 2015;20(6):673-679. [doi: 10.1080/10810730.2015.1018560] [Medline: 25868685]
7. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
8. Richardson JE, Ancker JS. Public perspectives of mobile phones' effects on healthcare quality and medical data security and privacy: A 2-Year nationwide survey. AMIA Annu Symp Proc 2015;2015:1076-1082 [FREE Full text] [Medline: 26958246]
9. 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]
10. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Dec 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
11. Prasad A, Sorber J, Stablein T, Anthony D, Kotz D. Exposing privacy concerns in mHealth. : USENIX Association; 2011 Presented at: 2nd USENIX conference on Health securityprivacy; 2011; San Francisco, CA p. 2-2.
12. Morera EP, de la Torre Díez I, Garcia-Zapirain B, López-Coronado M, Arambarri J. Security recommendations for mHealth apps: elaboration of a developer's guide. J Med Syst 2016 Jun;40(6):152. [doi: 10.1007/s10916-016-0513-6] [Medline: 27147515]
13. US Department of Health and Human Services - Office of Civil Rights. Breach Portal: Notice to the Secretary of HHS Breach of Unsecured Protected Health Information: Cases currently under investigation. 2017. URL: https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf [WebCite Cache ID 73cIgyDm4]
14. Mylonas A, Kastania A, Gritzalis D. Delegate the smartphone user? Security awareness in smartphone platforms. Comput Secur 2013 May;34:47-66. [doi: 10.1016/j.cose.2012.11.004]
15. Smith KA, Zhou L, Watzlaf VJ. User authentication in smartphones for telehealth. Int J Telerehabil 2017;9(2):3-12 [FREE Full text] [doi: 10.5195/ijt.2017.6226] [Medline: 29238444]
16. Apple. Privacy Policy. 2018. URL: https://www.apple.com/legal/privacy/en-ww/ [WebCite Cache ID 6xDBt8nU9]
17. Chin E, Felt A, Sekar V, Wagner D. Measuring User Confidence in Smartphone Security and Privacy. : ACM; 2012 Presented at: Symposium on Usable Privacy and Security (SOUPS'12); 2012; Washington, DC. [doi: 10.1145/2335356.2335358]
18. Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. Science 2015 Jan 30;347(6221):509-514. [doi: 10.1126/science.aaa1465] [Medline: 25635091]
19. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research, 2nd Ed. Cambridge, MA: Elsevier Inc; 2016.
20. Kate K, Hake J, Ahire S, Shelke H. Authentication of Smartphone Users Using Behavioral Biometrics And OPass Technique. 2017. URL: https://pdfs.semanticscholar.org/0308/44795d1e15979dd0ddfccd4027d7de92f520.pdf [WebCite Cache ID 73hJB2cvs]
21. Atienza AA, Zarcadoolas C, Vaughon W, Hughes P, Patel V, Chou WY, et al. Consumer attitudes and perceptions on mHealth privacy and security: findings from a mixed-methods study. J Health Commun 2015;20(6):673-679. [doi: 10.1080/10810730.2015.1018560] [Medline: 25868685]
22. Kowalski S, Goldstein M. Consumers' awareness of, attitudes towards and adoption of mobile phone security. 2006 Presented at: 20th International Symposium on Human Factors in Telecommunication; 20–23 March, 2006; Sophia Antipolis, France.
23. Ophoff J, Robinson M. Exploring End-User Smartphone Security Awareness within a South African context. : IEEE; 2014 Presented at: 2014 Information Security for South Africa; 13-14 August, 2014; Johannesburg, South Africa. [doi: 10.1109/ISSA.2014.6950500]

24. Parker F, Ophoff J, Van BJ, Karia R. Security awareness and adoption of security controls by smartphone users. : IEEE; 2015 Presented at: Second International Conference on Information Security and Cyber Forensics (InfoSec); 15-17 November, 2015; Cape Town, South Africa. [doi: 10.1109/InfoSec.2015.7435513]
25. Albrechtsen E, Hovden J. Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study. Comput Secur 2010 Jun;29(4):432-445. [doi: 10.1016/j.cose.2009.12.005]
26. Abawajy J. User preference of cyber security awareness delivery methods. Behav Inf Technol 2012 Aug 21;33(3):237-248. [doi: 10.1080/0144929x.2012.708787]
27. Jagatic TN, Johnson NA, Jakobsson M, Menczer F. Social phishing. Commun ACM 2007 Oct 01;50(10):94-100. [doi: 10.1145/1290958.1290968]
28. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012 Feb 10;14(1):e25 [FREE Full text] [doi: 10.2196/jmir.1912] [Medline: 22328237]
29. Browning RB, McGillicuddy JW, Treiber FA, Taber DJ. Kidney transplant recipients' attitudes about using mobile health technology for managing and monitoring medication therapy. J Am Pharm Assoc (2003) 2016;56(4):450-454.e1 [FREE Full text] [doi: 10.1016/j.japh.2016.03.017] [Medline: 27450140]

Abbreviations

ANOVA: analysis of variance
IT: information technology
mHealth: mobile health
PHI: Protected Health Information
PIN: personal identification number
PL: protection level
PSSUQ: Post-Study System Usability Questionnaire
RBAC: role-based access control
TLS: Transport Layer Security

Edited by G Eysenbach; submitted 01.06.18; peer-reviewed by D Klein, M Abdelhamid, S Albakri; comments to author 19.07.18; revised version received 12.09.18; accepted 23.09.18; published 11.12.18.

Please cite as:
Zhou L, Parmanto B, Alfikri Z, Bao J
A Mobile App for Assisting Users to Make Informed Selections in Security Settings for Protecting Personal Health Data: Development and Feasibility Study
JMIR Mhealth Uhealth 2018;6(12):e11210
URL: https://mhealth.jmir.org/2018/12/e11210/
doi:10.2196/11210
PMID:30538088

Original Paper

Identifying Evidence-Informed Physical Activity Apps: Content Analysis

Mihiretu Kebede[1,2,3], MPH; Berit Steenbock[1,4], MSc; Stefanie Maria Helmer[1,5], PhD; Janna Sill[1,2], MSc; Tobias Möllers[6], MSc; Claudia R Pischke[1,7], PhD

[1]Prevention and Evaluation, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

[2]Health Sciences, University of Bremen, Bremen, Germany

[3]Institute of Public Health, University of Gondar, Gondar, Ethiopia

[4]Center of Clinical Psychology and Rehabilitation, University of Bremen, Bremen, Germany

[5]Institute for Social Medicine, Epidemiology, and Health Economics, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

[6]Network Aging Research, University of Heidelberg, Heidelberg, Germany

[7]Institute of Medical Sociology, Centre for Health and Society, Medical Faculty, University of Duesseldorf, Duesseldorf, Germany

Corresponding Author:
Claudia R Pischke, PhD
Institute of Medical Sociology
Centre for Health and Society, Medical Faculty
University of Duesseldorf
Universitätsstraße 1
Duesseldorf, 40225
Germany
Phone: 49 221 81 08599
Fax: 49 211 81 08614
Email: ClaudiaRuth.Pischke@med.uni-duesseldorf.de

Abstract

Background: Regular moderate to vigorous physical activity is essential for maintaining health and preventing the onset of chronic diseases. Both global rates of smartphone ownership and the market for physical activity and fitness apps have grown rapidly in recent years. The use of physical activity and fitness apps may assist the general population in reaching evidence-based physical activity recommendations. However, it remains unclear whether there are evidence-informed physical activity apps and whether behavior change techniques (BCTs) previously identified as effective for physical activity promotion are used in these apps.

Objective: This study aimed to identify English and German evidence-informed physical activity apps and BCT employment in those apps.

Methods: We identified apps in a systematic search using 25 predefined search terms in the Google Play Store. Two reviewers independently screened the descriptions of apps and screenshots applying predefined inclusion and exclusion criteria. Apps were included if (1) their description contained information about physical activity promotion; (2) they were in English or German; (3) physical activity recommendations of the World Health Organization or the American College of Sports Medicine were mentioned; and (4) any kind of objective physical activity measurement was included. Two researchers downloaded and tested apps matching the inclusion criteria for 2 weeks and coded their content using the Behavioral Change Technique Taxonomy v1 (BCTTv1).

Results: The initial screening in the Google Play Store yielded 6018 apps, 4108 of which were not focused on physical activity and were not in German or English. The descriptions of 1216 apps were further screened for eligibility. Duplicate apps and light versions (n=694) and those with no objective measurement of physical activity, requiring additional equipment, or not outlining any physical activity guideline in their description (n=1184) were excluded. Of the remaining 32 apps, 4 were no longer available at the time of the download. Hence, 28 apps were downloaded and tested; of these apps, 14 did not contain any physical activity guideline as an app feature, despite mentioning it in the description, 5 had technical problems, and 3 did not provide objective physical activity measurement. Thus, 6 were included in the final analyses. Of 93 individual BCTs of the BCTTv1, on average,

9 (SD 5) were identified in these apps. Of 16 hierarchical clusters, on average, 5 (SD 3) were addressed. Only BCTs of the 2 hierarchical clusters "goals and planning" and "feedback and monitoring" were identified in all apps.

Conclusions: Despite the availability of several thousand physical activity and fitness apps for Android platforms, very few addressed evidence-based physical activity guidelines and provided objective physical activity measurement. Furthermore, available descriptions did not accurately reflect the app content and only a few evidence-informed physical activity apps incorporated several BCTs. Future apps should address evidence-based physical activity guidelines and a greater scope of BCTs to further increase their potential impact for physical activity promotion.

(JMIR Mhealth Uhealth 2018;6(12):e10314) doi:10.2196/10314

KEYWORDS

guidelines; mHealth; mobile apps; physical activity; mobile phone

Introduction

Smartphone ownership among adults, including older adults, has rapidly increased worldwide [1]. In Germany, 97% of adults aged 30-49 years, 88% of adults aged 50-64 years, and 41% of those aged ≥65 years reported owning and using a smartphone in 2017 [2]. In tandem with the peak of smartphone ownership, there is an increase in consumer interest in physical activity (PA) measurement assisting individuals in recording their day-to-day activities. The results of a population-based survey conducted in Germany suggested that among German smartphone users (n=4144; age 57 years, SD 14), approximately 21% use health apps to change health behavior, including PA, or to reach certain health outcomes, such as weight loss [3]. Although this analysis is not based on a representative sample, these findings indicate that apps may represent an important vehicle for implementing population-based strategies aimed at health behavior change, including the promotion of regular PA in Germany. This takes on added significance considering the notable demographic change in Germany. Compared with other European countries, Germany is faced with an aging population [4]. Hence, the use of health and fitness apps may facilitate the uptake and maintenance of health and PA behavior [5] and may contribute to healthy aging. The use of such apps may assist older adults in maintaining muscular and cardiorespiratory fitness and bone and functional health [6-8].

In 2015, the Preventive Health Care Act was passed in Germany, which mandates health insurances and long-term insurance funds to invest >500 million euros in health promotion and primary prevention in the coming years [9]. As a result, health insurances have increased their efforts to offer apps for health promotion to their clients. So far, apps for stress reduction, smoking cessation, and making dietary changes have been made available to insurance holders [10]. In some cases, insurance agencies in Germany have even developed apps themselves (eg, fit mit AOK, Generali Vitality, and BARMER App FIT2GO) [10]. Similarly, the market for commercially available health and fitness apps is booming. In 2016, approximately 259,000 mobile health (mHealth) apps (ie, those listed in the medical and health and fitness app category of an app store) were available in major app stores. Google Play (Android) currently displays 97,345 mHealth apps, including apps from both health and fitness and medical categories [11]. It is estimated that in 2020, approximately 2.6 billion app users will have downloaded mHealth apps at least once and 551 million of these app users will actively use the apps [12].

However, it is unclear whether the apps recommended by health insurances or those commercially available are based on existing evidence stemming from research identifying effective intervention components or mechanisms for health behavior change. A growing body of research is examining whether the content, particularly that of PA apps, is based on current evidence on the underlying mechanisms for behavior change. This research suggests that only a few PA apps are evidence informed and address current guidelines for aerobic activity and strength and resistance and flexibility training [7,13,14]. This is a major shortcoming considering that health benefits associated with PA can only be obtained when these recommendations are reached and that particularly older adults (age≥60 years) rarely meet these recommendations [15]. In Germany, only 22% of adults aged ≥45 years meet the current PA recommendations [16]. Thus, making information on PA recommendations salient in PA and fitness apps or designing features around facilitating weekly moderate PA of 150 minutes recommended by the World Health Organization (WHO) are necessary steps to support particularly older users in starting or developing a PA routine [13,14].

In addition, it is unknown which behavior change techniques (BCTs) are used in apps and whether their use is associated with increased behavior change among users. Harries et al [17] suggested that feedback on a person's personal PA level is itself sufficient to prompt increased walking; in their study, participants who wore an always-on, accelerometer-based smartphone app experienced a substantial increase in walking. In addition, several content analyses have been conducted, predominantly in the Netherlands and the United Kingdom, to identify active components of various types of apps, including PA apps [18,19]. Several content analyses used the Behavioral Change Technique Taxonomy v1 (BCTTv1), a comprehensive and reliable tool for assisting researchers in retrospectively identifying active components of interventions, particularly behavioral interventions. It includes 93 BCTs considered to be effective for behavior change and 16 hierarchical clusters [20]. Middelweerd et al [19] analyzed PA apps, using an earlier taxonomy developed by Michie et al [21], and found that, on average, 5 (range, 2-8) of 23 possible BCTs were used in the reviewed apps. The most frequently used BCTs were "feedback on performance" and "goal setting", whereas other BCTs of the taxonomy were not identified. In another content analysis, apps

for medication adherence were examined using the BCTTv1 [22]. Here, the number of BCTs contained in an app ranged from 0 to 7 out of 96 possible BCTs, and the most commonly used BCTs were "action planning" and "prompt/cues," which were included in 96% (160/166) of the total of 166 medication adherence apps investigated [22].

In sum, there are many PA and fitness apps commercially available, as well as an increasing number of apps made available by health insurances. However, there is still a lack of research on whether these apps are based on evidence-based PA guidelines and which BCTs are employed. To date, no content analysis evaluating the entire range of BCTs for evidence-informed PA (EIPA) apps with objective PA tracking available on the German market has been published. This is a major shortcoming considering the increased focus on population-based strategies for PA promotion for primary prevention in Germany owing to the Preventive Health Care Act and the need for low-threshold electronic health (eHealth) interventions, including PA apps, which can assist the general population in increasing PA. Hence, this study aims to identify EIPA apps and BCTs employed in both German and English PA and fitness apps, using the BCTTv1 taxonomy.

Methods

Definition of Evidence-Informed Physical Activity Apps

The global PA recommendation of the WHO for adults is to engage in 150 minutes of moderate-intensity aerobic PA throughout the week or at least 75 minutes of vigorous-intensity aerobic PA or an equivalent combination of moderate-intensity and vigorous-intensity activity [8]. Furthermore, aerobic activity should be performed in bouts of at least 10 minutes, and whole-body strength training activities for major muscle groups on at least 2 days per week are recommended. The American College of Sports Medicine (ACSM) outlines that at least 10,000 steps per day are needed for adults [7]. These are the two most commonly used guidelines for designing evidence-based PA interventions, including eHealth interventions for PA promotion (Eysenbach defines eHealth as an "intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the internet and related technologies" (pg 1) [23]). Therefore, a PA app designed following any of these guidelines is considered as an EIPA app [14].

Identification of Physical Activity and Fitness Apps

Apps were identified in a comprehensive systematic search in the Google Play Store. The search took place between August 3 and October 6, 2015. We used 25 search terms in German and English to search across all categories in the Google Play Store (ie, Bewegung, Sport, Aktivität, Übungen, Training, Laufen, Gehen, Joggen, Sportliche Aktivität, Fitness, Gleichgewicht, Kräftigung, move, sports, exercise, activity, exercises, workout, walk, run, step, jogging, physical activity, balance, and strengthening). Two reviewers independently screened all results available in the Google Play Store for each single search term during the day of the search. In cases where searches were performed on different days and there were fewer or more apps because of updates, only results of the later search were included.

Screening Procedure

Because of the comprehensiveness of our search, we divided the screening procedure into the following three different steps. The first step was to identify PA and fitness apps; each search term was entered in the Google Play Store. The names and descriptions of the apps were reviewed based on *a priori* defined inclusion and exclusion criteria. Apps were included if the app description contained information about PA promotion and if they were in English or German. Conversely, they were excluded if the content was focused on topics, such as allergies, babies, nutrition, hypnosis, smoking, pregnancy, and stress, or if they were in a language other than English and German. The name, description, number of installations, and the price of the included apps were extracted into an excel sheet (Multimedia Appendix 1). In case of any discrepancies between the 2 reviewers while applying the inclusion and exclusion criteria, the consensus was reached after discussing with a third reviewer. In the second step, duplicates of apps were removed manually. In the third step, app descriptions were screened, according to the following additional inclusion and exclusion criteria: Inclusion criteria—(1) the app description contained any of the recommendations of the ACSM or the WHO and (2) the app description contained information about an objective assessment of PA (eg, step count, distance in kilometers, or minutes of being active measured through mobile phone's built-in acceleration sensor or global positioning system); exclusion criteria—(1) a specific group other than the user was targeted (eg, gym owners); (2) the app only focused on training of particular areas of the body and did not target the whole body; (3) necessitated external devices for use (eg, heart rate monitor and accelerometer-based activity monitor). If a full and light version of the same app was available, the light version was excluded. When necessary, we paid for the full version of the app.

Testing Phase of Evidence-Informed Physical Activity Apps and Behavioral Change Technique Rating Procedure

Apps that met all inclusion criteria were downloaded, installed, and tested on different smartphones running Android operating systems (Samsung Galaxy S6 Edge and Samsung Galaxy S5) by 5 raters from February to May 2017. The content evaluation of the apps was based on the BCTTv1, and BCTs were independently identified by 2 trained raters using the taxonomy [20]. The raters ran all apps for at least 2 weeks on their smartphone to check all features of the apps and extract data on app characteristics (eg, user rating, download rates, and language) and additional functionalities. If an app outlined PA recommendations in the description but none of the recommendations were found in the features of the app during the 2-week testing phase, the app was excluded from the content analysis. After the end of the testing period, the 2 raters independently coded BCTs for each of the remaining EIPA apps. Of note, the results of this study are solely based on the content of each EIPA app. No additional information about the

apps was collected from websites of the developers. The information collected and the BCTs coded by the 2 raters were discussed with a third researcher to solve any discrepancies.

Behavioral Change Technique Taxonomy v1

The BCTTv1 is a tool for assisting researchers in retrospectively identifying effective components of behavioral interventions, including eHealth interventions such as Web-based interventions or smartphone apps. It includes 93 BCTs considered to be effective for behavior change, which are organized into 16 hierarchical clusters [20]. The BCTTv1 has been validated and is used to design and retrospectively evaluate the effects of behavioral health interventions [24].

Data Analysis

We calculated the interrater reliability between the 2 raters using the commonly used interrater agreement indices: Cohen kappa and prevalence-adjusted and bias-adjusted kappa. Descriptive statistics were used to analyze the number of BCTs addressed in the examined apps. Data analysis was performed using IBM SPSS Statistics for Windows, version 24.0 (IBM Corp).

Results

Figure 1 outlines the entire search process. The initial screening in the Google Play Store yielded 6018 apps (screening phase: August 3, 2015 to October 6, 2015). After eliminating ineligible apps (n=4108) and removing duplicates and light versions (n=694), 1216 apps passed the first assessment with regard to mentioning a PA guideline and objective PA measurement in their description. Briefly, 2.6% (32/1216) apps mentioned a PA guideline and an objective PA measurement in their description. Four apps were no longer available for download when the testing phase started (date of availability check: February 13, 2017). After downloading and testing the remaining 28 apps, 14 did not contain any of the PA guidelines as a feature of the app, despite originally mentioning them in their description. Another 5 of 28 apps had technical problems and at least 2 different types of mobile phones failed to run the apps, and 3 did not provide objective PA measurement. In addition, 0.5% (6/1216) apps addressed PA guidelines, provided objective PA tracking, and had no technical problems—"Pacer Health/Schrittzähler & Abnehm Trainer," "Health Mate," "Lark Chat," "The Walk: Fitness Tracker Game," "Step Counter," and "Pedometer." Multimedia Appendix 2 provides screenshots of the selected apps.

Table 1 shows the characteristics of the 6 EIPA apps included in the content analysis. During the coding of the app content, all EIPA apps had >3.5 stars of user ratings. The highest user rating and download rate were noted for "Pacer Health/Schrittzähler & Abnehm Trainer." "Health Mate," "Lark Chat," and "The Walk: Fitness Tracker Game" are English language apps, while "Pacer Health/Schrittzähler & Abnehm Trainer" is available in both English and German. "Step counter" and "Pedometer" provide Google translated versions for users. Except for "The Walk: Fitness Tracker Game," all the apps were free of charge. For PA tracking, half of the apps ("Pedometer," "The Walk: Fitness Tracker Game," and "Health Mate") used acceleration sensors, and the other half ("Step Counter," "Pacer Health/Schrittzähler & Abnehm Trainer," and "Lark Chat") used a combination of accelerometer sensors and global positioning system. Of the 6 apps, 4 (ie, "Pedometer," "Step counter," "Pacer Health/Schrittzähler & Abnehm Trainer," and "Health Mate") contained the recommendation to walk 10,000 steps per day. "The Walk: Fitness Tracker Game" and "Lark Chat" were based on the PA guideline of the WHO to engage in moderate PA for at least 30 minutes per day. Furthermore, 2 apps (ie, "Health Mate" and "Pacer Health/Schrittzähler & Abnehm Trainer") allowed users to connect their PA tracking data with other PA tracking apps such as "S Health."

Multimedia Appendix 3 includes the definitions of the BCTs identified in the examined apps, as well as examples of how these BCTs were used in the apps.

Table 2 presents the interrater reliability between raters. The maximum agreement statistics were noted for Pedometer achieving an 89.2% prevalence-adjusted and bias-adjusted kappa and a kappa value of .42. The 2 raters had the lowest agreement in rating "Lark Chat."

Multimedia Appendix 4 presents the type and the total number of BCTs rated for each EIPA app. Of 93 individual BCTs included in the BCTTv1, 29% (27/93) were found in the examined apps. On average, 9 (SD 5.06) were identified in apps (minimum: 4 and maximum: 18). In addition, 75% (12/16) of 16 hierarchical clusters of the BCTTv1 taxonomy were identified in the included EIPA apps. "Discrepancy between current behavior and goal standard," "feedback on behavior," "goal setting (behavior)," and "self-monitoring of behavior" were the most frequently included BCTs. On average, 5 (SD 3.07) hierarchical clusters were addressed in the EIPA apps. Only 2 BCTs (ie, "goals and planning" and "feedback and monitoring") were included in all apps. The median number of BCT hierarchical clusters and BCTs included in the app was 4.5 (range 2-9) and 7.5 (range 4-18), respectively. In "Lark Chat" followed by "Health Mate," the highest number of individual BCTs, as well as hierarchical clusters, was addressed. The maximum number of BCTs included in an EIPA app was identified for "Lark Chat" with 18 BCTs identified. The minimum number of BCTs was coded for "Pedometer," with only 4 BCTs coded. Moreover, a maximum of 10 and a minimum of 2 hierarchical clusters were addressed in the apps included in this study. Of 16 hierarchical clusters of the taxonomy, only BCTs of 2 clusters (ie, "goals and planning" and "feedback and monitoring") were identified in all of the apps included in this study. BCT hierarchical clusters namely "scheduled consequences" and "social support" were only identified in "Lark Chat" and "Health Mate," respectively. Furthermore, 5 hierarchical clusters, namely "associations," "comparison of behavior," "comparison of outcomes," "social support," and "scheduled consequences" were included only once.

Figure 1. Overview of steps for selecting physical activity (PA) and fitness apps. BCTTv1: Behavioral Change Technique Taxonomy v1.

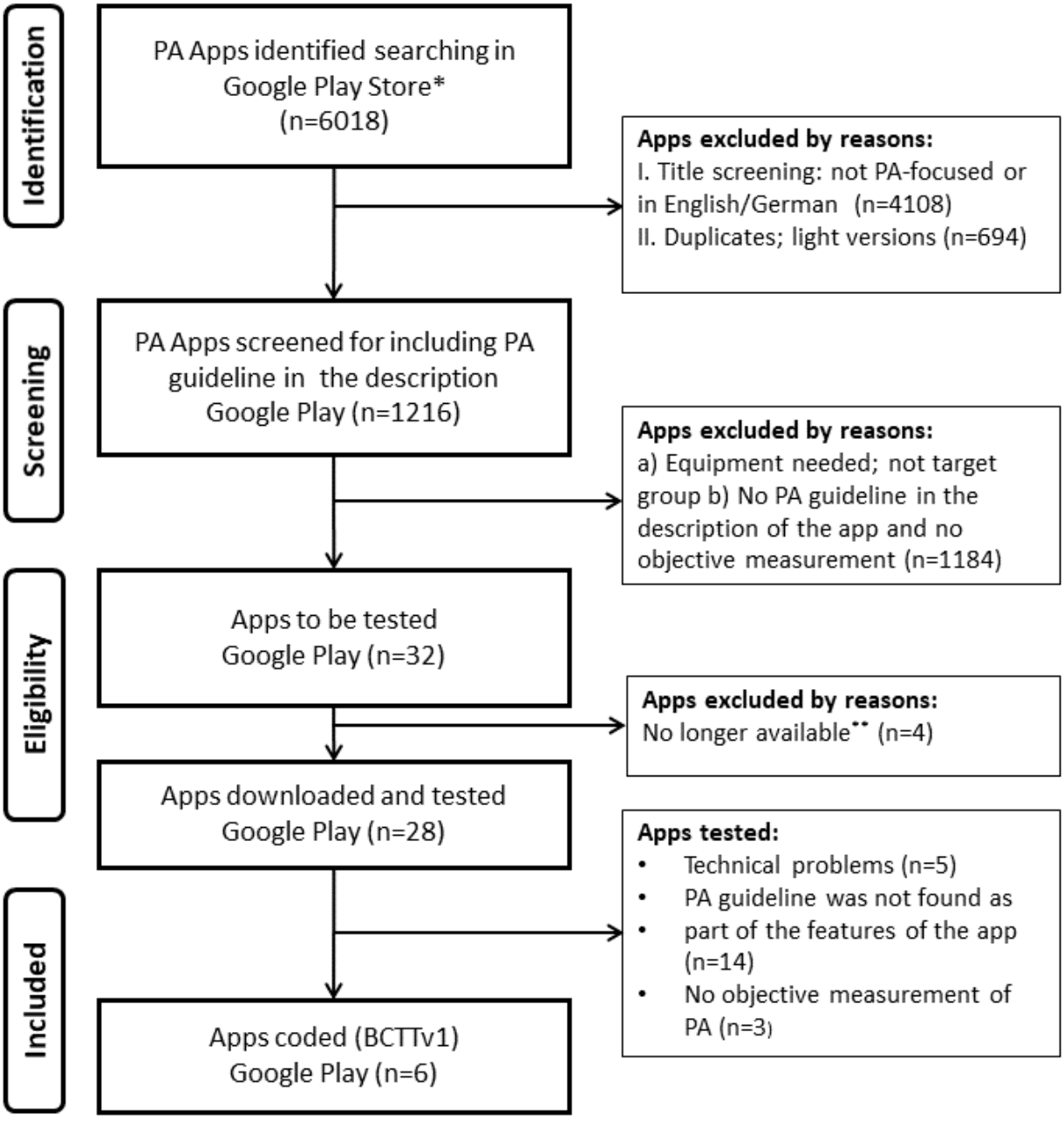


Table 1. Characteristics of physical activity and fitness apps.

Name of app	Developer(s)	Price in €	App store user rating in stars	Language	Number of app store raters	Number of downloads	Additional features
Pedometer	Ivon Liu	Free	4.0	English	470	10,000-50,000	Measures distance covered and calorie calculator
The Walk: Fitness Tracker Game	Six to Start (dev. with NHS and the UK's Department of Health)	3.39	3.7	English	1139	10,000-50,000	Google fit step tracking
Step Counter (Schrittzähler)	xstep	Free	3.5	English	363	100,000-500,000	Distance, speed, calories, heart rate, and elapsed time
Pacer (eng)/Schrittzähler & Abnehm Trainer (de)	Pacer Health	Free	4.5	English and German	330,283	10,000,000-50,000,000	Sleep tracking, nutrition (eg, water intake, vegetable intake, alcohol), pregnancy, synchronization with my fitness pal or S Health, weight
Health Mate	Withings	Free	3.6	English and German	15,222	1,000,000-5,000,000	Weight monitoring, other activities, heart rate, blood pressure, and food diary
Lark Chat	Stanford and Harvard health	Free	4.1	English	2702	50,000-100,000	Sleep tracking functionality, track other activities, such as biking, weight monitoring, and food logging

Table 2. Interrater reliability scores.

Name of app	Kappa	Prevalence-adjusted and bias-adjusted kappa (%)
Pedometer	.417	89.2
The Walk Fitness Tracker Game	.075	72.0
Step Counter	.332	85.0
Schrittzähler & Abnehm Trainer	.240	86.0
Health Mate	.285	72.0
Lark Chat	.031	52.6

Discussion

Principal Findings

This study aimed to identify EIPA apps and BCTs employed in these apps. The results revealed that <1% of the examined 1216 apps mentioned PA guidelines, an objective PA measurement, and worked properly. Regarding BCTs included in those EIPA apps, approximately one-third of the BCTs outlined in the BCTTv1 were used. Moreover, BCTs of 75% (12/16) of the 16 hierarchical clusters of the taxonomy were identified, the overarching clusters “goals and planning” and “feedback and monitoring” were included in all of them.

Comparison With Prior Studies

In comparison to our results, a review of 2400 PA apps conducted by Knight et al reported that none of the examined apps were based on evidence-based guidelines for aerobic PA, and only approximately 2% (8/379) of 379 apps deemed eligible were implementing evidence-based guidelines for resistance training [14]. Modave et al reviewed 30 popular PA apps and reported that only 3 apps reflected parts of the guidelines set forth by the ACSM [13]. Our search may have yielded different results because we searched the German Google Play Store, our search strategy was different, and the search was conducted during a later point in time. The updated versions of apps or newly developed apps may increasingly address PA guidelines and incorporate a larger number of BCTs.

Taking a somewhat similar approach to our study, Direito et al downloaded the top-20 paid and top-20 free PA and dietary behavior apps from the New Zealand Apple App Store Health and Fitness category and coded each app for the presence or absence of BCTs [25]. They coded approximately 20% of BCTs from the BCTTv1 compared with 30% in our content analysis. Similar to our study, they found, on average, 8 BCTs (range 2-18) in these apps. Furthermore, they found that paid apps included more BCTs [25]. Whether there was a statistically significant difference between paid apps and those free of charge regarding the number of BCTs was not investigated in this study. The most commonly identified BCTs in Direito et al’s study were “provide instruction,” “set graded tasks,” and “prompt

self-monitoring" [25].This is contrary to the BCTs most commonly found in our study (ie, "discrepancy between current behavior and goal standard," "feedback on behavior," "goal setting (behavior)," and "self-monitoring of behavior"). However, this discrepancy of results may be attributed to the differences in the scope of the 2 searches. Direito et al only examined the most downloaded apps [25]. Results comparable to our results were obtained in an extensive content analysis conducted by Middelweerd et al based on an earlier version of the taxonomy [19].

In another review of weight management apps, Bardus et al rated the app quality and content of the most popular health and fitness apps on Google Play and iTunes to determine the number of BCTs included [26]. In their study, 10 techniques were identified per app (range 1-17) and approximately 37% of BCTs from the BCTTv1 were applied with "goal setting" and "self-monitoring" among the most frequently identified. In addition, Bardus et al found that the number of BCTs included correlated with the app quality and the number of different technical features of apps [26]. Unfortunately, the app quality was not assessed in our study. Furthermore, which combination of BCTs is most effective in changing PA behavior and whether there is an association between including a higher number of BCTs in PA apps and changes in PA remain questions to be addressed in future studies. There is some indication of a review of studies examining the effects of interventions targeting healthy eating and PA that a combination of self-monitoring with at least one other technique derived from control theory was more effective in promoting behavior change than other single-technique interventions [27]. In addition, different BCTs may be effective for promoting short-term versus long-term changes in health behavior. Samdal et al demonstrated that "goal setting" and "self-monitoring of behavior" were associated with both short-term and long-term changes in healthy eating and PA in overweight and obese adults, whereas several other BCTs (eg, "goal setting of outcome," "feedback on outcome of behavior," "implementing graded tasks," "adding objects to the environment," such as step counters) predicted behavior change in the long term [28].

User ratings of >3.5 stars were noted for all EIPA apps examined in this study. The relationship between user rating and the number of BCTs included in apps is, however, still unclear and requires further investigation. In addition, it remains unknown whether EIPA apps include more BCTs than generic PA apps and therefore receive higher user ratings. The findings of a previous study indicated that user ratings positively correlate with the number of features included in PA apps [29]. However, a subgroup analysis on Google Play versus iTunes apps performed by Mollee et al to evaluate their potential for increasing PA yielded contradictory associations of user ratings and the number of features for Android versus iTunes PA apps [29]. User ratings were not associated with the number of features for Android PA apps [29], while in another study, a 15% increment of user ratings was noted for each additional BCT included in iTunes PA apps [30]. However, these associations may be misleading because user ratings can be easily manipulated. It has previously been reported that app developers can recruit users with as low as US $5 to negatively review or badly rate apps developed by their competitors [31,32]. Nevertheless, future studies should explore the interrelationships between user ratings, number of BCTs, additional features, price, PA measurement accuracy, and effectiveness of apps for PA promotion further. In addition, new tools, such as the Mobile App Rating Scale, may assist researchers in determining the quality of apps [33]. For example, differences in the accuracy of measuring distance were noted by Pobiruchin et al [34]. Hence, PA measurement accuracy of PA apps regarding objective indicators (eg, distance covered, steps counted, and timing of exercises) needs further evaluation and calibration, using gold standards.

Limitations and Strengths

This study has several limitations. First, the study was limited to Android PA apps. The inclusion of iTunes PA apps might have produced different results. Second, one problem encountered during the search was that the screening process took a long time because the search was not limited to the most popular or downloaded PA apps. Some of the apps were no longer available, or updates were available when the testing phase was reached. The restricted search strategies followed in the studies outlined above may have prevented our search from being outdated by the time the content was coded in detail. However, we ensured that the apps tested for 2 weeks were still available at the beginning of the testing phase. In some cases, there were updated versions, which had been further developed by the same or a different company. Another limitation was that the maximum number of apps available per search term was limited to 250 apps and the underlying algorithm for this limitation was unclear. Considering the rapid development and release of PA apps, a new search would produce very different findings. In addition, it remains unclear whether searches in other regions of Europe or the world would produce similar results. Therefore, the generalizability of the results is limited. Another issue encountered during the search was that app descriptions were in many cases different from the functions offered in the apps, resulting in a retrospective exclusion of apps during the testing phase. We may have excluded apps that did not explicitly mention PA guidelines in their descriptions but were guideline-informed. However, the strength of our study was the relatively high interrater reliability for identifying BCTs in the final selection of apps suggesting that raters were well versed in the use of the taxonomy.

In sum, our results are in line with existing research indicating that only a limited number of BCTs is currently included in such interventions despite growing evidence suggesting that the effectiveness of digital health interventions can be enhanced by incorporating BCTs [27,28]. In addition, the existing evidence suggests that the theoretical constructs of BCTs are only rarely considered during app development [18,19,35,36]. Hence, there appears to be a need for collaboration between PA app developers and public health, health promotion, and behavior change experts [35,37].

Conclusions

To conclude, this study indicates that despite the availability of several thousand PA apps for Android platforms, very few of them are evidence informed and simultaneously provide

objective PA measurement. In addition, only a few of them incorporate a large number of BCTs. Future apps should address evidence-based PA guidelines and a greater scope of BCTs to further increase their potential impact for PA promotion in the general population. Furthermore, it is important that researchers make recommendations regarding the use of EIPA apps in the general population or advise health insurances in selecting and disseminating the EIPA apps identified in this study to insurance holders as opposed to representatives of entities with commercial interests. The widespread use of EIPA apps may boost other population-based strategies for PA promotion for primary prevention in Germany, which are currently ongoing as a result of the Preventive Health Care Act.

Acknowledgments

We thank Max Julius Hassenstein for assisting us with creating the content for several tables. The study was funded by an internal grant for innovative research of the Leibniz-Institute for Prevention Research and Epidemiology – BIPS.

Authors' Contributions

MK has substantially contributed to the extraction of data for the apps, data analysis, and the composition of the manuscript. BS was involved in the conception and design of the study, conducted the app search and screening, extracted data for the apps, and participated in drafting the manuscript. SMH was involved in the conception and design of the study, conducted the app search and screening, extracted data for the apps, and critically revised the manuscript. JS was involved in the abstract and title screening, extraction of the data, and in the drafting of the manuscript. TM was involved in the app search and screening, extraction of the data for the apps, and critically revised the manuscript. CRP conceived the study, contributed to the extraction of the data, and wrote and critically revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results of the testing phase (n=32 apps).

[XLSX File (Microsoft Excel File), 18KB - mhealth_v6i12e10314_app1.xlsx]

Multimedia Appendix 2

Screenshots of the included apps.

[PDF File (Adobe PDF File), 917KB - mhealth_v6i12e10314_app2.pdf]

Multimedia Appendix 3

Definitions of behavioral change techniques (BCTs) addressed in the included evidence-informed physical activity apps and examples for application of BCTs in apps.

[PDF File (Adobe PDF File), 38KB - mhealth_v6i12e10314_app3.pdf]

Multimedia Appendix 4

Individual behavior change techniques and hierarchical clusters addressed in the evidence-informed physical activity apps included in the content analysis.

[PDF File (Adobe PDF File), 47KB - mhealth_v6i12e10314_app4.pdf]

References

1. eMarketer. Statista - The Statistics Portal. Statista; 2017. Number of smartphone users worldwide from 2014 to 2020 (in billions) URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [WebCite Cache ID 6wMYs54Pq]
2. Statista Das Statistik-Portal. Festnetz & Mobilfunk. Statista; 2018. Anteil der Smartphone-Nutzer in Deutschland nach Altersgruppe im Jahr 2017 URL: https://de.statista.com/statistik/daten/studie/459963/umfrage/anteil-der-smartphone-nutzer-in-deutschland-nach-altersgruppe/ [WebCite Cache ID 6wMZa0ZNS]
3. 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]

4. Frank S. Demographic Change in Germany and Europe.: Bundeszentrale für politische Bildung; 2014 Jan 28. URL: http://www.bpb.de/gesellschaft/migration/kurzdossiers/177962/germany-and-europe [accessed 2018-10-29] [WebCite Cache ID 6wMaJRkDA]
5. Hoj TH, Covey EL, Jones AC, Haines AC, Hall PC, Crookston BT, et al. How Do Apps Work? An Analysis of Physical Activity App Users' Perceptions of Behavior Change Mechanisms. JMIR Mhealth Uhealth 2017 Aug 03;5(8):e114 [FREE Full text] [doi: 10.2196/mhealth.7206] [Medline: 28778846]
6. Payne HE, Moxley VB, MacDonald E. Health Behavior Theory in Physical Activity Game Apps: A Content Analysis. JMIR Serious Games 2015;3(2):e4 [FREE Full text] [doi: 10.2196/games.4187] [Medline: 26168926]
7. Linda S. P, Ross A, Deborah R, Paul D. T. ACSM's Guidelines for Exercise Testing and Prescription 9th Ed. 2014. In: The Journal of the Canadian Chiropractic Association. Philadelphia: Lippincott Williams & Wilkins; 2014.
8. World Health Organization. Global recommendations on physical activity for health. Geneva, Switzerland: WHO Press; 2010. URL: http://www.who.int/sorry/ [accessed 2018-11-21] [WebCite Cache ID 746NcpXET]
9. Federal Ministry of Health. Promoting preventive healthcare. Berlin, Germany; 2015. Preventive Health Care Act URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/Publikationen/Ministerium/Broschueren/BMG_Hausbroschuere_2016_EN.pdf [accessed 2018-10-29] [WebCite Cache ID 6wMaxyIIW]
10. Stiftung Warentest. Versicherungen. 2016 Jul 19. Gesundheits-Apps der Krankenkassen: Prävention ja, Therapie nein URL: https://www.test.de/Gesundheits-Apps-Die-richtige-finden-aber-wie-5009792-0/ [accessed 2018-10-29] [WebCite Cache ID 6wMb6cyKX]
11. AppBrain. Top Categories. Most popular Google Play categories online: AppBrain; 2018 Oct 29. Most popular Google Play categories URL: http://www.appbrain.com/stats/android-market-app-categories [accessed 2018-10-29] [WebCite Cache ID 6wMbKQMEd]
12. research2guidance. The current status and trends of the mHealth app market. Berlin, Germany: R2G; 2016. mHealth App Developer Economics 2016 URL: https://research2guidance.com/r2g/r2g-mHealth-App-Developer-Economics-2016.pdf [accessed 2018-10-29] [WebCite Cache ID 73XRBP7rA]
13. Modave F, Bian J, Leavitt T, Bromwell J, Harris IC, Vincent H. Low Quality of Free Coaching Apps With Respect to the American College of Sports Medicine Guidelines: A Review of Current Mobile Apps. JMIR Mhealth Uhealth 2015;3(3):e77 [FREE Full text] [doi: 10.2196/mhealth.4669] [Medline: 26209109]
14. Knight E, Stuckey MI, Prapavessis H, Petrella RJ. Public health guidelines for physical activity: is there an app for that? A review of android and apple app stores. JMIR Mhealth Uhealth 2015;3(2):e43 [FREE Full text] [doi: 10.2196/mhealth.4003] [Medline: 25998158]
15. Sun F, Norman IJ, While AE. Physical activity in older people: a systematic review. BMC Public Health 2013;13:449 [FREE Full text] [doi: 10.1186/1471-2458-13-449] [Medline: 23648225]
16. Finger J, Mensink G, Lange C, Manz K. Gesundheitsfördernde körperliche Aktivität in der Freizeit bei Erwachsenen in Deutschland. Journal of Health Monitoring 2017;2(2):37-44 DOI: 10.17886/RKI-GBE-2017-027. [doi: 10.17886/RKI-GBE-2017-027]
17. 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 Dec 02;16:925 [FREE Full text] [doi: 10.1186/s12889-016-3593-9] [Medline: 27590255]
18. 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]
19. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
20. 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]
21. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
22. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior Change Techniques in Apps for Medication Adherence: A Content Analysis. Am J Prev Med 2016 May;50(5):e143-e146. [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
23. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
24. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188 [FREE Full text] [doi: 10.3310/hta19990] [Medline: 26616119]

25. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
26. 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;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
27. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
28. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. Int J Behav Nutr Phys Act 2017 Dec 28;14(1):42 [FREE Full text] [doi: 10.1186/s12966-017-0494-y] [Medline: 28351367]
29. Mollee J, Middelweerd A, Kurvers R, Klein M. What technological features are used in smartphone apps that promote physical activity? A review and content analysis. Personal and Ubiquitous Computing 2017;21(4):633-643 [FREE Full text]
30. Bondaronek P, Murray E, Hamilton F. The Public Health Potential of the Current Health Apps for Increasing Physical Activity. In: In Proceedings of the 2017 International Conference on Digital Health (DH '17). 3079480: ACM; 2017 Presented at: International Conference on Digital Health; 2017-07-02; London, United Kingdom p. 229-230 URL: https://dl.acm.org/citation.cfm?id=3079452.3079480
31. Ezra S. Apptentive. 2014 May 27. Fake Reviews in Google Play and Apple App Store URL: http://www.apptentive.com/blog/2014/05/27/fake-reviews-google-play-apple-app-store [accessed 2018-10-29] [WebCite Cache ID 6wMel1pDD]
32. 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]
33. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
34. Pobiruchin M, Suleder J, Zowalla R, Wiesner M. Accuracy and Adoption of Wearable Technology Used by Active Citizens: A Marathon Event Field Study. JMIR Mhealth Uhealth 2017 Feb 28;5(2):e24 [FREE Full text] [doi: 10.2196/mhealth.6395] [Medline: 28246070]
35. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
36. Breton ER, Fuemmeler BF, Abroms LC. Weight loss-there is an app for that! But does it adhere to evidence-informed practices? Transl Behav Med 2011 Dec;1(4):523-529 [FREE Full text] [doi: 10.1007/s13142-011-0076-5] [Medline: 24073074]
37. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]

Abbreviations

ACSM: American College of Sports Medicine
BCT: Behavior change technique
BCTTv1: Behavioral Change Technique Taxonomy v1
eHealth: electronic health
EIPA: Evidence-informed physical activity
mHealth: mobile health
PA: physical activity
WHO: World Health Organization

Edited by G Eysenbach; submitted 14.03.18; peer-reviewed by D Regan, C Delisle, RM Payo; comments to author 17.08.18; revised version received 30.08.18; accepted 14.09.18; published 18.12.18.

Please cite as:
Kebede M, Steenbock B, Helmer SM, Sill J, Möllers T, Pischke CR
Identifying Evidence-Informed Physical Activity Apps: Content Analysis
JMIR Mhealth Uhealth 2018;6(12):e10314
URL: https://mhealth.jmir.org/2018/12/e10314/
doi:10.2196/10314
PMID:30563810

Original Paper

Mobile Apps to Support Healthy Family Food Provision: Systematic Assessment of Popular, Commercially Available Apps

Chelsea E Mauch[1,2], BNutrDiet, BSc (Hons); Thomas P Wycherley[3], BEd, BApSc, BSc (Hons), MEp, PhD; Rachel A Laws[4], BSc, MSc (NutrDiet), PhD; Brittany J Johnson[1,2], BNutrDiet, BHSc (Hons); Lucinda K Bell[1], BNutrDiet (Hons), PhD; Rebecca K Golley[1,2], BSc (Hons), BNutrDiet, PhD

[1]Nutrition and Dietetics, College of Nursing and Health Sciences, Flinders University, Bedford Park, Australia

[2]School of Pharmacy and Medical Sciences, Division of Health Sciences, University of South Australia, Adelaide, Australia

[3]School of Health Sciences, Division of Health Sciences, University of South Australia, Adelaide, Australia

[4]Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Science, Deakin University, Geelong, Australia

Corresponding Author:
Chelsea E Mauch, BNutrDiet, BSc (Hons)
Nutrition and Dietetics
College of Nursing and Health Sciences
Flinders University
Level 7, Flinders Medical Centre
Flinders Drive
Bedford Park, 5042
Australia
Phone: 61 882047075
Fax: 61 882046406
Email: chelsea.mauch@flinders.edu.au

Abstract

Background: Modern families are facing conflicting demands on their time and resources, which may be at the detriment of child and family diet quality. Innovative nutrition interventions providing parents with behavioral support for the provision of healthy food could alleviate this issue. Mobile apps have the potential to deliver such interventions by providing practical behavioral support remotely, interactively, and in context.

Objective: This review aimed to identify and assess popular, commercially available food- and nutrition-related mobile apps that offer support for the provision of healthy family food by (1) describing app scope and characteristics, (2) assessing app quality, and (3) conducting a behavioral analysis of app content and features.

Methods: Searches in the Google Play Store and Apple App Store between August 2017 and November 2017 identified apps addressing the food provision process. Apps were included if they were applicable to parents or families, written in English, and with a user rating of ≥4 stars. Weight loss and diet monitoring apps and subscription apps with no free versions were excluded. App quality was assessed using the Mobile App Rating Scale (4 domains: engagement, functionality, aesthetics, and information). App content and features were extracted and behavior change techniques (BCTs) identified.

Results: Of the 2881 apps screened, 1.77% (51/2881) were included for assessment, comprising 23 recipe and recipe manager apps, 12 meal planning apps, 10 shopping list apps, 4 family organizers, and 2 food choice apps. Half (n=26) of the apps functioned primarily through user data input. Food choice and family organizer apps scored highest for app quality (mean 3.5 [SD 0.6] out of 5), whereas most apps scored well for functionality and poorly for engagement. Common app features with the potential to support healthy food provision included meal planners (n=26), shopping lists (n=44), and the ability to share app content (n=48). Behavioral support features mapped to relatively few BCTs (mean 3.9 [SD 1.9] per app), with *Adding objects to the environment* present in all apps, and 65% (33/51) including *Instruction on how to perform the behavior*.

Conclusions: Recipe and recipe manager apps, meal planning apps, and family organizers with integrated meal planning and shopping lists scored well for functionality and incorporated behavioral support features that could be used to address barriers to healthy food provision, although features were focused on planning behaviors. Future apps should combine a range of features such as meal planners, shopping lists, simple recipes, reminders and prompts, and food ordering to reduce the burden of the food

provision pathway and incorporate a range of BCTs to maximize behavior change potential. Researchers and developers should consider features and content that improve the engagement quality of such apps.

(JMIR Mhealth Uhealth 2018;6(12):e11867) doi:10.2196/11867

KEYWORDS

diet; nutrition; family; mobile applications; behavior modification

Introduction

Background

Excessive consumption of energy-dense, nutrient-poor foods is a key cause of poor diet quality [1-4] and is contributing to the high prevalence of overweight and obesity globally [5-7]. In Australia, these foods are contributing 30% to 40% of the total daily energy intake of children and adolescents [2]. Similar figures have been reported in the United States and Canada, with children and adolescents consuming at least one-third of their daily energy intake in the form of energy-dense, nutrient-poor foods [4,8]. The increasing reliance on these generally highly processed foods may be in part because of the conflicting demands that the modern lifestyle places on the resources available for family food provision [9].

Food provision, encompassing the planning, purchasing, and preparation of food, requires significant time and both mental (eg, food preparation knowledge and planning skills) and physical (eg, food preparation facilities) resource [10-12]. The use of food coping strategies (such as meal planning, shopping list writing, use of convenience ingredients or preprepared meals, and seeking support) can enable families to overcome resource-related barriers to food provision (including time or income scarcity). Although some strategies, such as the purchase of fast or convenience food, occur at the detriment of diet quality [13-17], other strategies, such as meal planning and shopping list use, have been associated with healthier food preparation [13]. Nutrition interventions supporting the use of healthy food coping strategies are warranted and in fact desired by parents [18]. However, interventions supporting parents to improve their children's dietary intake are primarily focused on education rather than skill development and are of moderate effectiveness [19]. Addressing resource-related food provision barriers and supporting the adoption of healthy food coping behaviors may enhance the effectiveness of interventions to improve child and family diet quality [11,12,20].

Health interventions delivered by mobile apps have the potential to address resource-related barriers to healthy food provision by offering practical behavioral support, remotely, interactively, and in context [21]. The unique placement of mobile phones within our daily lives, along with technological advancements such as global positioning system, machine learning, and data tracking, means that apps are positioned to deliver ecological momentary interventions [21,22]. Although the initial time and monetary outlay for app development can be substantial, they are highly scalable, and with mobile phone ownership nearing saturation, they have the potential to reach a diverse population [23,24]. Furthermore, interventions can be personalized based on user input, which may improve user engagement and intervention fidelity [22,23]. The current popularity of health and nutrition-related apps in both the general public and in research, along with the opportunities that the technology provides, makes it an important platform to explore for future family nutrition interventions [23,25].

Reviews of Mobile Health Interventions and Commercially Available Apps

Reviews of nutrition-related mobile health interventions have examined their effectiveness in relation to behavioral and weight-related outcomes [26-29]. A meta-analysis of 12 diet and physical activity–focused app studies found that delivery of an intervention via a mobile app significantly reduced weight compared with controls (−1.04 kg, 95% CI −1.80 to −0.27 kg) [27]. Similarly, a systematic review found moderate evidence that diet and physical activity apps lead to improvements in health-related behaviors and outcomes (19 of 27 apps) [29]. However, these reviews have generally focused on apps for weight loss or diet monitoring, with limited relevance to family food provision [26-29].

A recent scoping review identified studies describing apps relevant to families, although the focus was primarily on apps supporting parent food practices (ie, responding to vegetable refusal and food portions) and monitoring of family members snack intake [30]. The same review identified a small subset (19%, 9/47) of mainly app development studies describing food access and food purchasing apps [30]. These apps were found to utilize environmental support features such as recipe suggestions and augmented reality tagging of products in the supermarket aisle [30]. Therefore, although there is evidence of the development of apps providing behavioral support for aspects of food provision, there is a paucity of published research exploring the use of apps for families that consider a range of food provision processes. To understand the potential role of apps in addressing a range of food provision processes, it is crucial to look toward existing, commercially available apps to support innovation in future research studies [23].

Reviews of apps in the commercial space have assessed app features and quality as well as identified the behavior change technique (BCT) content of nutrition, physical activity, and weight management apps targeting adults [31-33] and children [34,35]. These reviews found that there remains a need to enhance app quality and utilize behavior change theory in app development as important precursors to app effectiveness [31,33,34]. The focus of these apps on diet and weight-related outcomes (such as calorie counting and weight monitoring), rather than the behaviors leading to healthy dietary intake and weight, may limit their behavior change potential [33]. Similar to reviews of published app studies, commercial apps pertaining to food provision in a family context have yet to be explored. To ensure that current technological and behavior change

potential in this area is fully understood, and to understand gaps in the commercial space, a review of existing, commercial apps addressing family food provision is required.

Objectives

Thus, the purpose of this review was to identify and assess popular, commercially available food and nutrition-related mobile apps that have the potential to offer behavioral support for the provision of healthy family food. Specifically, the objectives of this systematic assessment were to describe app scope and characteristics, assess app quality, and conduct a behavioral analysis of app content and features.

Methods

Search Strategy

Systematic searches were conducted in the Google Play Store and Apple App Store between August 2017 and November 2017. The search strategy was modeled on prior systematic assessments in similar fields of research [31,32,34,36]. Google Play searches were conducted on a personal computer in a Google Chrome Web browser without Google account log-in. App Store searches were performed using the app on an iPad, as the store does not include a search function when used on a personal computer [32]. Search terms relating to the food provision process were selected, and pilot searches in both stores resulted in the following primary terms being used to identify apps for inclusion:

- WHO: child, children, toddler, kid, kids, preschooler, family, families, and parent
- WHAT: nutrition, food, meal, menu, recipe, recipes, and diet
- HOW: planning, planner, shopping, supermarket, grocery, budget, cook, cooking, prep, and preparation

Terms were combined into groups reflecting the various stages of the food provision process, including meal planning; food budgeting; nutrition, food, and cooking knowledge; food purchasing; and meal preparation. Combinations of 2 to 3 words were then generated for each group (eg, meal planner and child meal plan), and the first combination from a group was entered, with the first 50 results being checked by title and description against the inclusion and exclusion criteria. This was repeated for subsequent search terms from that group until a term returned no new apps that met the inclusion criteria. The search was then deemed saturated for that group and the next group of search terms applied.

App Selection

Apps were included if they were applicable to parents with children, written in the English language, and had a user rating of at least four stars in the Google Play Store (to ensure that only popular, functional apps were reviewed) [31]. This limit was not applicable in the App Store as most apps had insufficient reviews to be given a star rating. All free, paid, and freemium apps were included, except where the app was subscription only with no freemium version. The following app types were excluded: (1) weight loss, diet monitoring, and calorie counter apps; (2) generic apps with only 1 food-related component (ie, personal organizers with a shopping list); (3) infant food and feeding apps; (4) apps focused on child feeding practices, electronic books, or magazines; and (5) recipe apps focused on unhealthy food (ie, cakes) or 1 key ingredient or cuisine. Apps were also excluded if their use was contingent upon involvement in a research study or a face-to-face component. The initial screen using these criteria was conducted using the app name, description, and screenshots of the app found within the stores. Approximately 10% of the screened apps (selected randomly, using the random number function within Microsoft Excel 2016) were checked by a second reviewer for correct inclusion and exclusion. Agreement was 93.7% (256/273), with discrepancies discussed and consensus reached [36].

Due to large numbers of similar and generic apps (eg, basic shopping list apps), a second and third screen was undertaken with additional exclusion criteria. At the second screen, apps with only 1 food-related component (ie, recipes only), less than 20 reviews in the Google Play store [34], and duplicates between stores were excluded. Apps were then grouped according to their primary purpose as described in the Google Play Store or App Store, and a third screen applied to ensure that the final sample provided good representation of the features available in such apps. Using the app description in the Google Play Store and App Store, apps were included if they had at least one unique feature not yet described in another app from that group of apps, or features in a unique combination.

Data Extraction and Assessment

Once all eligible apps were identified, an Apple iPad Mini Version 4 (Model A1550) and Lenovo Tab3 7 Essential (Model TB3-710F) were used to download apps for assessment. Where apps were *freemium* (ie, available for free but with some features only accessible with payment), the paid version was purchased, except where subscription was required. These apps were downloaded and assessed in the free version. Apps were used for a minimum of 10 min before any data extraction or assessment took place [37]. Reviewers used individual apps for a period of time (generally on a number of occasions) that was sufficient to familiarize themselves with the apps features and functionality. The time spent using apps varied because of the significant heterogeneity of the included apps. Data extraction was checked, and apps were assessed independently by a second reviewer in a random sample of 22% (11/51).

App Characteristics

App information including app and developer name, operating system availability, version, affiliations, cost structure, user rating and number of downloads (where available), and app scope (ie, target audience and behavior) was extracted into a purpose-designed Microsoft Excel 2016 spreadsheet. The primary direction of data into or out of the app was determined and described as input, output, or both. App content such as information, videos, images, and recipes were defined as *output*, whereas features requiring user input, such as entering items into shopping lists or meal planners, were defined as *input*.

App Quality

App quality was assessed using the Mobile App Rating Scale (MARS), an objective and reliable measure of the quality of

health-related apps [37]. The domains assessed by the MARS tool include engagement, functionality, aesthetics, and information [37]. An optional domain regarding subjective app quality was not included in this study. Apps were rated between 1 and 5 for each of the criteria, with 4 mean domain scores and an overall mean score across all 4 domains being indicative of app quality (a score of 5 indicating the best performing apps). Both reviewers viewed a Web-based training video before app assessment [38]. Inter-rater reliability of the overall MARS score was tested on the sample of double-assessed apps using the two-way random effects intraclass correlation coefficient (ICC) [39]. The resulting ICC value of .74 indicated good inter-rater reliability [40].

App Content and Features

Data regarding app content and features were sorted into 2 distinct categories: (1) "Behavioral support content and features" and (2) "Technical features." "Behavioral support content and features" were those that may enable the performance of a behavior relating to the provision of healthy family food. "Technical features" did not offer behavioral support but were important to the overall functioning of the app. App content was then assessed for the presence of BCTs against the BCT taxonomy version 1 (BCTTv1) [41]. Both reviewers underwent Web-based training before coding [42]. The agreement between reviewers regarding the presence of BCTs was tested in the 11 double-assessed apps using kappa and prevalence adjusted and bias adjusted kappa (PABAK) and was near perfect (kappa mean 0.82 [range 0.66-1], PABAK 0.97 [range 0.94-1]) [43].

Statistical Analysis

Means (SD) for each MARS subscale and the overall MARS score were calculated using Microsoft Excel 2016 for each app. A summary score was calculated for each app type (ie, recipe and recipe managers, meal planners, shopping lists, family organizers, and food choice apps) along with an overall mean score for all apps. The mean (SD) number of BCTs per app and app type was calculated, and the total number of apps from each app type incorporating the BCT was presented graphically. The presence of behavioral content and features and technical features was tallied for each app type and for all apps.

Results

App Selection

A total of 2881 apps were screened across the Google Play Store and Apple App Store. The final number included for assessment was 51 (see Figure 1).

App Characteristics

Selected apps fell into 5 categories of app type: (1) recipe and recipe manager apps, which provided recipes or digital storage of recipes; (2) meal planning apps, which allowed the planning and recording of meals in advance; (3) shopping list apps, which allowed recording of grocery items for purchase; (4) family organizer apps, which included meal planners and shopping lists synced between family members; and (5) food choice apps, which provided nutrition or produce information to support food purchasing (see Multimedia Appendix 1 for app details and MARS scores).

Recipe and recipe manager apps were the most common app type in the sample (45%, 23/51), followed by meal planning apps (24%, 12/51). Almost all apps were developed by commercial enterprises, with the exception of 1 app developed by a government body and another by a nongovernment research institute in collaboration with a private health insurer. Approximately one-third (31%, 16/51) of apps were free to download and use (see Multimedia Appendix 2). The primary behavioral targets of the apps included food purchasing (90%, 46/51), meal preparation (76%, 39/51), meal planning (47%, 24/51), and food choice (10%, 5/51). Half (51%, 26/51) of the apps operated primarily on input from the app user, with shopping lists and family organizers being most reliant on user data input. Only one-quarter of apps incorporated both significant user data input along with app information output (25%, 13/51).

App Quality

The mean MARS score for app quality was highest for food choice apps and family organizer apps (mean 3.5 [SD 0.6] out of 5 for each), followed by recipe and recipe manager apps (mean 3.4 [SD 0.5]). Shopping list apps had the lowest overall MARS scores, with half of the apps scoring below 2.5 (for MARS scores by app type, see Table 1, and by individual app, see Multimedia Appendix 1). Engagement was the lowest scoring domain for each app type, with shopping lists and meal planners performing the worst. Most app types scored well for functionality (mean across all app types 3.6 [SD 0.7]).

App Content and Features

Behavioral Support Content and Features

App content and features relating to the provision of healthy family food are presented by app type (see Table 2, and for details by app, see Multimedia Appendix 2). Several common app features supported the use of key healthy food coping strategies, for example, meal planners, shopping lists, and social supports. Meal planners were the primary feature of all 12 meal planning apps and featured in around half of the overall sample (51%, 26/51). Shopping lists featured almost universally (86%, 44/51) and where incorporated into other app types (as opposed to a stand-alone shopping list app), they generally offered automated list generation. Similarly, almost all (94%, 48/51) apps included the ability to share app content by email and/or social media.

Recipes and recipe managers (the primary feature of recipe and recipe manager apps, n=23) were present in more than half of the overall sample (recipes 33/51, 65% and recipe managers 28/51, 55%). Food preparation skills instructions were uncommon (14%, 7/51) and included either text, image, or video-based instructions. Reminders and/or prompts were included in almost a third of apps (27%, 27/51). A small number of apps included general and produce-related nutrition information (16%, 8/51), whereas only 3 apps (6%) included the ability to purchase food for delivery.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for popular, commercially available food and nutrition-related apps addressing parental food provision.

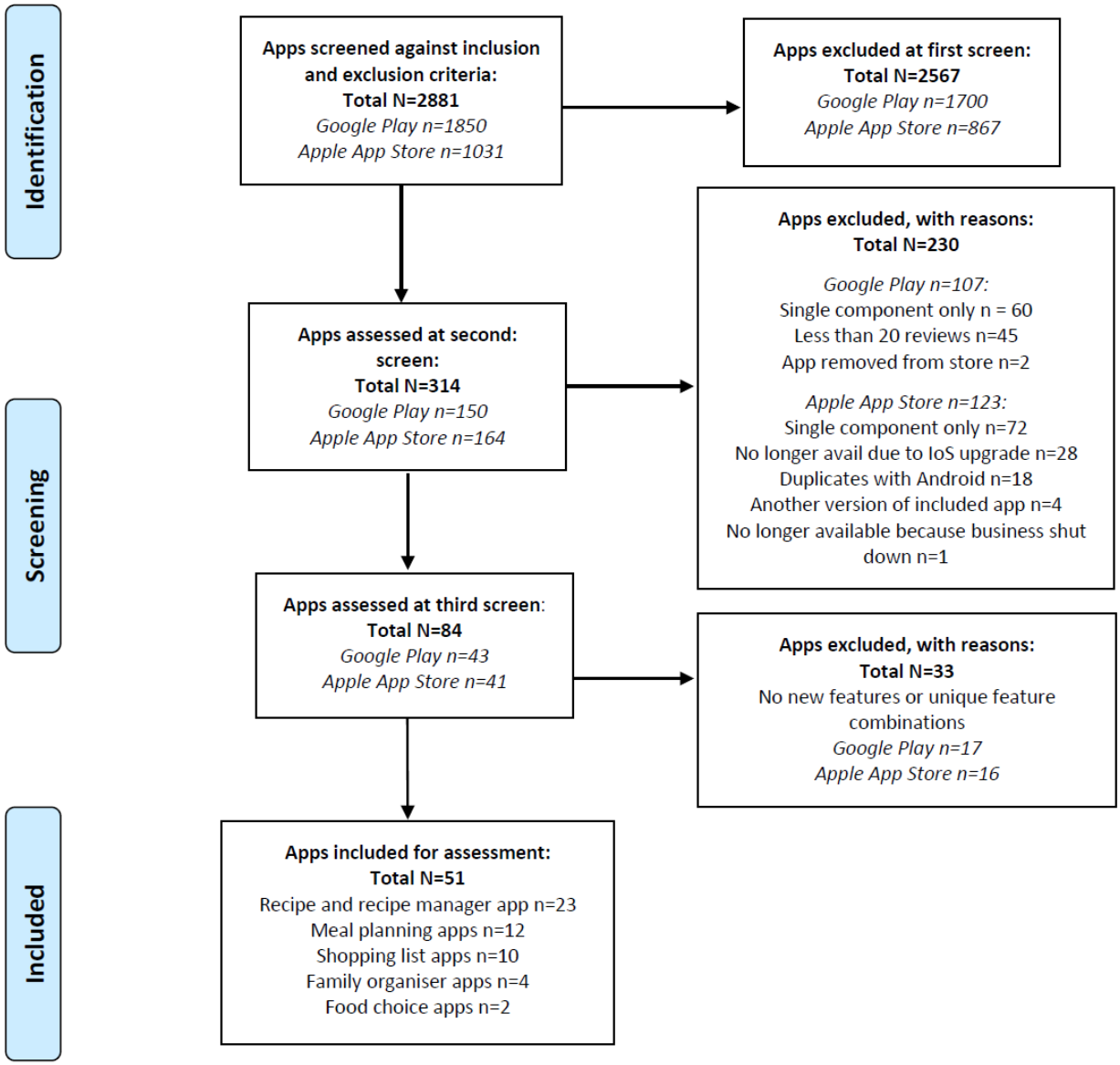


Table 1. Mean (SD) Mobile App Rating Scale (MARS) subscores and overall scores by app type.

MARS domain sub-scores and overall score	Recipe and recipe manager apps (n=23), mean (SD)	Meal planning apps (n=12), mean (SD)	Shopping list apps (n=10), mean (SD)	Family organizer apps (n=4), mean (SD)	Food choice apps (n=2), mean (SD)	All apps (N=51), mean (SD)
Engagement	2.7 (0.6)	2.5 (0.8)	2.1 (0.4)	3.2 (0.7)	2.7 (1.3)	2.6 (0.7)
Functionality	3.8 (0.6)	3.8 (0.7)	3.0 (0.9)	3.7 (0.6)	4.4 (0.2)	3.6 (0.7)
Aesthetics	3.6 (0.8)	3.2 (1.0)	2.9 (0.9)	3.7 (0.9)	2.8 (0.2)	3.3 (0.9)
Information	3.4 (0.4)	3.2 (0.6)	2.9 (0.5)	3.6 (0.5)	4.0 (0.7)	3.3 (0.6)
Overall score	3.4 (0.5)	3.1 (0.7)	2.7 (0.6)	3.5 (0.6)	3.5 (0.6)	3.2 (0.6)

Table 2. App behavioral support content and features presented by app type and across all apps

Behavioral support content or feature	Recipe and recipe manager apps (n=23), n (%)	Meal planning apps (n=12), n (%)	Shopping list apps (n=10), n (%)	Family organizer apps (n=4), n (%)	Food choice apps (n=2), n (%)	All apps (N=51), n (%)
Meal planners and meal plans	10 (44)	12 (100)	2 (20)	2 (50)	0 (0)	26 (51)
Shopping list	20 (87)	9 (75)	10 (100)	4 (100)	1 (50)	44 (86)
Social community or connectivity[a]	10 (44)	4 (33)	0 (0)	0 (0)	0 (0)	14 (27)
Other social supports[b]	23 (100)	11 (92)	9 (90)	4 (100)	1 (50)	48 (94)
Recipes	19 (83)	6 (50)	4 (40)	3 (75)	1 (50)	33 (65)
Recipe managers	13 (57)	6 (50)	7 (70)	2 (50)	0 (0)	28 (55)
Pantry or fridge manager	1 (4)	1 (8)	5 (50)	0 (0)	0 (0)	7 (14)
Food preparation skills instructions	6 (26)	1 (8)	0 (0)	0 (0)	0 (0)	7 (14)
Reminders and prompts[c]	4 (17)	4 (33)	5 (50)	1 (25)	0 (0)	14 (27)
Encouragement and incentives[d]	8 (35)	1 (8)	4 (40)	2 (50)	0 (0)	15 (29)
Produce purchasing information	0 (0)	0 (0)	0 (0)	0 (0)	1 (50)	1 (2)
Produce storage information	1 (4)	0 (0)	0 (0)	0 (0)	1 (50)	2 (4)
Produce nutrition information	1 (4)	1 (8)	1 (10)	0 (0)	2 (100)	5 (10)
Recipe nutrition information	6 (26)	3 (25)	1 (10)	0 (0)	0 (0)	10 (20)
Other nutrition information	2 (9)	1 (8)	0 (0)	0 (0)	0 (0)	3 (6)
Food purchase and delivery	1 (4)	1 (8)	1 (10)	0 (0)	0 (0)	3 (6)

[a]Community (with following), upload recipes or images, rate, review, like, and comment.

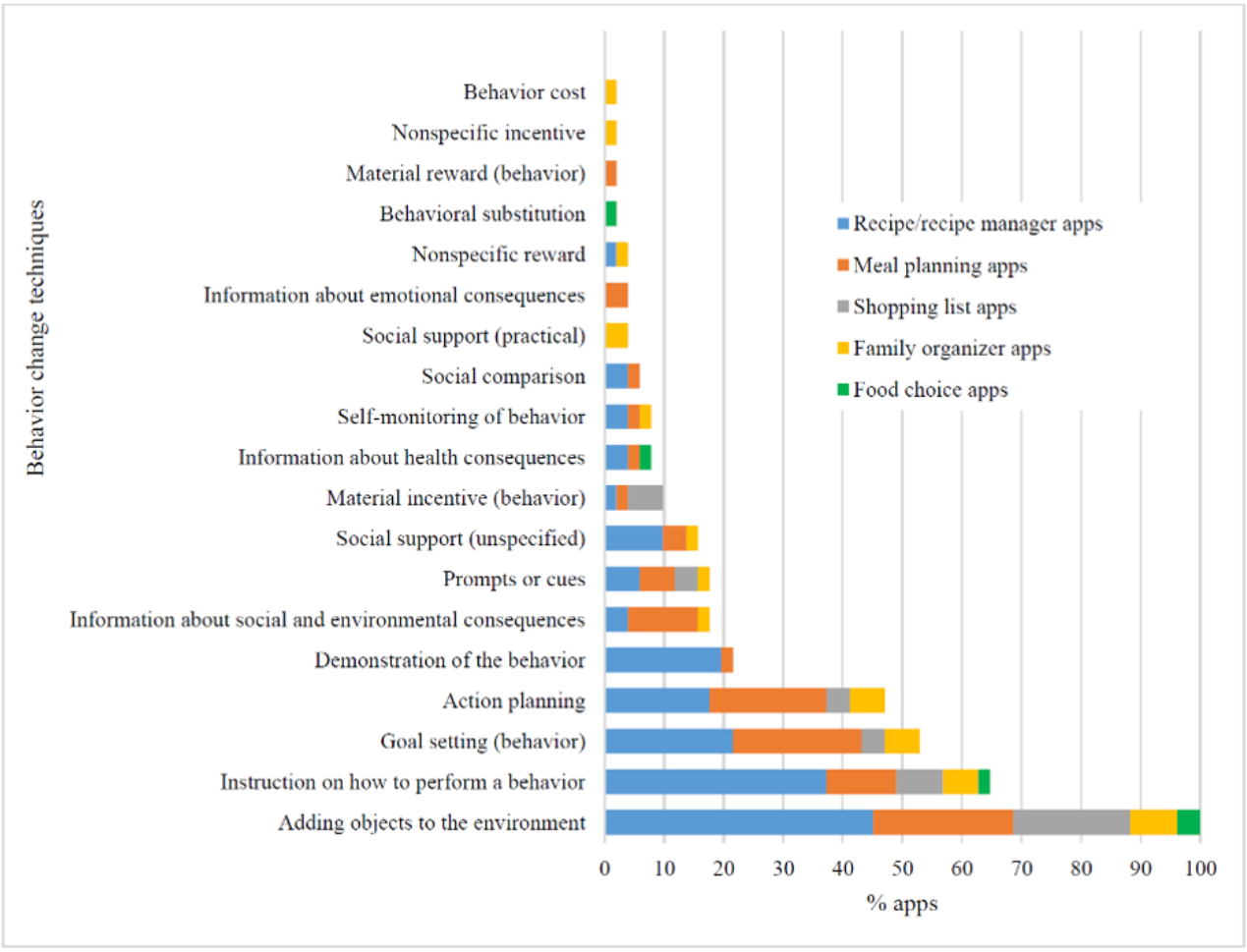
[b]Sharing to social media, sending via email, shared calendar, and private texting.

[c]Recipe suggestions on entering the supermarket, supermarket proximity alert, and reminders (to cook, plan meals, and shop).

[d]Positive messages, points, rewards, competitions, sales or discounts, and other notifications (eg, new content and offers).

Of 93 BCTs in the taxonomy, 19 were identified as being present across the 51 apps, with a mean (SD) of 3.9 (1.9) per app ranging from 1 to 10 (see Figure 2 and Multimedia Appendix 3). Family organizer apps followed by meal planning apps were identified as having the greatest number of BCTs (mean 5.5 [SD 3.1] and mean 4.8 [SD 1.9], respectively). Recipe and recipe manager apps included an average of 4 BCTs per app (mean 3.9 [SD 1.5]), whereas food choice apps and shopping list apps were identified as including the least number of BCTs (mean 2.5 [SD 0.7] and mean 2.3 [SD 0.8], respectively). The only BCT that was identified as being present across all apps was *Adding objects to the environment*. This was because of features such as shopping lists and meal planners that were thought to add objects to the environment that may subsequently enable a behavior relating to healthy food provision. Recipe and recipe managers commonly included the BCT *Instruction on how to perform the behavior* (83%, 19/23), owing to the inclusion of recipes with step-by-step instructions. Furthermore, 92% (11/12) and 83% (11/12) of meal planning apps included BCTs *Goal setting (behavior)* and *Action planning*, owing primarily to the ability to plan meals in advance.

Figure 2. Proportion of apps identified with Behavior Change Technique present, by mobile app type.



Technical Features

Technical features were grouped separately as they were unlikely to directly support behavior but remained important to the overall functioning and engagement of the mobile apps (Table 3 and Multimedia Appendix 2). Two-thirds of apps (69%, 35/51) allowed some level of personalization, such as a customized recipe display based on food preferences, dietary requirements, or number of serves required. More than half of all apps (57%, 29/51), predominantly recipe and recipe manager apps, included practical features such as cooking timers, unit converters (ie, cups to milliliters), voice input of data, hands free commands, and automatic screen lock to prevent the device from sleeping while the app is in use. A little over half of the apps allowed syncing between devices and cloud backup (59%, 30/51 and 57%, 29/51, respectively).

Table 3. Technical features presented by app type and across all apps.

Technical feature	Recipe and recipe manager apps (n=23), n (%)	Meal planning apps (n=12), n (%)	Shopping list apps (n=10), n (%)	Family organizer apps (n=4), n (%)	Food choice apps (n=2), n (%)	All apps (N=51), n (%)
Personalization[a]	20 (87)	9 (75)	4 (40)	1 (25)	1 (50)	35 (69)
Practical features[b]	17 (74)	4 (33)	7 (70)	1 (25)	0 (0)	29 (57)
Syncing between devices	12 (52)	6 (26)	8 (80)	4 (100)	0 (0)	30 (59)
Cloud backup	14 (61)	5 (42)	7 (70)	3 (75)	0 (0)	29 (57)
User or family profile[c]	7 (30)	3 (25)	0 (0)	3 (75)	0 (0)	13 (25)
Miscellaneous and optional purchases[d]	4 (17)	2 (17)	5 (50)	2 (50)	0 (0)	13 (25)
Search and display options[e]	19 (83)	5 (22)	8 (80)	4 (100)	1 (50)	37 (73)
Other input options[f]	6 (26)	6 (26)	10 (100)	3 (75)	1 (50)	26 (51)
Requires log-in	12 (52)	6 (26)	7 (70)	3 (75)	0 (0)	28 (55)
Web access required	21 (91)	10 (83)	5 (50)	4 (100)	2 (100)	42 (82)

[a]Food preferences, dietary requirements, favorites lists, scale recipes to serves required, and add notes or rating to recipes (private).

[b]Prevents device from sleeping, voice command, audio reading, hands free, smart watch compatible, cooking timers, and unit conversions.

[c]Individual profile or profile of individual family members or family as a whole.

[d]To-do lists and optional purchases (eg, hard copy cookbook and cooking equipment).

[e]Search functions, for example, by ingredient, recipe name, and category (eg, vegetarian), and novel search functions, for example, by shaking device and by photo.

[f]Common items lists, history or recurring items, barcode scanners, add images, coupons, and loyalty cards.

Discussion

Principal Findings

This review identified and assessed commercially available food and nutrition-related mobile apps addressing family food provision. Most apps provided behavioral support for the use of healthy food coping strategies, although supports were biased toward planning behaviors, which may appeal to some but not all users. App features and content mapped to relatively few BCTs, with the higher quality family organizer apps, meal planning apps, and recipe and recipe manager apps incorporating the greatest number of techniques, respectively. Recipe and recipe manager apps, meal planning apps, and family organizers with integrated meal planning and shopping lists were found to be highly functional with regards to their performance and ease of use and incorporated a range of behavioral support features that could be used to address barriers to healthy food provision, such as time scarcity and cognitive load.

App Characteristics and Quality

The majority of apps targeted meal planning and shopping list use, both considered healthy food coping strategies [13]. Although these food coping strategies are associated with healthier food preparation practices, they are best suited to those more inclined to plan [15]. Few apps effectively addressed food coping strategies such as preparing meals with few ingredients on hand, utilizing healthy convenience foods (ie, frozen or canned products and meal box kits), or seeking support. Furthermore, observed features often required extensive data input (eg, recipe managers and family organizers), which may be a barrier to app engagement or use [44].

Although most apps were generally functional in terms of their performance, ease of use, navigation, and gestural design, their low ratings for the engagement domain of the quality assessment was a concern, given this is a key predictor of long-term use [23]. A recent review of 11 weight loss apps addressing food-purchasing behavior reported similar findings [33], whereas, others have identified concerns regarding information quality and highlighted the need for evidence-based content [34]. However, as the information within the apps assessed in this review was mostly limited to recipes or food skills, the information quality rating is less relevant. The evidence base of such apps should be in their delivery of behavioral supports, to ensure that they have a positive influence on the food provision process.

Behavioral Analysis

Mobile app behavioral supports such as shopping lists, meal planners, and recipe managers have the advantage of delivering BCTs in the real world, when behaviors are likely to occur, thus improving the chance of positively shaping behavior [21,22]. However, the number of BCTs identified in the present sample of apps was lower compared with similar reviews of weight loss and general nutrition apps [31,34], reflecting the development of these apps for commercial purposes rather than for behavior change or health promotion. This indicates

significant scope for increasing the behavior change potential of future apps in this space.

There were a number of app types and features that should be considered in the development of future evidence-based, behavioral change theory–driven apps targeting food provision in families. Meal planning apps and features, supporting the formation of intentions to prepare a healthy meal, were identified as including the second largest mean number of BCTs. Most notably, they incorporated *Goal Setting (behavior)* and *Action Planning*. The 2 meal planning apps with the highest MARS scores and largest number of BCTs allowed the user to outsource some aspects of the planning and purchasing process. One included automated meal plans and shopping lists produced using an internal bank of recipes, whereas the other offered meal box kit ordering and delivery. These apps could be suitable for those not naturally inclined to plan and willing to relinquish some decision making regarding meals. However, inadequate personalization, complex recipes, and the high cost associated with ingredients and box kits may be barriers to the widespread use of such apps.

Shopping lists as a stand-alone app type generally failed to offer more than the conventional paper and pen method, so it was unsurprising that they performed poorly on all domains of the MARS and mapped against very few BCTs. Where shopping lists were incorporated into other app types and allowed automatic list generation through recipes, they have the potential to reduce the time burden associated with shopping list writing. Linking to Web-based grocery ordering would add a further efficiency; however, this feature was surprisingly uncommon, only being incorporated into 2 of 51 apps.

Another feature with the potential to increase efficiencies relating to food purchasing is the ability to sync grocery lists between family members (ie, a shared shopping list). This feature could be utilized to share the mental and physical load of planning and purchasing food. Family organizers generally offered the ability to share such tasks among family members but most were expensive (eg, up to Aus $69.99 per year subscription), requiring an ongoing subscription to access such features. Furthermore, they required significant data input and are likely suited to those with established planning skills.

Few apps incorporated timely reminders and prompts, which is a missed opportunity to take advantage of mobile apps ability to offer ecological momentary intervention [22]. If used appropriately (ie, not overwhelmingly) and timed to coincide with the performance of food-related behaviors, reminders and prompts in the form of push notifications could act to reduce the mental load of the food provision process. Supermarket proximity alerts and reminders of the planned evening meal were effective, albeit uncommon, examples of such push notifications, delivering the BCT *Prompts and cues.*

Most of the apps assessed provided limited information, generally in the form of recipes and food skills, which is consistent with the move toward more data input style apps. This content was associated with *Instruction on how to perform a behavior* and where video or image content was included, *Demonstration of the behavior*. However, most apps providing recipes or food skills were not focused on healthy food preparation or use of healthy food coping strategies (ie, utilizing frozen or canned foods, cooking from few ingredients), and few directly targeted families. Nutrition information delivered in the context of food purchasing, such as in 1 reviewed app that suggested healthier alternatives to scanned products, may be more likely to support behavior change than generic nutrition information. However, it is possible that the way information is presented and the functionality of the app delivering it determines its efficacy in changing behavior; For example, the convenience of the information (ie, barcode scanners for searching) and the pairing of recipes with relevant food skills videos, hands free commands, single directions displayed per page, and text to speech functions.

Review Strengths and Limitations

Although the search strategy of this review was systematic and based on similar reviews of commercial apps for nutrition and weight management [31,32,34], it was limited by the lack of standard methodology for searching commercial mobile app stores. Lack of standardized search methods and limited and variable information provided in app descriptions made it difficult to ensure all eligible apps were captured, particularly high-quality apps. There were also limitations relating to the use and interpretation of the MARS score. The information quality domain was limited to assessing the accuracy of the app description and the credibility of the app developer in the absence of assessable information and should, therefore, be interpreted with caution. Moreover, although family organizer apps and food choice apps scored the highest MARS ratings, they were based on only 4 and 2 apps, respectively. Finally, the coding of BCTs was limited to features and content that could be accessed or viewed within the assessment period. Therefore, some push notifications may have been overlooked, whereas lengthy blogs within apps were excluded from detailed analysis.

Despite its limitations, this review assessed a large number of apps and provides unique information about their behavior change potential by not only describing and assessing app scope, characteristics, and quality but also through a behavioral analysis of app content and features. Reviewer training, along with the use of a second reviewer in a 20% sample, improves the objectivity and accuracy of the data extracted and assessed in this review. The present target group is families, but the findings have applications to food planning, purchasing, and preparation behaviors more generally.

Implications for Practice and Future Research

The findings of this review suggest that recipe and recipe manager apps, family organizer apps, and meal planning apps should be explored as viable options for nutrition promotion interventions. Future apps should combine a range of behavioral support features such as meal planners, shopping lists, simple recipes, reminders and prompts, and food ordering to reduce the burden of the food provision process and maximize behavior change potential. Consideration of food coping strategies other than meal planning, or the incorporation of skills training, prompts, and encouragement to plan meals, would make these apps applicable to people less inclined to plan. Although particular attention should be paid to personalization features, they should also provide a level of automation that reduces the

need for excessive data input. Finally, researchers and developers should be mindful of the needs of modern families and consider the engagement qualities of such apps to ensure their effectiveness and longevity.

Conclusions

This review, assessing commercially available food and nutrition-related apps for family food provision, demonstrates that apps could be used to deliver behavioral support for healthy food coping strategies. Future apps should include a wider range of features and BCTs to promote engagement and improve the behavior change potential of such apps.

Acknowledgments

CEM is supported by a National Health and Medical Research Council (NHMRC) postgraduate scholarship (GNT1114194) and a top-up scholarship funded by Flinders University (previously University of South Australia) and The Early Prevention of Obesity in Childhood Centre for Research Excellence (APP1101675). RAL and RKG are researchers within the NHMRC Centre for Research Excellence in The Early Prevention of Obesity in Childhood (APP1101675). BJJ was supported by an Australian Government Research Training Program Scholarship and is currently supported by a Flinders University Research Scholarship, Flinders University top-up scholarship, and a King and Amy O'Malley Trust Postgraduate Scholarship.

Authors' Contributions

CEM led the research, under the supervision of RKG, TPW, RAL, and LKB guiding study design and search strategy. CEM conducted all searches, data extraction, and app assessment, with BJJ undertaking double screening and assessment. CEM drafted the manuscript, with all authors contributing to the interpretation of results and reviewing of drafts. All authors read and approved of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Included app details and Mobile App Rating Scale scores.

[PDF File (Adobe PDF File), 21KB - mhealth_v6i12e11867_app1.pdf]

Multimedia Appendix 2

App content and features.

[PDF File (Adobe PDF File), 46KB - mhealth_v6i12e11867_app2.pdf]

Multimedia Appendix 3

Behavior change technique presence within apps, according to the behavior change technique taxonomy version 1.

[PDF File (Adobe PDF File), 21KB - mhealth_v6i12e11867_app3.pdf]

References

1. Johnson BJ, Bell LK, Zarnowiecki D, Rangan AM, Golley RK. Contribution of discretionary foods and drinks to Australian children's intake of energy, saturated fat, added sugars and salt. Children (Basel) 2017 Dec 01;4(12) [FREE Full text] [doi: 10.3390/children4120104] [Medline: 29194425]
2. Australian Bureau of Statistics. Canberra; 2014 May 09. Australian Health Survey: Nutrition First Results-Foods and Nutrients URL: http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/4683FD7315DFDFDBCA257D080014F9E0/$File/australian%20health%20survey%20nutrition%20first%20results%20-%20food%20and%20nutrients,%202011-12.pdf [accessed 2018-08-08] [WebCite Cache ID 71VXdKLrJ]
3. Golley RK, Hendrie GA. Dietary guideline index for children and adolescents: what is the impact of the new dietary guidelines? Nutrition & Dietetics 2014 Sep 26;71(3):210-212. [doi: 10.1111/1747-0080.12136]
4. Keast DR, Fulgoni VL, Nicklas TA, O'Neil CE. Food sources of energy and nutrients among children in the United States: National Health and Nutrition Examination Survey 2003–2006. Nutrients 2013 Jan 22;5(1):283-301 [FREE Full text] [doi: 10.3390/nu5010283] [Medline: 23340318]
5. Australian Institute of Health and Welfare. Canberra: Australian Institute of Health and Welfare; 2017. A picture of overweight and obesity in Australia URL: https://www.aihw.gov.au/getmedia/172fba28-785e-4a08-ab37-2da3bbae40b8/aihw-phe-216.pdf.aspx?inline=true [accessed 2018-08-08] [WebCite Cache ID 71VY7Gzgo]

6. National Health & Medical Research Council. Canberra: Australian Government; 2013. Australian Dietary Guidelines URL: https://www.nhmrc.gov.au/_files_nhmrc/file/publications/n55_australian_dietary_guidelines1.pdf [accessed 2018-08-08] [WebCite Cache ID 71VYMsAXA]
7. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet 2017 Dec 16;390(10113):2627-2642 [FREE Full text] [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]
8. Moubarac JC, Batal M, Louzada ML, Martinez Steele E, Monteiro CA. Consumption of ultra-processed foods predicts diet quality in Canada. Appetite 2017 Dec 01;108:512-520. [doi: 10.1016/j.appet.2016.11.006] [Medline: 27825941]
9. Li J, Akaliyski P, Schäfer J, Kendall G, Oddy WH, Stanley F, et al. Non-linear relationship between maternal work hours and child body weight: evidence from the Western Australian Pregnancy Cohort (Raine) Study. Soc Sci Med 2017 Dec;186:52-60. [doi: 10.1016/j.socscimed.2017.05.046] [Medline: 28582656]
10. Jabs J, Devine CM. Time scarcity and food choices: an overview. Appetite 2006 Sep;47(2):196-204. [doi: 10.1016/j.appet.2006.02.014] [Medline: 16698116]
11. Neumark-Sztainer D, MacLehose R, Loth K, Fulkerson JA, Eisenberg ME, Berge J. What's for dinner? Types of food served at family dinner differ across parent and family characteristics. Public Health Nutr 2014 Jan;17(1):145-155 [FREE Full text] [doi: 10.1017/S1368980012004594] [Medline: 23083836]
12. Rose D. Food Stamps, the Thrifty Food Plan, and meal preparation: the importance of the time dimension for US nutrition policy. J Nutr Educ Behav 2007;39(4):226-232. [doi: 10.1016/j.jneb.2007.04.180] [Medline: 17606249]
13. Morin P, Demers K, Turcotte S, Mongeau L. Association between perceived self-efficacy related to meal management and food coping strategies among working parents with preschool children. Appetite 2013 Jun;65:43-50. [doi: 10.1016/j.appet.2013.01.012] [Medline: 23402715]
14. Devine CM, Farrell TJ, Blake CE, Jastran M, Wethington E, Bisogni CA. Work conditions and the food choice coping strategies of employed parents. J Nutr Educ Behav 2009;41(5):365-370 [FREE Full text] [doi: 10.1016/j.jneb.2009.01.007] [Medline: 19717121]
15. Jabs J, Devine CM, Bisogni CA, Farrell TJ, Jastran M, Wethington E. Trying to find the quickest way: employed mothers' constructions of time for food. J Nutr Educ Behav 2007;39(1):18-25. [doi: 10.1016/j.jneb.2006.08.011] [Medline: 17276323]
16. Pescud M, Pettigrew S. 'I know it's wrong, but...': a qualitative investigation of low-income parents' feelings of guilt about their child-feeding practices. Matern Child Nutr 2014 Jul;10(3):422-435. [doi: 10.1111/j.1740-8709.2012.00425.x] [Medline: 22708589]
17. Venn D, Strazdins L. Your money or your time? How both types of scarcity matter to physical activity and healthy eating. Soc Sci Med 2017 Dec;172:98-106. [doi: 10.1016/j.socscimed.2016.10.023] [Medline: 27839899]
18. Fulkerson JA, Kubik MY, Rydell S, Boutelle KN, Garwick A, Story M, et al. Focus groups with working parents of school-aged children: what's needed to improve family meals? J Nutr Educ Behav 2011;43(3):189-193 [FREE Full text] [doi: 10.1016/j.jneb.2010.03.006] [Medline: 21367663]
19. Johnson BJ, Zarnowiecki D, Hendrie GA, Mauch CE, Golley RK. How to reduce parental provision of unhealthy foods to 3- to 8-year-old children in the home environment? A systematic review utilizing the Behaviour Change Wheel framework. Obes Rev 2018 Aug 09;19(10):1359-1370. [doi: 10.1111/obr.12702] [Medline: 30092606]
20. Fulkerson JA, Story M, Neumark-Sztainer D, Rydell S. Family meals: perceptions of benefits and challenges among parents of 8- to 10-year-old children. J Am Diet Assoc 2008 Apr;108(4):706-709. [doi: 10.1016/j.jada.2008.01.005] [Medline: 18375230]
21. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. J Behav Med 2017 Feb;40(1):85-98 [FREE Full text] [doi: 10.1007/s10865-016-9818-7] [Medline: 28058516]
22. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
23. Hingle M, Patrick H. There are thousands of apps for that: navigating mobile technology for nutrition education and behavior. J Nutr Educ Behav 2016 Mar;48(3):213-8.e1. [doi: 10.1016/j.jneb.2015.12.009] [Medline: 26965099]
24. Drumm J, White N, Swiegers M, Davey M. Deloitte. Mobile Consumer Survey 2017 URL: https://www2.deloitte.com/au/mobile-consumer-survey [accessed 2018-08-08] [WebCite Cache ID 73muqrfLD]
25. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
26. DiFilippo KN, Huang WH, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: a systematic literature review. J Telemed Telecare 2015 Jul;21(5):243-253. [doi: 10.1177/1357633X15572203] [Medline: 25680388]
27. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]

28. 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]
29. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
30. Tonkin E, Brimblecombe J, Wycherley TP. Characteristics of smartphone applications for nutrition improvement in community settings: a scoping review. Adv Nutr 2017 Mar;8(2):308-322 [FREE Full text] [doi: 10.3945/an.116.013748] [Medline: 28298274]
31. 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;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
32. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular nutrition-related mobile apps: a feature assessment. JMIR Mhealth Uhealth 2016 Aug 01;4(3):e85 [FREE Full text] [doi: 10.2196/mhealth.5846] [Medline: 27480144]
33. Flaherty S, McCarthy M, Collins A, McAuliffe F. Can existing mobile apps support healthier food purchasing behaviour? Content analysis of nutrition content, behaviour change theory and user quality integration. Public Health Nutr 2018 Feb;21(2):288-298. [doi: 10.1017/S1368980017002889] [Medline: 29081322]
34. Schoeppe S, Alley S, Rebar AL, Hayman M, Bray NA, Van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. Int J Behav Nutr Phys Act 2017 Dec 24;14(1):83 [FREE Full text] [doi: 10.1186/s12966-017-0538-3] [Medline: 28646889]
35. Burrows TL, Khambalia AZ, Perry R, Carty D, Hendrie GA, Allman-Farinelli MA, et al. Great 'app-eal' but not there yet: a review of iPhone nutrition applications relevant to child weight management. Nutr Diet 2015 Apr 16;72(4):363-367 [FREE Full text] [doi: 10.1111/1747-0080.12184]
36. Taki S, Campbell KJ, Russell CG, Elliott R, Laws R, Denney-Wilson E. Infant feeding websites and apps: a systematic assessment of quality and content. Interact J Med Res 2015 Sep 29;4(3):e18 [FREE Full text] [doi: 10.2196/ijmr.4323] [Medline: 26420339]
37. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
38. Stoyanov SR. YouTube. 2016. MARS training video URL: https://www.youtube.com/watch?v=25vBwJQIOcE [accessed 2018-08-06] [WebCite Cache ID 71SppRTEx]
39. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
40. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23-34 [FREE Full text] [doi: 10.20982/tqmp.08.1.p023}] [Medline: 22833776]
41. 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]
42. BCT Taxonomy. London: University College London; 2018. BCT Taxonomy v1 Online Training URL: http://www.bct-taxonomy.com/ [accessed 2018-08-06] [WebCite Cache ID 71SqJxqgN]
43. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. J Med Internet Res 2015 May 14;17(5):e118 [FREE Full text] [doi: 10.2196/jmir.4060] [Medline: 25977135]
44. 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]

Abbreviations

BCT: behavior change technique
BCTTv1: behavior change technique taxonomy version 1
ICC: intraclass correlation coefficient
MARS: Mobile App Rating Scale
PABAK: prevalence adjusted and bias adjusted kappa

Edited by G Eysenbach; submitted 26.08.18; peer-reviewed by M Kebede, P Krebs; comments to author 26.09.18; revised version received 09.10.18; accepted 09.10.18; published 21.12.18.

Please cite as:
Mauch CE, Wycherley TP, Laws RA, Johnson BJ, Bell LK, Golley RK
Mobile Apps to Support Healthy Family Food Provision: Systematic Assessment of Popular, Commercially Available Apps
JMIR Mhealth Uhealth 2018;6(12):e11867
URL: http://mhealth.jmir.org/2018/12/e11867/
doi:10.2196/11867
PMID:30578213

Original Paper

Mobile Phone, Computer, and Internet Use Among Older Homeless Adults: Results from the HOPE HOME Cohort Study

Maria C Raven[1,2], MSc, MPH, MD; Lauren M Kaplan[3,4], PhD; Marina Rosenberg[5], BA; Lina Tieu[3,4], MPH; David Guzman[3,4], MSPH; Margot Kushel[2,3,4], MD

[1]Department of Emergency Medicine, School of Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco, San Francisco, CA, United States

[3]Center for Vulnerable Populations, Zuckerberg San Francisco General Hospital and Trauma Center, University of California, San Francisco, San Francisco, CA, United States

[4]Division of General Internal Medicine, Department of Medicine, Zuckerberg San Francisco General Hospital and Trauma Center, University of California, San Francisco, San Francisco, CA, United States

[5]School of Public Health, University of California, Berkeley, Berkeley, CA, United States

Corresponding Author:
Maria C Raven, MSc, MPH, MD
Department of Emergency Medicine
School of Medicine
University of California, San Francisco
505 Parnassus Avenue
San Francisco, CA, 94143 0209
United States
Phone: 1 917 499 5608
Email: maria.raven@ucsf.edu

Abstract

Background: The median age of single homeless adults is approximately 50 years. Older homeless adults have poor social support and experience a high prevalence of chronic disease, depression, and substance use disorders. Access to mobile phones and the internet could help lower the barriers to social support, social services, and medical care; however, little is known about access to and use of these by older homeless adults.

Objective: This study aimed to describe the access to and use of mobile phones, computers, and internet among a cohort of 350 homeless adults over the age of 50 years.

Methods: We recruited 350 participants who were homeless and older than 50 years in Oakland, California. We interviewed participants at 6-month intervals about their health status, residential history, social support, substance use, depressive symptomology, and activities of daily living (ADLs) using validated tools. We performed clinical assessments of cognitive function. During the 6-month follow-up interview, study staff administered questions about internet and mobile technology use. We assessed participants' comfort with and use of multiple functions associated with these technologies.

Results: Of the 343 participants alive at the 6-month follow-up, 87.5% (300/343) completed the mobile phone and internet questionnaire. The median age of participants was 57.5 years (interquartile range 54-61). Of these, 74.7% (224/300) were male, and 81.0% (243/300) were black. Approximately one-fourth (24.3%, 73/300) of the participants had cognitive impairment and slightly over one-third (33.6%, 100/300) had impairments in executive function. Most (72.3%, 217/300) participants currently owned or had access to a mobile phone. Of those, most had feature phones, rather than smartphones (89, 32.1%), and did not hold annual contracts (261, 94.2%). Just over half (164, 55%) had ever accessed the internet. Participants used phones and internet to communicate with medical personnel (179, 64.6%), search for housing and employment (85, 30.7%), and to contact their families (228, 82.3%). Those who regained housing were significantly more likely to have mobile phone access (adjusted odds ratio [AOR] 3.81, 95% CI 1.77-8.21). Those with ADL (AOR 0.53, 95% CI 0.31-0.92) and executive function impairment (AOR 0.49; 95% CI 0.28-0.86) were significantly less likely to have mobile phones. Moderate to high risk amphetamine use was associated with reduced access to mobile phones (AOR 0.27, 95% CI 0.10-0.72).

Conclusions: Older homeless adults could benefit from portable internet and phone access. However, participants had a lower prevalence of smartphone and internet access than adults aged over 65 years in the general public or low-income adults. Participants

faced barriers to mobile phone and internet use, including financial barriers and functional and cognitive impairments. Expanding access to these basic technologies could result in improved outcomes.

(JMIR Mhealth Uhealth 2018;6(12):e10049) doi:10.2196/10049

KEYWORDS

homelessness; internet; cell phone; smartphone; aged and middle aged

Introduction

Background

In the past 25 years, the median age of individuals experiencing homelessness in the United States has risen [1]. Approximately half of the single adult homeless population is aged 50 years and older [1]. Adults with a current or recent experience of homelessness (homeless-experienced) have a high prevalence of chronic disease, functional and cognitive impairment, and substance use [2-4]. Homeless-experienced older adults' competing needs for food and shelter, lack of stable mailing address, and limited social support complicate the receipt of longitudinal health care needed to manage these conditions [2,5].

Appropriate longitudinal health care relies on intervisit communication [6-9]. Mobile phones, email, and patient portals increase the consistency of intervisit communication between patients and clinicians and improve self-management of chronic diseases in the general population [5,10-14]. None of these requires a permanent address, and therefore, they could be used by people experiencing homelessness [11,15,16].

In addition to improving health care communication, these technologies have other potential health benefits for homeless individuals, including decreasing social isolation, connecting to social services, and identifying housing resources [17-21]. However, little is known about how older homeless-experienced adults use mobile and internet technologies.

Low-income housed individuals report barriers to technology use, such as lack of high-speed broadband access, limited English proficiency, and limited digital and linguistic literacy [14,22,23]. Low-income populations rely on smartphones, rather than computers, for internet access [14]. Older adults in the general population use technology at lower rates than younger adults [24,25]. Cost; low digital literacy; and cognitive, executive, and sensory impairments may limit use in this population [24-27].

Objectives

The limited literature about mobile phone and internet access among homeless populations has focused on younger populations [18,28]. Little is known about the use of mobile phones and internet by older adults who experience homelessness. In a population-based cohort of 350 homeless-experienced adults aged 50 years and older, we examined the prevalence of mobile phone (smartphones and feature phones), computer and internet access, purposes of use, types of service contracts and charging locations, and the factors associated with access to mobile phones.

Methods

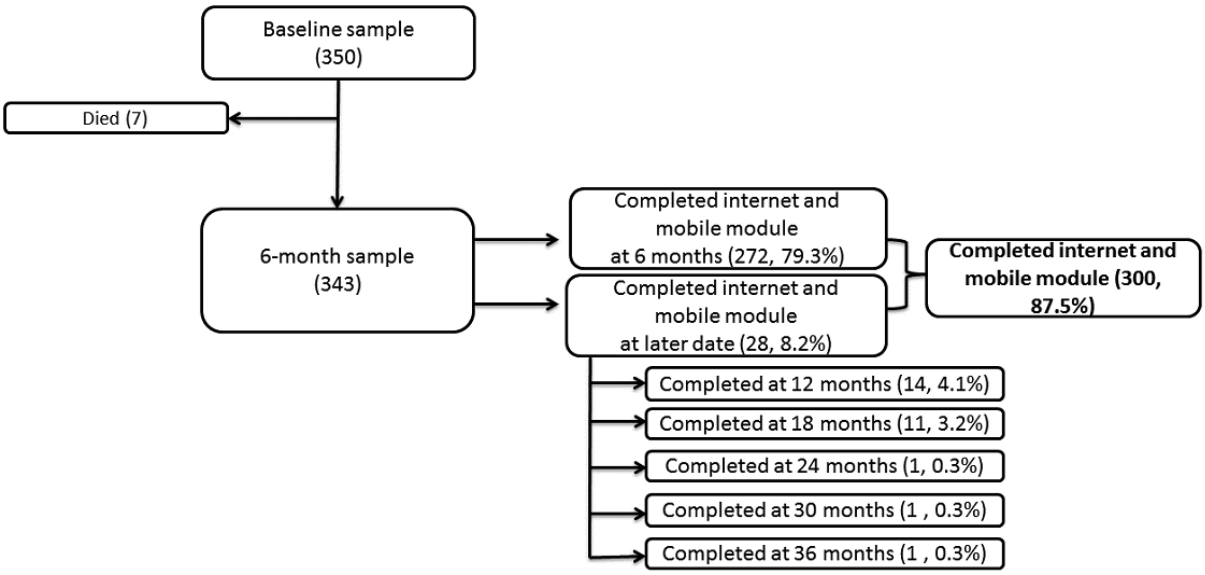
Participants and Setting

The overall goal of the Health Outcomes of People Experiencing Homelessness in Older Middle Age (HOPE HOME) study is, among older homeless adults, to describe the life course events and level of geriatric conditions and to explore the association between life course events and geriatric conditions with acute health service utilization. Between July 2013 and June 2014, we used population-based sampling to recruit 350 homeless individuals aged 50 years or older in Oakland, California [29]. We recruited from homeless encampments, all overnight homeless shelters that served single adults over the age of 25 years (n=5), one recycling center close to homeless service agencies, and all free and low-cost meal programs serving at least 3 prepared meals a week (n=5). We constructed our sampling frame to approximate the source population; we randomly selected potential participants at each recruitment site [30,31].

After an initial screening for eligibility, we invited individuals to complete a detailed eligibility interview within 1 week. Participants were eligible for the study if they were English-speaking, aged 50 years or older, defined as homeless by the Homeless Emergency Assistance and Rapid Transition to Housing Act (HEARTH), [32] and able to give written informed consent as determined by a teach-back method [33]. We gave participants gift certificates worth US $25 for completing the screening and baseline interviews and US $20 for each semiannual follow-up visit. We gave participants a US $5 gift certificate for each monthly check-in between scheduled visits. The majority of study activities took place at St Mary's Center, a nonprofit community-based organization serving low-income older adults. The institutional review board of the University of California, San Francisco approved the study.

Trained study staff administered structured baseline interviews and follow-up interviews at 6-month intervals. At the initial interview, study staff collected extensive contact information on participants, including a phone number, if the participant had one. Participants checked in monthly between study visits, by phone or in person, to enhance the follow-up process. During structured interviews at baseline and follow-up, participants reported information about housing history, demographic information, health history, health care utilization, drug and alcohol use, mental health, and social support, and completed assessments of functional and cognitive impairment. Participants remained in the study independent of their housing status at the time of follow-up. During the 6-month follow-up interview or, if missed, the next attended interview, study staff administered a module centered on the use of internet and mobile technology, as shown in Figure 1.

Figure 1. Recruitment flowchart.



In this analysis, we use all time-varying variables at the interview at which the participant completed the internet and mobile technology module. To assess differential loss to follow-up, we assessed whether participants who were eligible for, but did not complete, a mobile phone and internet interview were less likely to report having a phone number at enrollment than those who completed the interview.

Measures

Demographics

Demographic variables included age, sex, and race or ethnicity (black, white, Hispanic or Latino, Asian, other or mixed). We dichotomized participants as having completed a high school or General Educational Development (GED) degree versus having completed less than a high school equivalent degree. Participants reported their total income in the past 30 days, categorized as US $0-$150, $151-$700, $701-$1150, and over $1150. To assess health literacy, we used a validated one-item health literacy screen "How confident are you filling out medical forms by yourself?" (Not at all, A little bit, Somewhat, Quite a bit, Extremely) [34]. On the basis of validation studies within low-income populations, we considered those who reported being somewhat confident or less as having limited health literacy [35,36].

Focal Variables

Focal Independent Variable: Housing Status

At each interview, we determined whether participants still met the HEARTH criteria for homelessness, categorizing the participants' current living situation as homeless, housed, or in an institution. As participants were either currently homeless, or had recently been homeless, and in keeping with the transient nature of homelessness, we described the sample as homeless-experienced [14].

Mobile Phone Access, Use, and Service Type

Participants reported if they had ever used a mobile phone (feature, smartphone, or both). We adapted Pew survey items based on prior research on information technologies among homeless populations [37,38]. We defined feature phones as phones allowing users to "make and receive phone calls and text messages, take pictures and perform basic Web browsing." We defined smartphones as "a phone with a larger screen that allows functions like a mini computer and lets you check your email and use a number of different applications." We asked participants if they had ever used a mobile phone; if yes, we asked whether they had current access to a mobile phone or had access in the past. We defined having access to a mobile phone as owning a mobile phone, borrowing one long term, being able to borrow one if needed, or being able to find one in an emergency. Participants reported whether they had current access to mobile phones, past access, or never had access. Our focal dependent variable was current access to a mobile phone.

We asked participants to report what type of mobile service they used (contract, month-to-month, prepaid, free phone, or other). If participants reported ever having access to a mobile phone, we asked them to report what they used it for (making phone calls, receiving phone calls, voicemail, or text messages). If participants had ever used a smartphone, we asked them if they used it to check and send email, access social networking sites, look up information on the internet, look up bus schedules, or get directions. We asked participants to report whether they used a mobile phone to contact others, and if so, whom they contacted. We asked participants if they had ever had their mobile phones stolen. If participants reported having had their phones stolen, we asked them how many times. We asked participants if they had ever lost a mobile phone. If they reported losing a phone, we asked how many times.

Ease of Use and Charging Locations

We asked participants to report, on a 5-point Likert scale (1=very easy to 6=I don't know how to do this), how comfortable they were with performing the following actions

on a mobile phone: making a call, answering a call, contacting 911 or emergency medical services, checking voicemail, and using text messaging. Participants rated the difficulty of using basic components of their phone, such as the buttons and screen. We asked participants where they charged their phones. To assess barriers to phone charging, we asked participants whether there were times they had not had mobile phone service because they did not have a place to charge their phones.

Computer, Internet, and Email Use

We asked participants if they had ever used a computer. If so, we asked if they had ever used the internet. We asked those who had ever used the internet if they had done so in the past 30 days. Among those with recent use, we asked where they used the internet and what they used the internet for. Potential venues included the following: on a mobile phone, in a public or university library, drop-in center or shelter, friend or relative's house, internet café, coffee shop or restaurant, social service agency, motel or hotel lobby, church, and others. Uses included reading or sending email; getting news online; watching a video, downloading a music file or playing a game; browsing the internet for fun; searching for a fact or to answer a question; looking for information about a shelter or place to live, a hobby or interest, health or medical information, or about a job; checking social networking sites; doing research for school, training, or education; sending instant messages; refilling a prescription; and looking for a sex partner. For each of these response categories, we asked participants to note all those that were applicable.

We asked participants whether they had ever used email and if they had a current email account. We asked what they used their email for: staying in touch with family or friends, job searches, housing searches, staying in touch with a case manager or social worker, staying in touch with a health care provider, and other. We asked participants to note all that were applicable.

Descriptive Variables

Health History

We asked participants to rate their health status, dichotomized as poor or fair versus good, very good, or excellent [39]. On the basis of the National Health and Nutrition Examination Survey, we asked participants to report whether a health care provider had ever told them they had any of the 10 chronic conditions [3,40]. We created a composite variable for the total number of chronic conditions, categorized as none, 1, 2, or 3 or more. We asked participants if they had difficulty performing any activities of daily living (ADL): dressing, bathing or showering, eating, getting in or out of bed, or using the toilet [41]. We dichotomized participants as having any difficulty versus no ADL difficulty.

We administered the Modified Mini-Mental (3MS) Examination to assess global cognitive impairment [42] and the Trail Making Test B (Trails B) [43] to assess executive function. Comparing scores with age- and education-adjusted reference values, we categorized scores below the seventh percentile on the 3MS as cognitive impairment [44]. We classified the participants' performance as "unable to complete" if their time to complete the Trails B lasted longer than 5 min. We interpreted scores with the demographically adjusted (age, gender, and race or ethnicity) norms for the Halstead-Reitan Neuropsychological Test Battery, which uses the Halstead-Reitan Battery (HRB) Norms scoring program [45].

Mental Health, Substance Use, and Social Support

We assessed depressive symptoms using the Center for Epidemiologic Studies Depression Scale (CES-D) [46]. On the basis of past studies with older adults, we classified scores of ≥22 as indicative of major depressive symptoms [47,48]. We considered participants who reported drinking ≥6 drinks on 1 occasion every month as heavy drinkers [49]. Using the World Health Organization's Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST) with a lengthened time frame of the past 6 months, we assessed illicit drug use for cocaine, opioids, or amphetamines [50]. We classified scores of ≥4 for any illicit drug as moderate to high risk use of an illicit substance. To assess social support, we asked participants how many close friends or relatives they had in whom they could confide (0, 1-5, or 6 or more) [51,52].

Analysis

We performed a descriptive analysis to assess the prevalence of mobile phone, computer, internet, and email use and the purposes of usage. To identify facilitators to mobile phone and internet use among older homeless adults, we assessed the ease of use, types of service, and charging locations. We assessed bivariable associations between current mobile phone use and a priori independent variables using logistic regression. We built our multivariable model by including variables with bivariable type 3 P values <.20, and reduced the model using backwards elimination retaining variables with P values <.05 for our final multivariable model. We implemented our models in SAS 9.4 (SAS Institute, Cary, NC).

Results

Follow-Up

Of the 350 individuals enrolled in the study, 7 died before the 6-month follow-up. Of the 343 participants alive by the 6-month follow-up, 300 (87.5% (300/343) completed the module on internet and mobile phone use. Of these, 79.3% (272/343) completed the mobile phone and internet module at 6 months and 8.2% (28/343) completed the module at a later date. (Figure 1). One-third (32.6%, 14/43) of those who were eligible but did not complete a mobile phone interview reported having a phone number at enrollment, compared with 68.0% (204/300) of those who completed the interview (P<.001).

Demographics

The median age of participants was 57.5 years (interquartile range 54-61). Most participants (74.7%) were male and black (81.0%; see Table 1). Approximately one-fourth had less than a high school or equivalent (eg, GED) education (24.7%). Most participants (74.3%) remained homeless at their follow-up interview. Approximately three-fourths (74.9%) reported having at least one confidant. Over ten percent (10.3%) reported heavy drinking, and approximately one-third met the criteria for moderate- to high-risk cocaine use (29.0%).

Table 1. Participant characteristics of mobile phone use.

Descriptive, health, and health-related variables	Total (N=300)	Currently own or have access to a mobile phone (N=217)	Owned or have access to a mobile phone (N=60)	Never owned or had access to a mobile phone in the past (N=23)
Age in years, median (interquartile range)	57.5 (54.0-61.0)	57.0 (54.0-61.0)	58.0 (54.0-61.0)	58.0 (55.0-65.0)
Male, n (%)	224 (74.7)	161 (74.2)	44 (73)	19 (83)
Black, n (%)	243 (81.0)	183 (84.3)	45 (75)	15 (65)
Completed less than high school degree[a], n (%)	74 (24.5)	57 (26.2)	10 (177)	7 (30)
Total income in past 30 days, n (%)				
US $0-150	62 (20.6)	41 (18.8)	13 (22)	8 (35)
US $151-700	76 (25.3)	55 (25.3)	19 (32)	2 (9)
US $701-1150	128 (42.7)	92 (42.3)	24 (40)	12 (52)
More than US $1150	34 (11.3)	29 (13.3)	4 (7)	1 (4)
Homeless at follow-up interview[b]	223 (74.3)	149 (68.7)	55 (92)	19 (83)
Social support, n (%)	224 (74.9)	169 (77.9)	40 (67)	15 (68)
Number of confidants[c]				
None	76 (25.3)	49 (22.5)	20 (33)	7 (32)
1	78 (26.0)	55 (25.3)	17 (28)	6 (27)
2	46 (15.3)	33 (15.2)	9 (15)	4 (18)
3 or more	99 (33.0)	80 (36.8)	14 (23)	5 (23)
Fair or poor health, n (%)	166 (56.1)	113 (52.8)	39 (66)	14 (61)
Number of chronic conditions, n (%)				
None	76 (25.3)	58 (26.7)	13 (22)	5 (22)
1	103 (34.3)	76 (35.0)	18 (30)	9 (39)
2	88 (29.3)	61 (28.1)	19 (32)	8 (35)
3 or more	33 (11.0)	22 (10.1)	10 (17)	1 (4)
Activities of daily living impairment, n (%)	128 (42.7)	85 (39.1)	31 (52)	12 (52)
Cognitive impairment (3MS, baseline)[d], n (%)	73 (24.3)	46 (21.1)	17 (28)	10 (44)
Executive function impairment (Trails B)[e], n (%)	100 (33.6)	65 (29.9)	27 (45)	8 (35)
Moderate-to-severe depressive symptoms[f], n (%)	94 (31.3)	63 (29.0)	22 (37)	9 (43)
Heavy drinking[g], n (%)	31 (10.3)	16 (7.3)	11 (4)	3 (13)
Moderate-to-high risk amphetamines use[h], n (%)	19 (6.3)	8 (3.6)	9 (15)	2 (9)
Moderate-to-high risk cocaine use[i], n (%)	87 (29.0)	59 (27.1)	20 (33)	8 (35)
Moderate-to-high risk opioids use[j], n (%)	19 (6.3)	12 (5.5)	5 (8)	2 (9)

[a]Completion of high school degree included General Education Development (GED).

[b]Homeless as defined by the Homeless Emergency Assistance and Rapid Transition to Housing (HEARTH) Act.

[c]Confidant defined as “a close friend or family member in whom you can confide or talk about yourself and your problems.”

[d]Modified Mini-Mental State Examination; less than seventh percentile based on Z-scores used.

[e]Trail Making Test; more than 5-min completion time on Trails B.

[f]Score of ≥22 on the Center for Epidemiologic Studies Depression Scale (CES-D).

[g]Drinking ≥6 drinks on one occasion every month.

[h]Score of ≥4 for any amphetamines using the World Health Organization’s Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST).

[i]Score of ≥4 for any cocaine using the World Health Organization’s Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST).

[j]Score of ≥4 for any opioids using the World Health Organization’s Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST).

Approximately one-third (31.8%) met the criteria for major depressive symptoms. Over half (56.1%) described their health as fair to poor and approximately three-fourths had at least one chronic condition (74.7%). Over 40% had ADL impairment (42.7%), approximately one-fourth had cognitive impairment (24.3%), and one-third had problems in executive functioning (33.6%).

Mobile Phone Access, Use, and Service Type

Almost all participants currently owned or had access to a mobile phone (72.3%) or had owned or had access to a mobile phone in the past (20.0%; see Table 2). Among those with current mobile phone access, 204 owned their phones, 4 borrowed phones for long term, 6 borrowed phones for short term, and 3 could access a mobile phone in an emergency. Among participants who currently or had ever had a phone (n=277), two-thirds had basic mobile phones, as opposed to smartphones.

More than three-fourths of participants with current or prior access to a phone (n=277) reported using phones to contact relatives (82.3%) and friends (77.6%; see Table 2). A majority of participants used phones to contact medical personnel (66.6%), and nearly half of them used phones to contact social service agencies (49.5%). Almost one-third used phones to contact shelters or other housing providers (30.7%). Approximately one-fourth used them to contact potential employers (23.1%). A smaller proportion used phones to contact potential landlords (19.5%; see Table 2). Over half reported having had their mobile phones stolen (53.1%) or lost (52.9%).

Ease of Use and Charging Locations

Over 80% of participants with experience with mobile phones reported that it was easy to use them (Table 3). Participants reported charging their phones at a variety of locations, most commonly at a relative or friend's (34.3%) or a drop-in center or shelter (32.5%; see Table 3). Over half (56.1%) of those with past access to mobile phones versus approximately one-third (31.6%) of those with current mobile phone access reported not having service due to not having a place to charge their phones.

Computer and Internet Use

A majority of the participants reported using a computer (64.8%) or accessing the internet (55.0%) during their lifetime (see Table 4). Approximately one-third of the participants had used a computer (37.9%) or the internet (39.3%) in the past 30 days. Participants accessed the internet from a variety of locations, most of which were public. They used the internet for multiple functions including email (24.8%) and looking for information about housing (16.8%), medical information (15.1%), or a job (14.4%; see Table 4).

Approximately one-third had a current email account (35.2%). The most common uses of email were staying in touch with family or friends and searching for jobs and housing (Table 4).

Factors Associated With Mobile Phone Access

In an adjusted multivariable regression model (see Table 5), we found that individuals who were housed at the time of this interview had 3.81 (adjusted odds ratio [AOR] 3.81, 95% CI 1.77-8.21) higher odds of currently owning a mobile phone, compared with those who were not housed (see Table 5). Moreover, 3 factors were associated with significantly lower odds of current mobile phone ownership: ADL impairment (AOR 0.53, 95% CI 0.31-0.92), executive function impairment (AOR 0.49, 95% CI 0.28-0.86), and moderate to high use of amphetamines (AOR 0.27, 95% CI 0.10-0.72).

Table 2. Mobile phone use.

Mobile phone use	Total (N=277), n (%)	Currently own or have access to a mobile phone (N=217), n (%)	Owned or had access to a mobile phone in the past (N=60), n (%)	*P* value
Type of phone use or used				
Feature phone	186 (67.1)	143 (65.9)	43 (72)	.57
Smartphone	89 (32.1)	72 (33.1)	17 (28)	—[a]
Both	2 (0.7)	2 (0.9)	0 (0)	—
Type of service				
Contract	16 (5.7)	13 (5.9)	3 (5)	.10
Month-to-month	167 (60.3)	130 (59.9)	37 (62)	—
Free phone	52 (18.7)	46 (21.1)	6 (10)	—
Prepaid	26 (9.3)	16 (7.3)	10 (17)	—
Other or don’t know	16 (5.7)	12 (5.5)	4 (7)	—
Mobile phone features used				
Make and receive phone calls	277 (100.0)	217 (100.0)	60 (100)	>.99
Check and receive voicemails	195 (70.4)	162 (74.7)	33 (55)	.003
Send and receive text messages	172 (62.9)	145 (66.8)	27 (45)	.002
Smartphone features used				
Look up information on the internet	66 (23.8)	55 (25.3)	11 (65)	.42
Check and send email	53 (19.1)	46 (21.1)	7 (41)	.11
Get directions	49 (17.6)	41 (18.8)	8 (47)	.53
Look up bus route or schedule	37 (13.3)	34 (15.6)	3 (18)	.03
Check social networking sites	24 (8.6)	23 (10.5)	1 (6)	.03
Uses of phone to contact others				
Use or used phone to contact relatives	228 (82.3)	178 (82.0)	50 (83)	.81
Use or used phone to contact friends	215 (77.6)	171 (78.8)	44 (73)	.37
Use or used phone to contact medical personnel	179 (64.6)	150 (69.1)	29 (48)	.003
Use or used phone to contact social service agencies	137 (49.5)	115 (53.0)	22 (37)	.03
Use or used phone to contact shelters or other housing providers	85 (30.6)	71 (32.7)	14 (23)	.16
Use or used phone to contact (potential) employer	64 (23.1)	51 (23.5)	13 (22)	.77
Use or used phone to contact (potential) landlord	54 (19.4)	44 (20.2)	10 (17)	.53
Use or used phone to contact emergency services	29 (10.4)	25 (11.5)	4 (7)	.28
Ever had phone stolen	146 (53.1)	106 (49.3)	40 (67)	.02
Number of times phone stolen				
0	129 (47.4)	109 (51.2)	20 (344)	.06
1-2	109 (40.1)	80 (36.8)	29 (49)	.06
≥3	34 (12.2)	24 (11.0)	10 (17)	.06
Ever lost phone	145 (52.9)	113 (52.6)	32 (54)	.81
Number of times lost phone				
0	129 (47.4)	102 (47.9)	27 (46)	.30
1-2	106 (39.0)	79 (36.4)	27 (46)	.30
≥3	37 (12.2)	32 (14.7)	5 (9)	.30

[a]Not applicable.

Table 3. Ease of using mobile phone features and charging among participants who had ever used a mobile phone (N=277).

Ease of use and charging locations	Total (N=277), n (%)	Currently own or have access to a mobile phone, (N=217), n (%)	Owned or had access to a mobile phone in the past, (N=60), n (%)	*P* value
Proportion reporting very easy to neither easy nor difficult, n (%)				
Punching buttons on the screen	240 (86.6)	191 (88.0)	49 (82)	.20
Seeing the phone screen	217 (78.3)	174 (80.2)	43 (72)	.16
Hearing the phone ring	239 (86.3)	189 (87.1)	50 (83)	.45
Hearing people talk	222 (80.1)	179 (82.5)	43 (72)	.06
Using voicemail	191 (69.0)	160 (73.7)	31 (52)	.001
Using other mobile phone features (eg, contacts)	198 (71.5)	167 (77.0)	31 (52)	<.001
Charging locations				
A friend or relative's house	95 (34.2)	81 (37.3)	14 (23)	.04
A drop-in center or homeless shelter	90 (32.4)	70 (32.2)	20 (33)	.87
A library	37 (13.3)	25 (11.5)	12 (20)	.09
A coffee shop or restaurant	39 (14.0)	24 (11.0)	15 (25)	.01
A city power supply	19 (6.8)	9 (4.1)	10 (17)	<.001
A social service or case management agency	18 (6.4)	13 (5.9)	5 (8)	.51
A place where you pay to charge your phone	3 (1.0)	1 (0.5)	2 (3)	.06
No service due to lack of a charging location[a]	98 (35.3)	66 (30.4)	32 (56)	<.001

[a]11 participants had missing data.

Table 4. Computer, internet, and email use.

Computer, internet, and email use	n (%)
Ever used a computer[a]	193 (64.8)
Currently use a computer	113 (37.9)
Ever used internet	164 (55.0)
Used internet, last 30 days	117 (39.3)
Venues where internet was used[b]	
Public or university library	59 (19.8)
On mobile phone	51 (17.1)
Drop-in center or homeless shelter	27 (9.1)
Friend or relative's house	21 (7.0)
Internet café, coffee shop, or restaurant	11 (3.4)
Social service agency	8 (2.7)
Motel or hotel lobby	3 (1.0)
Church	1 (0.3)
Workplace	2 (0.7)
Other venue	26 (8.7)
Purpose of using the internet	
Read or send email	74 (24.8)
Get news online	66 (22.1)
Watch a video, download a music file, or play a game	61 (20.5)
Browse the internet for fun	56 (18.8)
Search for a fact or to answer a question	50 (16.8)
Look for information about a shelter or place to live	50 (16.8)
Look for information about a hobby or interest	46 (15.4)
Look for health or medical information	45 (15.1)
Look for information about a job	43 (14.4)
Check social networking sites (eg, Facebook or Twitter)	43 (14.4)
Do research for school or training, or obtain education	25 (8.4)
Send instant messages	18 (6.0)
Refill a prescription	16 (5.4)
Look for a (sex) partner	4 (1.3)
Email	
Know what email is	234 (78.5)
Know how to use email	144 (48.3)
Have an email account	105 (35.2)
Uses of email[c]	
Stay in touch with family or friends	67 (22.5)
Job searches	46 (15.4)
Housing searches	40 (13.4)
Stay in touch with health care providers	21 (7.0)
Stay in touch with case manager or social workers	16 (5.4)
Other	21 (7.0)

[a]Two participants are not included in the computer and internet use section because they did not report whether they had ever used a computer, N=298

[b]The denominator for internet venues and uses is 117.

[c]The denominator for email variables ranges from 296.3 to 300.

Table 5. Odds of current mobile phone use.

Independent variables	Unadjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Black	2.06 (1.12-3.80)	.02	—[a]	—
Housed[b]	3.75 (1.76-7.99)	<.001	3.81 (1.77-8.21)	<.001
Health history				
Good to excellent health	1.63 (0.96-2.78)	.07	—	—
ADL impairment[c]	0.60 (0.36-1.00)	.05	0.53 (0.31-0.92)	.02
Cognitive impairment (3MS, baseline)[d]	0.56 (0.32-0.99)	.04	—	—
Executive function impairment (Trails B)[e]	0.59 (0.35-1.01)	.05	0.49 (0.28-0.86)	.01
Heavy drinking[f]	0.58 (0.32-1.02)	.06	—	—
Moderate to high risk amphetamine use[g]	0.25 (0.10-0.65)	.004	0.27 (0.10-0.72)	.01

[a]Not applicable.

[b]Not homeless as defined by the Homeless Emergency Assistance and Rapid Transition to Housing (HEARTH) Act.

[c]Difficulty performing one or more activities of daily living (ADL): dressing, bathing or showering, eating, getting in or out of bed, or using the toilet.

[d]Modified Mini-Mental State Examination; less than seventh percentile based on Z-scores used.

[e]Trail Making Test; more than 5-min completion time on Trails B.

[f]Drinking ≥6 drinks on one occasion every month.

[g]Score of ≥4 for amphetamine use using the World Health Organization's Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST).

Discussion

Principal Findings

In a sample of homeless-experienced adults aged 50 years and older, almost three-fourths of participants had current access to a mobile phone. Participants with phones used them for social support and communication with health care providers; however, few had annual phone contracts. Without annual contracts, it is likely that participants' phone numbers changed frequently, limiting the utility for two-way communication.

Although compared with the general population, there was less use of the internet, a high proportion who reported ever having used the internet had used it in the prior 30 days, suggesting that individuals who had knowledge about the internet used it regularly [53]. Participants who did use the internet in the prior month used it to get directions, bus schedules, and to obtain information on employment and housing—all of which can be invaluable for individuals experiencing homelessness [5,54].

When we recruited our participants, all were homeless. A significantly higher proportion of those who were housed at the time of this interview had current access to a mobile phone. Those with current mobile phone access were significantly more likely to use phones to contact shelters or other housing providers than those without current access [55]. In our multivariable analysis, we found that being housed was significantly associated with current mobile phone ownership. A stable living situation may enable more consistent mobile phone ownership, or access to a mobile phone may have helped participants regain housing. Poor executive and cognitive functions and moderate to high risk amphetamine use were negatively associated with current mobile phone ownership. Each of these 3 factors can adversely affect an individual's ability to participate in the type of anticipatory planning and organization required to obtain and maintain technology—even simple technology such as mobile phones. A majority of participants had lost or had their phones stolen, reflecting one of many adversities in the experience of homelessness.

Limitations

Our study has several limitations. First, our estimates of mobile phone access were likely overestimates. We introduced the mobile phone and internet module at the first follow-up interview after the baseline interview. There was differential loss to follow-up. Participants without phones at enrollment were less likely to complete the follow-up interview. Second, not all participants remained homeless at the time of the interview; those with housing were more likely to have phones than those without housing. Finally, we used a liberal definition of access to mobile phones, including those who owned or borrowed phones, or had access to one. We relied on participants to self-report mobile phone and internet use and did not have any way to verify these reports with phone bills, direct observation, or other methods.

Comparison With Prior Work

Participants' access to mobile phones and the internet was much lower than the general population, of whom 95% own mobile phones (77% of which are smartphones) and 90% of whom use

the internet [53]. Participants in our study had lower rates of smartphone and internet use when compared with low-income adults of any age [53,56,57]. Of the minority of participants who were able to access the internet, they accessed it most commonly via smartphones and public libraries. The prevalence of internet access via smartphones in our study was lower than that of those with low incomes in the general population [53]. A majority of participants reported having phones stolen and lost. Having assets stolen or lost is a common experience for people experiencing homelessness. If phones increase the risk of robbery, they may present a safety hazard for older homeless adults.

There are several ways in which older homeless adults could benefit from internet and phone access. Participants used these technologies for health care communication and to seek housing and employment information. Increasing internet and mobile phone access among older homeless adults could allow older homeless adults to more easily apply for housing or to search for housing in areas outside of urban centers that may be lower-cost. Internet and mobile phone access could also facilitate contact with potential employers and increase access to employment and social networking sites.

Mobile phones can facilitate communication with family or friends who may be able to provide instrumental as well as emotional support [17]. Social support has been shown to be associated with better health [58]. In addition, homeless individuals need low-barrier access to outpatient primary care; mobile phones and internet access could facilitate this. A pilot study that examined the feasibility and potential efficacy of using text messages to remind homeless veterans about appointments found that the veterans liked receiving the messages, and those messages may have improved appointment attendance [59]. Two-thirds of our participants reported using their phones to communicate with their health care providers, suggesting both interest and feasibility.

Our participants did not have annual phone contracts. This limited the possibilities for bidirectional communication due to interruptions in service and changing phone numbers. Previous research has cited barriers to mobile phone use among homeless individuals, including cost, fear of loss or theft, and a lack of knowledge about how to use mobile technology [19,37]. The widespread use of month-to-month, instead of annual plans, the use of borrowed (instead of owned) phones, and the relatively low proportion of people who had current access to phones may reflect these barriers, particularly cost. Although there are some programs to address financial barriers to mobile phone use among low-income populations, we found participants had low rates of enrollment in such programs. The “Lifeline” program provides Federal Assistance recipients and those who provide proof of low income with free feature or smartphones and pays for voice calls and texting for a year, with the possibility of recertification [60]. Although most of our participants met the criteria for this program, few reported using its free mobile phones and service. The Lifeline program requires a mailing address. Many people experiencing homelessness lack a stable mailing address, which could cause phone service interruptions. Without the widespread adoption of phone contracts by homeless adults, health care providers should consider open access scheduling, which could allow homeless adults with any form of phone access to make appointments, while acknowledging their inability to receive appointment reminder calls and texts. Open access scheduling allows for same day appointments and does not rely on reminder calls for appointments scheduled far in advance. This could lower access barriers for individuals experiencing homelessness who may have minimal or no access to mobile phones and the internet. In addition, allowing mobile phone users to maintain the same phone number despite interruptions in service could increase their ability to communicate with health care providers.

Participants in our study did not report difficulty with using phone buttons or keyboard. However, impairments in ADLs and executive function were associated with lower odds of current mobile phone use. Given the levels of these impairments among our participants, more research is needed to match end users with appropriate training tools and technology. It is possible that many who use feature phones could make use of smartphones with appropriate access and training. Others may require improved access and training to make use of feature phone technology.

Another possibility is that impairments in ADLs and executive functioning indirectly decrease use of mobile phones by making it difficult to obtain mobile phones and maintain service. Participants without phones reported a higher likelihood of losing service due to not having a place to charge their phone. Therefore, multipronged approaches that include increasing access to phones, charging stations, and internet might be most effective in increasing the adoption of mobile technology among older adults experiencing homelessness.

Increased public access to high-speed internet and providing discounted smartphones for high-need, low-income individuals may increase access to the internet [61]. Private sector technology and telecommunication companies might also be incentivized to fund initiatives that increase the use of their services among underserved populations, increasing access to reliable mobile technology [61]. Older adults comprise an increasing proportion of the US population. One way for technology companies to increase adoption of mobile phones for older adults is to include them in participatory design and usability testing [62,63]. Adapting devices and tailoring online advanced features to meet the needs of older homeless adults could facilitate their inclusion in the digital economy.

Conclusions

This study is one of the first studies to examine mobile phone and internet use among a community-based sample of homeless adults over the age of 50 years [64]. The majority of participants with access to technology were able to take advantage of most mobile phone functions, although most of their mobile phones were feature phones with limited internet access. Participants used these technologies for health care communication, seeking information for housing, and looking for employment opportunities.

However, most participants did not have annual phone contracts—which can lead to new phone numbers with each new phone—and few had access to smartphones. Lowering financial barriers to allow annual mobile phone contracts and increasing the homeless individuals' ability to access the internet via smartphones could promote more reliable and widespread use of these basic technologies. In addition, providers can take steps to optimize the technology individuals experiencing homelessness have access to, by offering open access and same-day scheduling and communication. More research is needed to determine if increasing access to mobile phones and internet can positively impact downstream health and economic outcomes among individuals experiencing homelessness.

The high prevalence of functional and executive function impairment in our study population was negatively associated with access to mobile phones. Advanced technological features might be challenging for this segment of the homeless-experienced population. Initiatives to increase access to technology among older homeless adults must address the needs of those with impairments and create technological features that fit the individuals' needs and abilities.

Acknowledgments

The authors gratefully acknowledge their colleagues Claudia Ponath, Pamela Olsen, Nina Fiellin, Tauni Marin, and Kenneth Perez for their invaluable contributions to the Health Outcomes of People Experiencing Homelessness in Older Middle Age (HOPE HOME) study. They also acknowledge their colleague Angela Allen, who passed away in May 2015, for her incredible contributions to the study. The authors also thank the staff at St. Mary's Center and the HOPE HOME Community Advisory board for their guidance and partnership. This study was funded by grants from the National Institute on Aging at the National Institutes of Health [K24AG046372, R01AG041860, and P30AG044281] and by the National Institute on Minority Health and Health Disparities [R25MD006832]. These funding sources had no role in the preparation, review, or approval of the manuscript.

Conflicts of Interest

None declared.

References

1. Hahn JA, Kushel MB, Bangsberg DR, Riley E, Moss AR. BRIEF REPORT: the aging of the homeless population: fourteen-year trends in San Francisco. J Gen Intern Med 2006 Jul;21(7):775-778 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00493.x] [Medline: 16808781]
2. Brown RT, Kimes RV, Guzman D, Kushel M. Health care access and utilization in older versus younger homeless adults. J Health Care Poor Underserved 2010 Aug;21(3):1060-1070. [doi: 10.1353/hpu.0.0344] [Medline: 20693744]
3. Brown RT, Hemati K, Riley ED, Lee CT, Ponath C, Tieu L, et al. Geriatric conditions in a population-based sample of older homeless adults. Gerontologist 2017 Dec 01;57(4):757-766 [FREE Full text] [doi: 10.1093/geront/gnw011] [Medline: 26920935]
4. Spinelli MA, Ponath C, Tieu L, Hurstak EE, Guzman D, Kushel M. Factors associated with substance use in older homeless adults: results from the HOPE HOME study. Subst Abus 2017;38(1):88-94 [FREE Full text] [doi: 10.1080/08897077.2016.1264534] [Medline: 27897965]
5. Kushel MB, Gupta R, Gee L, Haas JS. Housing instability and food insecurity as barriers to health care among low-income Americans. J Gen Intern Med 2006 Jan;21(1):71-77 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.00278.x] [Medline: 16423128]
6. Arora S, Burner E, Terp S, Nok LC, Nercisian A, Bhatt V, et al. Improving attendance at post-emergency department follow-up via automated text message appointment reminders: a randomized controlled trial. Acad Emerg Med 2015 Jan;22(1):31-37 [FREE Full text] [doi: 10.1111/acem.12503] [Medline: 25388481]
7. da Costa TM, Salomão PL, Martha AS, Pisa IT, Sigulem D. The impact of short message service text messages sent as appointment reminders to patients' cell phones at outpatient clinics in São Paulo, Brazil. Int J Med Inform 2010 Jan;79(1):65-70. [doi: 10.1016/j.ijmedinf.2009.09.001] [Medline: 19783204]
8. Stubbs ND, Geraci SA, Stephenson PL, Jones DB, Sanders S. Methods to reduce outpatient non-attendance. Am J Med Sci 2012 Sep;344(3):211-219. [doi: 10.1097/MAJ.0b013e31824997c6] [Medline: 22475731]
9. Wright A, Poon EG, Wald J, Feblowitz J, Pang JE, Schnipper JL, et al. Randomized controlled trial of health maintenance reminders provided directly to patients through an electronic PHR. J Gen Intern Med 2012 Jan;27(1):85-92 [FREE Full text] [doi: 10.1007/s11606-011-1859-6] [Medline: 21904945]
10. Schickedanz A, Huang D, Lopez A, Cheung E, Lyles CR, Bodenheimer T, et al. Access, interest, and attitudes toward electronic communication for health care among patients in the medical safety net. J Gen Intern Med 2013 Jul;28(7):914-920 [FREE Full text] [doi: 10.1007/s11606-012-2329-5] [Medline: 23423453]
11. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: a systematic review of the evidence. J Diabetes Complications 2014;28(2):243-251 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.07.003] [Medline: 24332469]

12. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
13. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
14. Woods SS, Forsberg WC, Schwartz CE, Nazi M, Hibbard HJ, Houston KT, et al. The association of patient factors, digital access, and online behavior on sustained patient portal use: A prospective cohort of enrolled users. J Med Internet Res 2017;19(10):A. [Medline: 29042345]
15. Genz A, Kirk G, Piggott D, Mehta SH, Linas BS, Westergaard RP. Uptake and acceptability of information and communication technology in a community-based cohort of people who inject drugs: implications for mobile health interventions. JMIR Mhealth Uhealth 2015 Jun 25;3(2):e70 [FREE Full text] [doi: 10.2196/mhealth.3437] [Medline: 26111915]
16. Stennett CR, Weissenborn MR, Fisher GD, Cook RL. Identifying an effective way to communicate with homeless populations. Public Health 2012 Jan;126(1):54-56. [doi: 10.1016/j.puhe.2011.09.020] [Medline: 22136698]
17. Rice E, Kurzban S, Ray D. Homeless but connected: the role of heterogeneous social network ties and social networking technology in the mental health outcomes of street-living adolescents. Community Ment Health J 2012 Dec;48(6):692-698 [FREE Full text] [doi: 10.1007/s10597-011-9462-1] [Medline: 22075769]
18. Rice E, Barman-Adhikari A. Internet and social media use as a resource among homeless youth. J Comput Mediat Commun 2014 Jan 01;19(2):232-247 [FREE Full text] [doi: 10.1111/jcc4.12038] [Medline: 25328374]
19. McInnes DK, Fix GM, Solomon JL, Petrakis BA, Sawh L, Smelson DA. Preliminary needs assessment of mobile technology use for healthcare among homeless veterans. PeerJ 2015;3:e1096 [FREE Full text] [doi: 10.7717/peerj.1096] [Medline: 26246964]
20. McInnes DK, Sawh L, Petrakis BA, Rao SR, Shimada SL, Eyrich-Garg KM, et al. The potential for health-related uses of mobile phones and internet with homeless veterans: results from a multisite survey. Telemed J E Health 2014 Sep;20(9):801-809. [doi: 10.1089/tmj.2013.0329] [Medline: 25046280]
21. Rice E, Lee A, Taitt S. Cell phone use among homeless youth: potential for new health interventions and research. J Urban Health 2011 Dec;88(6):1175-1182 [FREE Full text] [doi: 10.1007/s11524-011-9624-z] [Medline: 22076445]
22. Zhang Y, Lauche R, Sibbritt D, Olaniran B, Cook R, Adams J. Comparison of health information technology use between American Adults with and without chronic health conditions: findings from the National Health Interview Survey 2012. J Med Internet Res 2017 Oct 05;19(10):e335 [FREE Full text] [doi: 10.2196/jmir.6989] [Medline: 28982644]
23. Ackerman SL, Sarkar U, Tieu L, Handley MA, Schillinger D, Hahn K, et al. Meaningful use in the safety net: a rapid ethnography of patient portal implementation at five community health centers in California. J Am Med Inform Assoc 2017 Mar 15. [doi: 10.1093/jamia/ocx015] [Medline: 28340229]
24. Arthur T, Jerzy L. The internet , ecommerce and older people: an actor network approach to researching reasons for adoption and use. Logistics Inf Manage 2003:56-63. [doi: 10.1108/09576050310453741]
25. Wildenbos GA, Peute L, Jaspers M. Facilitators and barriers of electronic health record patient portal adoption by older adults: a literature study. Stud Health Technol Inform 2017;235:308-312. [doi: 10.3233/978-1-61499-753-5-308] [Medline: 28423804]
26. Lee C, Coughlin JF. Perspective: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Prod Innov Manag 2014 Jun 03;32(5):747-759. [doi: 10.1111/jpim.12176]
27. Olphert C, Damodaran L, May A. Towards digital inclusion - engaging older people in the 'digital world'. In: Proceedings of the 2005 international conference on Accessible Design in the Digital World.: BCS Learning & Development Ltd; 2005 Presented at: 2005 international conference on Accessible Design in the Digital World; August 23-25, 2005; Dundee, Scotland p. 17-17 URL: https://ewic.bcs.org/content/ConWebDoc/3825
28. Jennings L, Lee N, Shore D, Strohminger N, Allison B, Conserve DF, et al. US minority homeless youth's access to and use of mobile phones: implications for mHealth intervention design. J Health Commun 2016 Jul;21(7):725-733. [doi: 10.1080/10810730.2015.1103331] [Medline: 27232544]
29. Burnam M, Koegel P. Methodology for obtaining a representative sample of homeless persons: the Los-Angeles skid row study. Eval Rev 1988;12(2):117-152. [doi: 10.1177/0193841X8801200202]
30. Lee CT, Guzman D, Ponath C, Tieu L, Riley E, Kushel M. Residential patterns in older homeless adults: results of a cluster analysis. Soc Sci Med 2016 Mar;153:131-140 [FREE Full text] [doi: 10.1016/j.socscimed.2016.02.004] [Medline: 26896877]
31. Henry M, Watt R, Rosenthal L, Shivji A. Washington: DC: The U.S. Department of Housing and Urban Development: Office of Community Planning and Development. 2017. The 2017 Annual Homeless Assessment Report (AHAR) to Congress URL: https://www.hudexchange.info/resources/documents/2017-AHAR-Part-1.pdf [WebCite Cache ID 72VZpP5hS]
32. US Congress. 2009. S. 896 HEARTH Act of 2009 URL: https://www.hudexchange.info/resource/1717/s-896-hearth-act/ [accessed 2018-01-01] [WebCite Cache ID 6x70x1qRc]
33. Dunn LB, Jeste DV. Enhancing informed consent for research and treatment. Neuropsychopharmacology 2001 Jun;24(6):595-607 [FREE Full text] [doi: 10.1016/S0893-133X(00)00218-9] [Medline: 11331139]
34. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]

35. Sarkar U, Schillinger D, López A, Sudore R. Validation of self-reported health literacy questions among diverse English and Spanish-speaking populations. J Gen Intern Med 2011 Mar;26(3):265-271 [FREE Full text] [doi: 10.1007/s11606-010-1552-1] [Medline: 21057882]
36. Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. J Gen Intern Med 2008 May;23(5):561-566 [FREE Full text] [doi: 10.1007/s11606-008-0520-5] [Medline: 18335281]
37. McInnes DK, Li AE, Hogan TP. Opportunities for engaging low-income, vulnerable populations in health care: a systematic review of homeless persons' access to and use of information technologies. Am J Public Health 2013 Dec;103(Suppl 2):e11-e24. [doi: 10.2105/AJPH.2013.301623] [Medline: 24148036]
38. Pew Research Center.: Pew Research Center; 2012. March 15-April 3, 2012 – Cell Phones URL: http://www.pewinternet.org/dataset/april-2012-cell-phones/ [accessed 2018-01-01] [WebCite Cache ID 6x71i9JIU]
39. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]
40. Centers for Disease Control and Prevention. Hyattsville, MD: Centers for Disease Control and Prevention; 2013. National Health and Nutrition Examination Survey Questionnaire (NHANES) URL: https://www.cdc.gov/nchs/data/nhanes/nhanes_13_14/2013-14_overview_brochure.pdf [accessed 2018-01-01] [WebCite Cache ID 6x71wyMfQ]
41. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 1983 Dec;31(12):721-727. [Medline: 6418786]
42. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. J Clin Psychiatry 1987 Aug;48(8):314-318. [Medline: 3611032]
43. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. Percept Mot Skills 1958;8(3):271-276. [doi: 10.2466/PMS.8.7.271-276]
44. Bravo G, Hébert R. Age- and education-specific reference values for the Mini-Mental and modified Mini-Mental State Examinations derived from a non-demented elderly population. Int J Geriatr Psychiatry 1997 Oct;12(10):1008-1018. [Medline: 9395933]
45. Heaton R. Revised Comprehensive Norms for an Expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults (HRB). Professional Manual. Lutz, Fla: Psychological Assessment Resources. Print 2004 [FREE Full text]
46. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
47. Cheng ST, Chan AC. The Center for Epidemiologic Studies Depression Scale in older Chinese: thresholds for long and short forms. Int J Geriatr Psychiatry 2005 May;20(5):465-470. [doi: 10.1002/gps.1314] [Medline: 15852439]
48. Haringsma R, Engels GI, Beekman ATF, Spinhoven P. The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. Int J Geriatr Psychiatry 2004 Jun;19(6):558-563. [doi: 10.1002/gps.1130] [Medline: 15211536]
49. Babor T, Higgins-Biddle J, Saunders J, Monteiro M. World Health Organization. Geneva; 2001. AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for use in primary care URL: http://apps.who.int/iris/handle/10665/67205 [accessed 2018-01-01] [WebCite Cache ID 6x74bV5PV]
50. Humeniuk R, Henry-Edwards S, Ali R, Poznyak V, Monteiro M. World Health Organization. Geneva; 2010. ASSIST"The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Manual for use in primary care URL: http://apps.who.int/iris/bitstream/10665/44320/1/9789241599382_eng.pdf [WebCite Cache ID 6x74nQ1R9]
51. Gielen AC, McDonnell KA, Wu AW, O'Campo P, Faden R. Quality of life among women living with HIV: the importance violence, social support, and self care behaviors. Soc Sci Med 2001 Jan;52(2):315-322. [doi: 10.1016/S0277-9536(00)00135-0] [Medline: 11144787]
52. Gielen AC, O'Campo PJ, Faden RR, Kass NE, Xue X. Interpersonal conflict and physical violence during the childbearing year. Soc Sci Med 1994 Sep;39(6):781-787. [Medline: 7802853]
53. Pew Research Center. Washington DC: Pew Research Center; 2017. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-01-01] [WebCite Cache ID 6x755TLAd]
54. Le Dantec CA, Edwards W. Designs on dignity: perceptions of technology among the homeless. 2008 Apr 10 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery; April 10, 2008; New York p. A-636 URL: https://www.cc.gatech.edu/~keith/pubs/chi2008-homeless.pdf
55. Eyrich-Garg KM. Sheltered in cyberspace? Computer use among the unsheltered 'street' homeless. Comput Human Behav 2011 Jan;27(1):296-303. [doi: 10.1016/j.chb.2010.08.007]
56. Pew Research Center. Washington DC: Pew Research Center; 2017. Tech adoption climbs among older adults URL: http://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2018-01-01] [WebCite Cache ID 6x765LbMv]
57. Pew Research Center. Washington DC: Pew Research Center; 2017. Internet/Broadband Fact Sheet URL: http://www.pewinternet.org/fact-sheet/internet-broadband/# [accessed 2018-01-01] [WebCite Cache ID 6x76FPkGQ]

58. Uchino BN, Cacioppo JT, Kiecolt-Glaser JK. The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. Psychol Bull 1996 May;119(3):488-531. [doi: 10.1037/0033-2909.119.3.488] [Medline: 8668748]
59. McInnes DK, Petrakis BA, Gifford AL, Rao SR, Houston TK, Asch SM, et al. Retaining homeless veterans in outpatient care: a pilot study of mobile phone text message appointment reminders. Am J Public Health 2014 Sep;104(Suppl 4):S588-S594. [doi: 10.2105/AJPH.2014.302061] [Medline: 25100425]
60. Lifeline Facts. 2017. Obama Phone URL: https://www.obamaphone.com/lifeline-facts [WebCite Cache ID 72VbVZXg5]
61. Jaeger P, Bertot J, Thompson K, Katz S, DeCoster E. The intersection of public policy and public access: digital divides, digital literacy, digital inclusion, and public libraries. Public Library Q 2012;31(1):1-20. [doi: 10.1080/01616846.2012.654728]
62. Gustafson DH, Maus A, Judkins J, Dinauer S, Isham A, Johnson R, et al. Using the NIATx Model to implement user-centered design of technology for older adults. JMIR Hum Factors 2016 Jan 14;3(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.4853] [Medline: 27025985]
63. Pak R, Price MM, Thatcher J. Age-sensitive design of online health information: comparative usability study. J Med Internet Res 2009 Nov 16;11(4):e45 [FREE Full text] [doi: 10.2196/jmir.1220] [Medline: 19917549]
64. Eyrich-Garg KM. Mobile phone technology: a new paradigm for the prevention, treatment, and research of the non-sheltered "street" homeless? J Urban Health 2010 May;87(3):365-380 [FREE Full text] [doi: 10.1007/s11524-010-9456-2] [Medline: 20397058]

Abbreviations

ADL: activity of daily living
AOR: adjusted odds ratio
ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test
CES-D: Center for Epidemiologic Studies Depression Scale
GED: General Educational Development
HEARTH: Homeless Emergency Assistance and Rapid Transition to Housing
HOPE HOME: Health Outcomes of People Experiencing Homelessness in Older Middle Age
Trails B: Trail Making Test
3MS: Modified Mini-Mental State

Edited by G Eysenbach; submitted 12.02.18; peer-reviewed by K McInnes, B Linas; comments to author 30.07.18; revised version received 20.09.18; accepted 25.09.18; published 10.12.18.

Please cite as:
Raven MC, Kaplan LM, Rosenberg M, Tieu L, Guzman D, Kushel M
Mobile Phone, Computer, and Internet Use Among Older Homeless Adults: Results from the HOPE HOME Cohort Study
JMIR Mhealth Uhealth 2018;6(12):e10049
URL: http://mhealth.jmir.org/2018/12/e10049/
doi:10.2196/10049
PMID:30530464

Original Paper

Smartphone-Based Physical Activity Telecoaching in Chronic Obstructive Pulmonary Disease: Mixed-Methods Study on Patient Experiences and Lessons for Implementation

Matthias Loeckx[1,2,3], MSc; Roberto A Rabinovich[4], MD, PhD; Heleen Demeyer[1,2,5], PhD; Zafeiris Louvaris[1,6], PhD; Rebecca Tanner[7], MSc; Noah Rubio[4], BSc; Anja Frei[8], PhD; Corina De Jong[9], PhD; Elena Gimeno-Santos[5,10,11], PhD; Fernanda M Rodrigues[1,2], MSc; Sara C Buttery[7], BSc; Nicholas S Hopkinson[7], MD, PhD; Gilbert Büsching[8]; Alexandra Strassmann[8], MSc; Ignasi Serra[5,10,11], MSc; Ioannis Vogiatzis[6,12], PhD; Judith Garcia-Aymerich[5,10,11], MD, PhD; Michael I Polkey[7], MD, PhD; Thierry Troosters[1,2], PhD

[1]Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium

[2]Respiratory Division, University Hospitals Leuven, Leuven, Belgium

[3]Department of Physiotherapy, LUNEX International University of Health, Exercise and Sports, Differdange, Luxembourg

[4]ELEGI Colt Laboratory, Centre for Inflammation Research, The Queen's Medical Research Institute, University of Edinburgh, Edinburgh, United Kingdom

[5]ISGlobal, Barcelona, Spain

[6]Faculty of Physical Education and Sports Sciences, National and Kapodistrian University of Athens, Athens, Greece

[7]National Institute for Health Research Respiratory Biomedical Research Unit, Royal Brompton and Harefield National Health Services Foundation Trust and Imperial College, London, United Kingdom

[8]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland

[9]Groningen Research Institute for Asthma and Chronic Obstructive Pulmonary Disease-Primary Care, Department of General Practice and Elderly Care, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

[10]CIBER Epidemiología y Salud Pública, Barcelona, Spain

[11]Universitat Pompeu Fabra, Barcelona, Spain

[12]Department of Sport, Exercise and Rehabilitation, Faculty of Health and Life Sciences, Northumbria University, Newcastle-upon-Tyne, United Kingdom

Corresponding Author:
Thierry Troosters, PhD
Department of Rehabilitation Sciences
KU Leuven
Labo Pneumologie
Herestraat 49 bus 706, Onderwijs & Navorsing I
Leuven, 3000
Belgium
Phone: 32 16330798
Email: thierry.troosters@kuleuven.be

Abstract

Background: Telecoaching approaches can enhance physical activity (PA) in patients with chronic obstructive pulmonary disease (COPD). However, their effectiveness is likely to be influenced by intervention-specific characteristics.

Objective: This study aimed to assess the acceptability, actual usage, and feasibility of a complex PA telecoaching intervention from both patient and coach perspectives and link these to the effectiveness of the intervention.

Methods: We conducted a mixed-methods study based on the completers of the intervention group (N=159) included in an (effective) 12-week PA telecoaching intervention. This semiautomated telecoaching intervention consisted of a step counter and a smartphone app. Data from a project-tailored questionnaire (quantitative data) were combined with data from patient interviews and a coach focus group (qualitative data) to investigate patient and coach acceptability, actual usage, and feasibility of the intervention. The degree of actual usage of the smartphone and step counter was also derived from app data. Both actual usage and perception of feasibility were linked to objectively measured change in PA.

Results: The intervention was well accepted and perceived as feasible by all coaches present in the focus group as well by patients, with 89.3% (142/159) of patients indicating that they enjoyed taking part. Only a minority of patients (8.2%; 13/159) reported that they found it difficult to use the smartphone. Actual usage of the step counter was excellent, with patients wearing it for a median (25th-75th percentiles) of 6.3 (5.8-6.8) days per week, which did not change over time (P=.98). The smartphone interface was used less frequently and actual usage of all daily tasks decreased significantly over time (P<.001). Patients needing more contact time had a smaller increase in PA, with mean (SD) of +193 (SD 2375) steps per day, +907 (SD 2306) steps per day, and +1489 (SD 2310) steps per day in high, medium, and low contact time groups, respectively; P for-trend=.01. The overall actual usage of the different components of the intervention was not associated with change in step count in the total group (P=.63).

Conclusions: The 12-week semiautomated PA telecoaching intervention was well accepted and feasible for patients with COPD and their coaches. The actual usage of the step counter was excellent, whereas actual usage of the smartphone tasks was lower and decreased over time. Patients who required more contact experienced less PA benefits.

Trial Registration: ClinicalTrials.gov NCT02158065; http://clinicaltrials.gov/ct2/show/NCT02158065 (Archived by WebCite at http://www.webcitation.org/73bsaudy9)

(JMIR Mhealth Uhealth 2018;6(12):e200) doi:10.2196/mhealth.9774

KEYWORDS

physical activity; COPD; telemedicine; smartphone; patient adherence; patient satisfaction; outcome and process assessment (health care)

Introduction

Background

Reduction in physical activity (PA) is a major feature of chronic obstructive pulmonary disease (COPD), occurring both as a consequence of disease and driving worse outcomes in the condition [1]. PA coaching has been recommended as a nonpharmacological treatment strategy for patients with COPD across all stages of the disease [2]. Telecoaching, where support is provided to achieve effective behavior change by use of electronic communication strategies [3], has received increasing attention in the recent years. It offers the possibility of coaching patients from a distance in an automated or semiautomated way, thereby reducing the burden of face-to-face interactions for patients and health care providers. The latter type of intervention is an example of a complex intervention, which consists of several interacting components [4]. This interaction between multiple components complicates the implementation of such interventions [4]. Therefore, process evaluations have been proposed by the UK Medical Research Council [5], which offer the possibility to investigate how the intervention was delivered (ie, why the intervention worked or did not work) in addition to whether it was effective or not. This is of crucial importance to health technology assessment bodies as it provides information on which components of an intervention were effective or noneffective and how the intervention can be improved and replicated in different settings and patient groups [4,5]. Process evaluation can also be of great value in evaluating PA telecoaching interventions, which have been shown to be effective in enhancing PA in some studies [6-8] but not in others [9]. In a recent multicenter PA telecoaching trial (MrPAPP), which had a positive outcome [6], a large variability in the effect of the intervention was noticed. Patients with better functional exercise capacity (ie, 6-minute walking distance [6MWD] ≥450 meters), fewer symptoms (ie, modified Medical Research Council [mMRC] dyspnea scale <2), those in Global Initiative for Chronic Obstructive Lung Disease (GOLD) quadrants A-B improved their PA to a greater extent [6]. In addition to these patient characteristics, intervention-specific characteristics and the way patients cope with the intervention might also have contributed to the success of the intervention.

Objectives

In this paper, 3 concepts, which are often assessed as part of a process evaluation, have been investigated: (1) acceptability, (2) actual usage, and (3) feasibility of the intervention from both a patient and a coach perspective. In addition, we aimed to investigate their association (ie, actual usage and feasibility) with the effectiveness of the intervention.

First, acceptability is a key concept in the development, evaluation, and the implementation of complex interventions and can have significant impact on the intervention's effectiveness [10]. It has been defined as "a multi-faceted construct that reflects the extent to which people delivering or receiving a health care intervention consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention" [10]. A potentially effective intervention might not reach its potential due to poor acceptability to patients or health care providers [10].

Second, the actual usage of the intervention by patients and health care providers forms an important part of the delivery of PA telecoaching interventions. Actual usage was assessed as the degree to which patients used the components of the interventions as it was designed [11]. It is often confused with the term *adherence* [12]. The latter term requires a rationale for the minimum intended use of the components of the intervention. As there is no established minimum usage of such PA telecoaching interventions, we used the term *actual usage*, with the assumption that the more usage, the better [12]. Although the actual usage of step counters is known to be relatively good in short-term coaching trials involving patients with COPD [7,13,14], actual usage of smartphone apps in coaching trials has been less intensively studied.

Third, the implementation of this intervention also depends on whether it was considered to be feasible by patients as well by the coaches. Feasibility is defined as "the extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting" [15,16]. The coach feasibility of the PA telecoaching program in this paper has already been partly assessed in the main paper of the MrPAPP trial, which reported that coaches contacted patients for a total duration of 50 min throughout the trial [6]. However, qualitative data on the perceived feasibility of both patient and coach are lacking.

Finally, the direct association between both coach feasibility (as assessed by contact time) and actual usage by patients with the effectiveness of the intervention was investigated. The latter insights could lead to improved design and implementation of PA telecoaching interventions in the future as well as optimized selection of patients.

Methods

Ethics Approval

This study was approved by the local ethics committee at each center (Commissie medische ethiek van de universitaire ziekenhuizen KU Leuven [Leuven, S-55919]; Medische ethische toetsingscommissie universitair medisch centrum Groningen [Groningen, Metc 2013.362]; RES Committee London—South East [London and Edinburgh, 13/LO/1660]; Scientific Council of the 'Sotiria' General Hospital for Chest Diseases (Athens, 27852/7-10-13); Kantonale Ethikkommission Zürich, and Ethikkommission Nordwest- und Zentralschweiz [Zurich, KEK-ZH-Nr. 2013-0469 and EKNZ2014-192, respectively]).

Study Population and Design

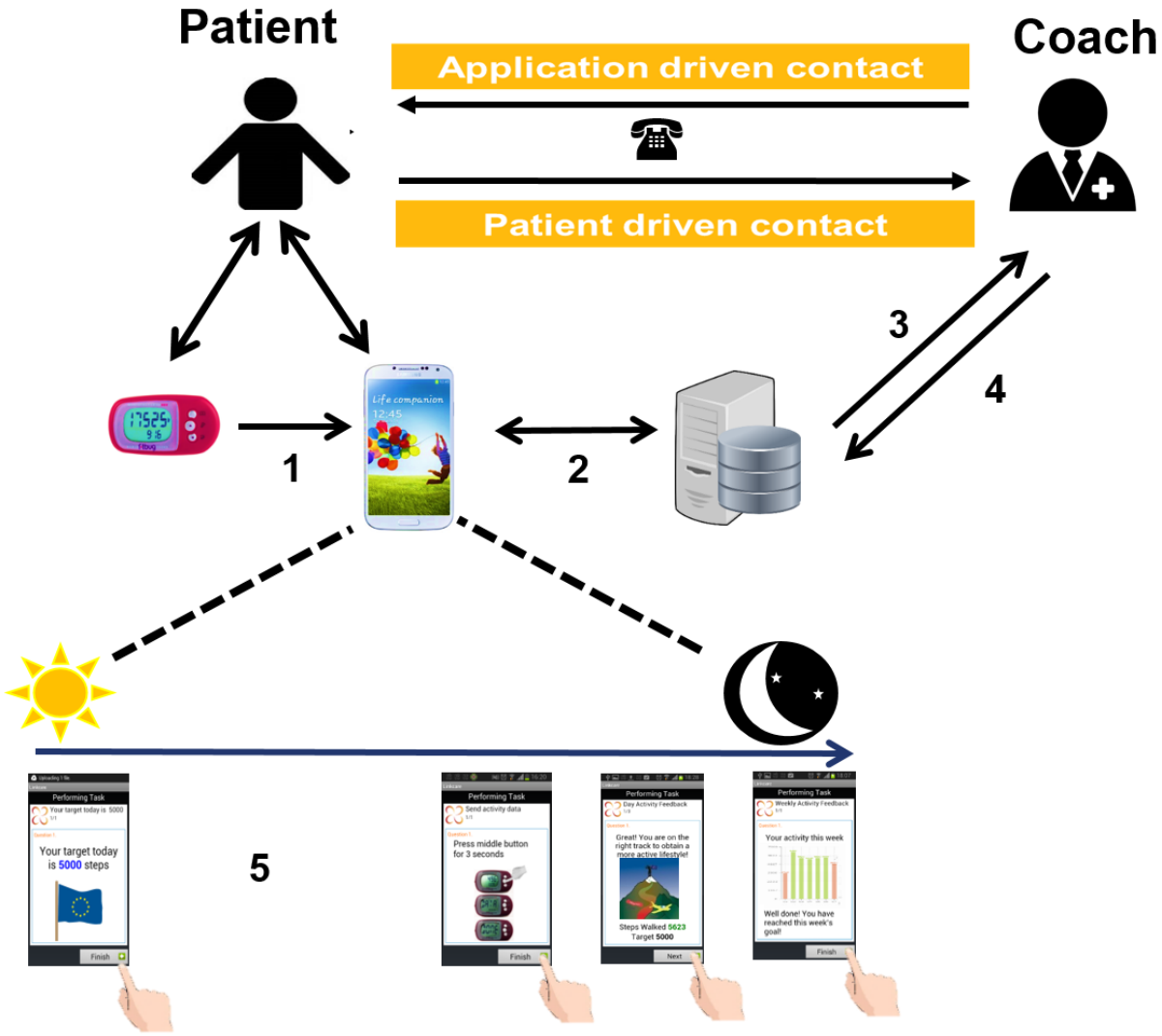
A convergent mixed-methods research design using quantitative and qualitative data was applied to evaluate the acceptability, actual usage, and feasibility of a PA telecoaching intervention. Both qualitative and quantitative data on the intervention were separately collected and analyzed. Later, these findings were compared for data triangulation, which allowed a more comprehensive understanding of the intervention [17-19].

This trial forms part of a 12-week, multicenter randomized controlled trial (1:1 randomization) conducted by the PROactive consortium [6]. The trial consisted of 3 visits—a screening visit (V1), a randomization visit (V2) 1 to 2 weeks later, and a final visit (V3) 12 weeks post randomization. In total, 171 patients were allocated to the intervention group in 6 centers across Europe (Leuven, Belgium; Athens, Greece; London and Edinburgh, United Kingdom; Zurich, Switzerland; and Groningen, The Netherlands) between June and December 2014, from which 159 patients completed the trial and were considered for the present analyses. More information on the study population and design has already been published elsewhere [6]. All patients provided informed consent before any data collection.

Physical Activity Telecoaching Intervention

Patients in the intervention group [6] received a multicomponent PA telecoaching intervention, consisting of a step counter and a smartphone app (Samsung Galaxy S4 mini; android version 4.4.2), in addition to usual care. Furthermore, patients in the intervention group received an exercise instruction booklet for home use and a one-to-one interview with a coach discussing motivation, barriers, favorite activities, and strategies to become more active. The exercise instruction booklet contained 3 different sessions of upper limb and lower limb stretching as well as balance and strengthening exercises with a standardized amount of sets and repetitions (see Multimedia Appendix 1). Patients were asked to wear the step counter (Fitbug air) during waking hours and to interact with the project-tailored smartphone app on a daily basis. They were instructed to access and review automated tasks that appeared on the smartphone's display and to press the *closebox* on the screen thereafter (ie, completion of a task). An audio reminder was provided for patients to send their step data at 8 pm to their smartphone (through Bluetooth) by pressing a single button of the step counter. The app provided patients with daily activity goals in the morning, which were set for 1 week. The patients' goal was adjusted according to their PA performance in the previous week and to their willingness to increase their goal. Goals were calculated based on the mean and median of the 4 most active days of the previous week. If the mean value was higher than the weekly goal (ie, patients reaching the goal), the patients had the opportunity to (1) not change or (2) increase their median goal by 500 steps through a *yes* or *no* option displayed on the app. If the mean of the 4 most active days of the previous week was lower (ie, patients not reaching their goal) and the median was more than 500 steps below the goal, the goal was reduced to the median of the 4 most active days+500 steps. In other cases, the goal remained the same. Coaches were asked to contact the patients (ie, tasks of the coaches) in case patients (1) did not send their step count data for 3 consecutive days, (2) did not reach their target for 2 consecutive weeks, (3) reached the target but they were not willing to increase for 2 consecutive weeks, and (4) were not adherent with wearing the step counter for 2 consecutive weeks. More details on when coaches were instructed to contact the patients (ie, flagging system) are published elsewhere [6]. Daily and weekly encouraging feedback messages were displayed on the smartphone using both text and pictograms (see Multimedia Appendix 2; slide 7). Throughout the whole intervention period, coaches could access patient data through their app-linked Web accounts to monitor patients' performed PA and their actual usage of the intervention (PROactive Linkcare app, Barcelona, Spain; see Multimedia Appendix 2). The use of the intervention was completely free of charge for all patients. No major bug fixes or changes to the intervention were made throughout the trial. A detailed overview of how the intervention works can be found in Figure 1.

Figure 1. Overview of the intervention; 1=sending of "steps data" to smartphone (through Bluetooth); 2=data sent to central database; 3=coach is able to access database; 4=coach is able to manually adjust goals, 5=accessing & closing the different tasks on the smartphone app (automated messages); i.e., (from left to right); morning goal, send activity in the evening, daily feedback (from Monday to Saturday) and weekly feedback (only on Sunday) tasks.



Outcomes

Acceptability

Acceptability was assessed through quantitative data (a project-tailored questionnaire [20 items, Multimedia Appendix 3]) and qualitative data collection (patient interview [4 open questions, Multimedia Appendix 4] and a coach focus group [Multimedia Appendix 5]).

During the final visit of the study (V3), patients were asked to fill in a 20-min self-administered, project-tailored, multiple-choice questionnaire on their experiences with the intervention and the usefulness of its components on a 10-point Likert scale (Multimedia Appendix 3). Each center collected and anonymized answers from all their patients into an Excel file, which was sent to 1 investigator (HD). HD pooled all data together into 1 Excel file, which was then used for analysis.

Patient interviews were conducted by local PA coaches in each center at V3. Each coach was informed and trained on how to conduct the interview during an investigator's meeting before the start of the trial. Interviewers from each center were asked to transcribe the answers of the patients to the discussion guide questions and forward them (anonymized) to one researcher (ML) who collected all quotes into 1 Excel file for analysis. In this pooled Excel file, each line represented the verbatim answer of each participant on a question with a number code and a letter representing, respectively, the patient's ID and the question of the discussion guide.

After completion of the trial, an audiotaped focus group was organized to capture the intervention experience from the perspective of the coaches. Local PA coaches with a diverse background (ie, medical doctor [RAR], physiotherapist [ML, HD], exercise physiologist [ZL], biomedical scientist [MS], and psychologist [AF]; n=6), and 2 experienced physiotherapists who were involved in the development of the intervention (n=2; EGS and Ane Arbillaga-Etxarri (AAE) from the center in Barcelona [IS GLOBAL]) discussed the feasibility, appreciation, possible future adaptations, time investment, and actual usage

of the different components of the intervention (Multimedia Appendix 5). A total of 2 PA coaches (ML and HD) facilitated the focus group.

Actual Usage

Actual usage of the intervention by patients was assessed objectively through the smartphone app log. A database was derived directly from the smartphone app. This included information about completion of the app tasks and step counter data on a day-by-day basis. Actual usage of the step counter was defined based on the presence of step count data (ie, ≥70 steps for that day). Self-reported actual usage of performing home exercise and the times patients looked at their step counter were assessed subjectively in the project-tailored questionnaire.

Actual usage by the coaches was assessed based on the closure of tasks in the app-linked Web accounts and discussed during the coach focus group.

Feasibility

Coach feasibility was already partly assessed in the main paper of the MrPAPP trial in terms of number of contacts and total amount of contact time between coaches and patients (quantitative data) [6]. As a secondary analysis, the evolution in efficiency of coaches, as measured by contact time throughout the study recruitment period, was assessed. In addition, coach perception of the feasibility of the intervention was also covered in the coach focus group (qualitative data). Intervention feasibility from the patient perspective was evaluated through the project-tailored questionnaire (quantitative data) and patient interviews (qualitative data).

Association of Actual Usage and Feasibility With the Effectiveness of the Intervention

Both actual usage by patients and coach feasibility (ie, contact time) with the intervention were separately linked to the effectiveness of the intervention. This effectiveness was assessed as the change in numbers of steps per day after 12 weeks, measured by the Actigraph GT3x (ACT, Actigraph LLC Pensacola, FL). The latter is a triaxial accelerometer validated for use in patients with COPD [20,21]. Further details on the PA assessment methodology and its validity criteria can be found elsewhere [6].

Statistical Analysis

All statistical analyses were performed with Statistical Analysis Software version 9.4 (SAS Institute, Cary, NC). Continuous variables were expressed as means with SD (normal distribution) or as medians (25th-75th percentiles [P25-P75]; skewed distribution), unless stated otherwise. Categorical variables were expressed as proportions and percentages. The level of significance was set at .05 for all statistical tests. The analyses were based on patients in the intervention group who completed the 12-week intervention (N=159).

Data from the project-tailored questionnaire were scored as categorical variables and reported as frequencies and percentages (ie, number of patients indicating each answer), except for the usefulness ratings of the components, which were expressed as median (P25-P75).

For analysis of the interview data, two researchers (HD and FMR) independently performed thematic analysis on the Excel file containing the verbatim transcriptions of the interview data [22] according to the 6-step framework as proposed by Braun and Clarkes [23]:

1. HD and FMR read the data multiple times and descriptively noted down their initial ideas of what is in the data and what is interesting about them.
2. HD and FMR independently generated an initial list of codes from the data and put the data systematically under certain headings.
3. Afterwards, they searched for reoccurring themes, which began to emerge from these codes to focus their analysis on a broader level.
4. HD and FMR refined and defined their themes taken into account the overall message of the analysis. Themes and subthemes were organized and ranked into categories.
5. HD and FMR came together for group discussion to find an agreement on defining the themes and subthemes, which led to the development of a (final) codebook.
6. Afterwards, one researcher (ML) applied the final codebook to all verbatim transcripts. After iterative group discussions, data were synthesized and representative example quotes were extracted to illustrate findings and were labeled by a unique participant's code together with the category of contact time and actual usage score of that participant.

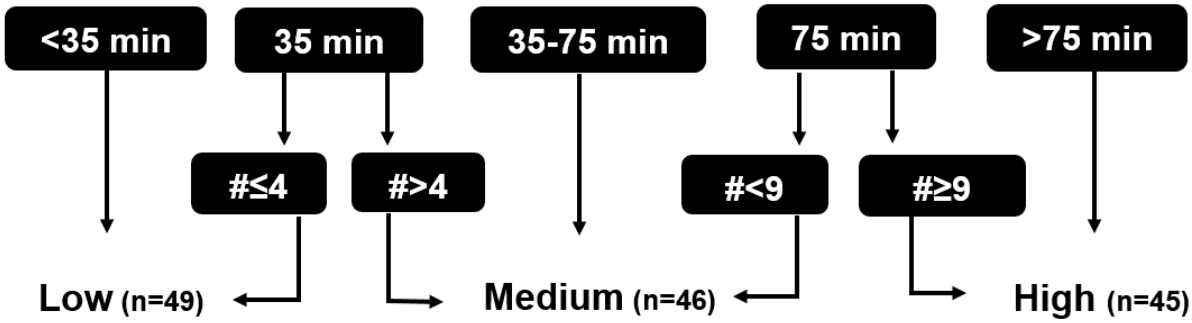
The thematic analysis was conducted inductively (ie, themes emerged from the data, hence without predetermined coding frame) in Excel, without the use of specialized analytic software. Further details on the methodological aspects of the latter analyses have been added to the COnsolidated criteria for REporting Qualitative research (COREQ) checklist (see Multimedia Appendix 6).

During the focus group, 1 PA coach (ML) wrote a consensus summary. A total of 2 PA coaches (HD and MS) independently reviewed the consensus summary based on the audio recording. Additional information that was considered as relevant was independently added by both coaches (HD and MS). Only minor interpretation disagreements occurred between the 2 PA coaches, which were discussed together with a third PA coach (ML). Later, a summary of the focus group was sent for revision to all PA coaches, including those who could not be present at the focus group. A consensus quote on the future implementation of this PA telecoaching intervention was formulated.

Actual usage was compared according to age (<65 vs ≥65 years, Mann-Whitney *U* test), gender (male vs female, Mann-Whitney *U* test), and over time in the trial (week 2-3 vs week 11-12, Wilcoxon signed-rank sum test). Actual usage of the step counter was expressed as the percentage of patients who wore the step counter for at least 90% of the days in the study. Actual usage of the different smartphone tasks was expressed as median (P25-P75).

In the larger centers (inclusion of at least 20 patients), the contact time with the first 10 patients was compared with the others (Mann-Whitney *U* test) to assess possible learning effect of the coaches.

Figure 2. Division (into 3 groups) of patients based on total duration and number of contacts between patients and coach. Min=minutes; #=number of contacts; n=number of patients in each group.



We attempted to create 3 equally balanced groups (low, medium, and high) of total contact time (Figure 2) and of an overall score of actual usage. This overall actual usage score was calculated by summing up each actual usage component (actual usage of all tasks and wearing the step counter) as a percentage of their recommended frequency. The 3 groups were compared (via analysis of variance test or Kruskal-Wallis test) to characterize those who required a lot of contact time and those who did not and those who had high actual usage of the intervention and those who did not. As a sensitivity analysis for the latter tertiles approach, we also analyzed contact time and actual usage score as continuous variables. The methodology used for the latter sensitivity analysis can be found in Multimedia Appendix 7.

To analyze the association between (1) the actual usage by patients of different components of the intervention and coach feasibility (ie, contact time) and (2) the effectiveness of the intervention, 2 separate generalized linear model analyses were used in completers with valid PA data (88.1% [140/159] of the completers sample). Change in PA was used as the outcome and contact time and actual usage as the class variables, respectively. Due to their possible influence on the intervention effect, baseline exercise capacity (6MWD), symptom score (mMRC-scale), forced expiratory volume in 1 second (FEV_1) % predicted, and the number of acute exacerbations in the previous 12 months were considered as possible (continuous) covariates of the association [6]. Details on sensitivity analyses for the latter tertiles approach (with contact time and actual usage scores as continuous variables) can be found in Multimedia Appendix 7. Finally, we hypothesized high contact time in the first 4 weeks to be an early sign of absence of response to the intervention. To that end, we calculated the likelihood of achieving the minimal important difference (MID) improvement of 1000 steps per day [24] in patients with a low (≤30 min) and high (>30 min) contact time in the first 4 weeks of the trial (as a possible early predictor for treatment failure).

Results

Study Population

Baseline characteristics of the 159 completers are outlined in Table 1. Information on the full study population (including further details about dropouts and the occurrence of adverse events) has been detailed elsewhere [6].

Outcomes

Acceptability

Overall, the PA telecoaching intervention was well received by the patients as 89.3% (142/159) indicated that they “enjoyed taking part in the intervention.” Furthermore, the majority of the patients (59.1%, 94/159) claimed that the intervention coached them “a lot” toward enhancing their PA. Approximately half of the patients (47.2%, 75/159) experienced the proposed weekly increases in step counts as “reasonable,” whereas 37.7% (60/159) and 10.1% (16/159) of the patients experienced these increases as “a little bit too high” and “much too high,” respectively.

Patients rated the usefulness of the step counter (median [P25-P75]; 10 [8-10]) and the telephone contacts with the coach in case of problems (9 [7-10]) as the most crucial parts of the intervention (see Figure 3). The display of a daily (educational) activity tip in the evening (6.5 [5-8]) and the booklet for home exercises (6 [4-8]) were rated as less useful.

When patients were asked to name the most important part of the intervention, 76.1% (121/159) of patients did choose the step counter as the most important part with 93.1% (148/159) of all patients willing to continue using the step counter in the future. In total, 45.9% (73/159) of all patients were willing to continue using the full intervention, with only 8.2% (13/159) of all patients reported to experience working with the smartphone as difficult.

Table 1. Baseline characteristics of the completers of the trial.

Variables	Intervention completers (n=159)
Age in years, mean (SD)	66 (8)
Gender (male), n (%)	89 (64)
BMI[a] (kg/m^2), mean (SD)	26.9 (5.3)
FEV_1[b] predicted (%), mean (SD)	53.9 (19.9)
6MWD[c] (m), mean (SD)	442 (107)
6MWD predicted (%), mean (SD)	70.3 (16.5)
CAT[d] score, mean (SD)	13 (8)
QF[e] (kg), mean (SD)	31.5 (10.9)
PA[f] (steps per day), median (P25-P75)[g]	4272 (2783-5768)

[a]BMI: body mass index.

[b]FEV_1: forced expiratory volume in 1 second.

[c]6MWD: 6-minute walking distance; 6MWD was missing in 2 patients.

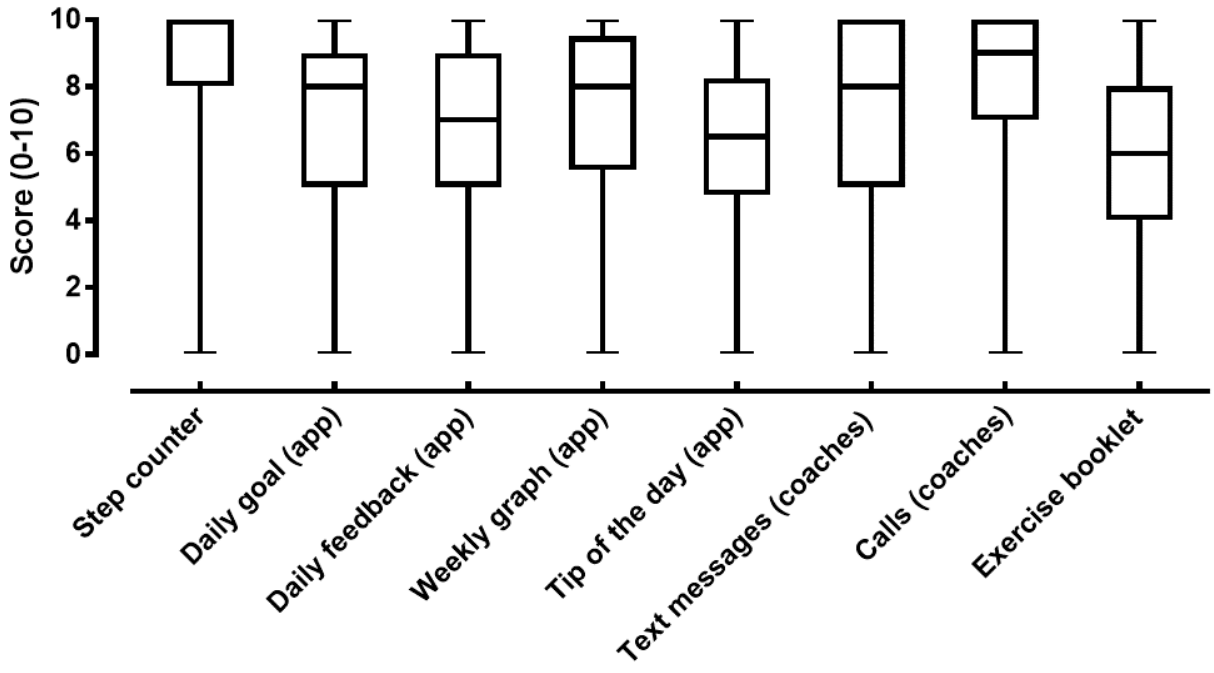
[d]CAT: chronic obstructive pulmonary disease (COPD) assessment test.

[e]QF: quadriceps force; QF was not measured in 2 centers and QF was missing in 27 patients.

[f]PA: physical activity; valid PA measurements were present in 140 patients.

[g]25th and 75th percentiles (P25-P75).

Figure 3. Boxplots depicting the usefulness score (0-10 Likert scale) of the different parts of the intervention from the patients' perspective. "app" between brackets represents messages displayed on the smartphone app.



In total, 91.2% of patients (145/159 of the completers sample) took part in the semistructured interviews at V3. Themes and subthemes that were derived from the verbatim responses of patients to the interview are presented in Textbox 1. Moreover, 2 major topics can be distinguished from the interview data: technical aspects and aspects related to the content of the intervention (see Textbox 1). Illustrative quotes, which support findings from the thematic analysis, are provided in Multimedia Appendix 8. Further information on the interview process, participants, and the interviewers can be found in the COREQ-checklist (see Multimedia Appendix 6).

Textbox 1. Findings of the thematic analysis of the interview data are categorized under (1) technical aspects and (2) aspects related to the content of the intervention.

Technical aspects:

Themes of (1) positive experiences and (2) issues or problems emerged from the data.

1. Positive experiences
   - No technical problems: A large portion of patients stated not to have encountered technical issues with any of the components of the intervention.
   - Working with app: The ease of use with the different components of the intervention was highlighted by patients. Furthermore, patients who had less a priori experience with managing a smartphone device expressed that the learning process of working with this device was smooth.
2. Issues or problems
   - Help from others: Few patients needed more than a familiarization period before they were able to feel confident about working with the smartphone and its app. Help from both the study team (through phone calls or face-to-face contacts) and from their relatives was considered essential when experiencing problems.
   - Speed of interaction with the app: Some patients felt the speed of the app was slow and perceived the interaction with it as time consuming. Especially, the transfer of step data onto the phone in the evening was delayed for several minutes.
   - App problems: Some patients reported during the interview that working with the app was often hindered (eg, tasks not opening and not possible to send data). Reasons for these app problems were mostly related to issues with the internet connection or Bluetooth problems.
   - Step counter: A small minority of patients expressed their frustration with the step counter that was not always able to detect all steps they performed. Activities such as slow walking, cycling, and arm movements were not measured accurately.

Aspects related to the content of the intervention:

Themes of (1) positive experiences, (2) issues or problems, and (3) outcome emerged from the data.

1. Positive experiences
   - Step counter: The step counter was judged as the essential part of the intervention by several patients because of its simplicity, feedback, and usefulness.
   - Graphs: Another highly rated aspect of the intervention was the graphical feedback displays that patients received based on the achievement of their goals. According to the patients, it was an interesting and excellent way of motivating them.
   - Nice experience: In general, the intervention was considered as motivating to a large majority of patients across the different centers. Patients claimed it was a fun and interesting experience that helped them toward being more active and feeling better and fitter.
   - Being monitored: One of the most important motivational reasons according to patients to become more active was the feeling of being monitored. Knowing that the coaches were following them up gave them an external motivational cue to be physically active.
   - Family participation: Next to the help from the coaches, patients' relatives often played an important supportive and stimulating role throughout the intervention. Close relatives of patients (mostly spouses) also bought a step counter to join their wife or husband throughout their coaching.
2. Issues or problems
   - Goals: One of the most important issues was the increase in the step count goal, which was often too high for patients. This caused some frustration among patients as it was perceived as demotivating to have too high goals and not being able to reach them.
   - Variations: As the intervention was used for a period of 12 weeks, the component of variation in the content of the app was deemed as important according to the patients. Some patients reported that because of the lack of variation, their actual usage of the intervention (in particular with the opening of the messages on the smartphone) lowered. The morning messages with the goal patients needed to achieve were repeated every day of that week and required more variation according to the patients.
   - Barriers: One of the major drawbacks of the intervention according to patients was that it did not take into account several barriers with which they were confronted. When a patient experienced an acute exacerbation, his or her goal was not adjusted immediately. Weather factors were not taken into account within the app. Furthermore, patients regretted that there was no option for them to make the intervention aware that they had other priorities (eg, holidays or days when they needed to watch their grandchildren).
   - Motivational issues: A few patients did not find the app interesting and did not like working with it.
3. Outcomes
   - New routine: Patients stated that the intervention and the goals resulted in the adoption of new lifestyle routines to be more physically active. They hoped to continue with these more active lifestyles after the intervention finished.

All coaches present at the focus group considered the intervention to be a useful addition to standard care in patients with COPD. The coaches rated the step counter as very useful, mainly attributed to the direct feedback it provided and its ease of use. Technical problems with the smartphone interface intermittently occurred (eg, Bluetooth connection or requests

for automatic updates). In addition, coaches reported that a minority of patients felt the smartphone app lacked variation. Considering future long-term use, coaches proposed a more individualized technical training based on individual patient needs (eg, more extensive in patients with difficulties and those needing more contact time). Finally, the coaches regretted that the home exercises did not result in higher step counts and lacked variation, which might explain the low use of the home exercise booklet by patients.

Actual Usage

Almost 60% (59.7%, 95/159) of patients wore the step counter for more than 90% of the days they were included in the coaching program, representing a median (P25-P75) of 6.3 (5.8-6.8) days per week with no difference over time within the trial (P=.98). Actual usage of the different smartphone app tasks is outlined in Table 2. Actual usage decreased significantly over time for all tasks (P<.001 for all) except for the weekly feedback task (P=.14). More specifically, actual usage of the daily goal, sending activity, and daily feedback tasks decreased from, respectively, 5 (3-7), 5 (2.5-6), and 3 (1-5) days per week at the start of the intervention to 4 (1.5-6.5), 3.5 (0.5-6.0), and 2 (0-4.5) days per week at the end of the trial (P<.001 for all). The actual usage did not differ between younger and older patients or between male and female patients (Multimedia Appendix 9).

In terms of self-reported actual usage, a large majority of the patients (76.7%, 122/159) stated that they looked *several times per day* at their step counter. Only 22.0% (35/159) of patients claimed to perform their home exercise *at least on a daily basis* and one-third stated they had *never* performed these exercises.

Coaches performed 1053 out of the 1161 contacts that appeared on the platform; however, no details on the time of solving the tasks were available.

Feasibility

Feasibility from the perspective of the patients was good as a large proportion of patients reported that the smartphone intervention was not too much of a burden to work with when they were asked how they had experienced the technical aspects of the intervention. Coaches spent significantly more time (P=.002) interacting with the first 10 of their patients compared with the ones who were recruited at a later stage in their center (see Figure 4). These findings were confirmed when the arbitrarily chosen cutoff point of comparing the first 10 patients was changed with the first 8 or 12 patients.

All PA coaches present in the focus group reached consensus that a follow-up of approximately 25 to 30 patients simultaneously for 1 coach would be feasible. It was felt to be beneficial to have 1 coordinating center to discuss day-to-day problems in patient management on a case-by-case approach.

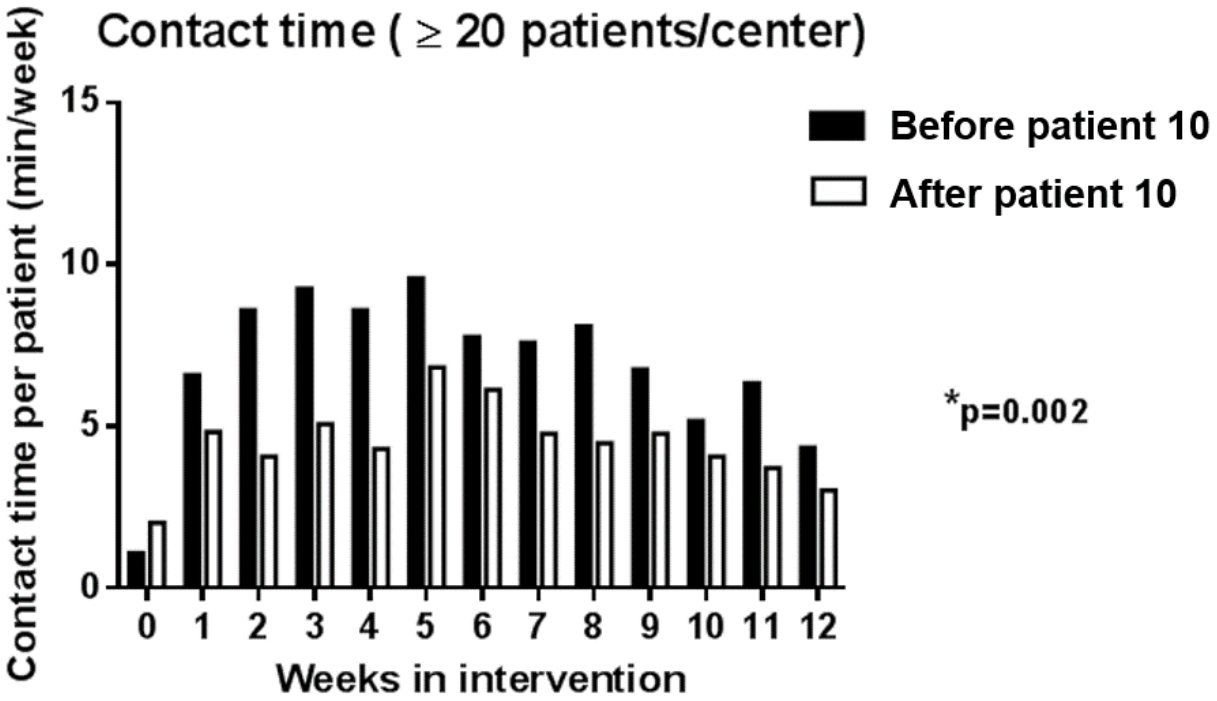
Table 2. Overview of the different components of the intervention. Definition of actual usage of the different components of the intervention of all completers (n=159 patients) and the minimum and maximum values one can achieve in terms of actual usage were reported when applicable. Actual usage and possible minimum-maximum are expressed as median (P25-P75) days per week for the step counter and the daily tasks on the app. Weekly feedback is expressed as median (P25-P75) percent of weeks in the intervention.

Components of the intervention	Actual usage		
	Definition of actual usage	Median (p25-p75)[a]	Possible minimum-maximum
One-to-one interview with coach discussing motivation, barriers, favorite activities, and strategies to become more active	N/A[b]	N/A	N/A
Step counter (Fitbug Air; days per week)	A day with ≥70 steps recorded	6.3 (5.8-6.8)	0-7
A project-tailored smartphone coaching app (Linkcare, Barcelona ES) with different tasks			
Send activity data task (days per week)	Patient closes task	4.1 (2.4-5.6)	0-7
Looking to the daily goal task (days per week)	Patient closes task	4.1 (2.1-5.9)	0-7
Looking at the daily feedback task (days per week)	Patient closes task	2.2 (0.7-4.1)	0-6
Looking at the weekly feedback task (% of weeks in the intervention)	Patient closes task	55 (29-78)	0-100
A booklet containing home exercises	N/A	N/A	N/A
Weekly group text messages with activity proposals sent by the coach	N/A	N/A	N/A
Contact with the coaches, which was triggered in the case of nonactual usage with wearing the step counter, failure to transmit data, or failure to progress	N/A	N/A	N/A

[a]25th and 75th percentiles (P25-P75).

[b]N/A: not applicable.

Figure 4. Contact time throughout the intervention (only including centers with more than 20 patients). The black bars represent the mean contact time (in min per week) per patient from the first 10 patients that were recruited in each center. White bars represent the mean contact time (in min per week) per patient from the patients that were recruited at a later stage. *P* value indicates difference between the total cumulated contact time over the 12 weeks between patients recruited in early stage versus later stage.



Association of Actual Usage and Feasibility With the Effectiveness of the Intervention

Patients in the low (n=49), medium (n=46), and high (n=45) contact time group had a median (P25-P75) total contact time of 25 (10-30), 50 (40-60), and 140 (105-185) min, respectively. Patients who had more contact time with the coaches during the time of the study, had more severe airflow obstruction, tended to have a lower functional exercise capacity (Table 3) and had a significant smaller increase in PA, also after adjusting for covariates (age, baseline FEV_1 [%predicted], baseline 6MWD, baseline mMRC-score, and the number of acute exacerbations in the last 12 months; P-for-trend=.01; Figure 5). The latter findings were confirmed when contact time was treated as a continuous variable (see Multimedia Appendix 7).

When groups were divided in 3 according to their overall actual usage score, neither patient characteristics nor effectiveness were different (see Table 4 and Figure 6). The latter findings were confirmed when actual usage score was treated as a continuous variable (sensitivity analyses in Multimedia Appendix 7).

Logistic univariate regression analysis revealed that patients with a low contact time (≤30 min; n=103) after 4 weeks were 3.58 times more likely of achieving the MID improvement of 1000 steps per day (95% CI 1.88-6.82; *P*<.001) compared with patients with more contact time.

Table 3. Patient baseline characteristics according to the total contact time (only including patients with valid PA measurement; n=140); data are expressed as mean (SD) unless stated otherwise. *P* value indicates differences between the 3 contact time groups.

Variables	Low contact time (n=49)	Medium contact time (n=46)	High contact time (n=45)	*P* value
Age in years, mean (SD)	65 (7)	65 (10)	68 (6)	.16
Gender (male), n (%)	28 (57)	34 (74)	27 (60)	.20
BMI[a] (kg/m^2), mean (SD)	27.8 (5.3)	26.1 (4.4)	27.0 (6.4)	.35
FEV_1[b] predicted percentage, mean (SD)	59.5 (22.6)	54.1 (16.5)	49.1 (20.5)[j]	.04
6MWD[c] (m), mean (SD)	444 (100)	459 (101)	411 (113)	.09
6MWD predicted percentage, mean (SD)	71.5 (14.5)	71.2 (15.0)	67.4 (19.6)	.29
CAT[d] score, median (p25-p75)[e]	10 (6-17)	13 (7-19)	16 (10-21)	.11
QF[f] (kg), mean (SD)	33.1 (13.2)	31.2 (10.0)	29.2 (10.5)	.33
PA[g] (steps per day), median (p25-p75)	4542 (3387-5587)	4377 (3016-6723)	3186 (2375-5339)	.15
Contact time first 4 weeks in minutes, median (p25-p75)	0 (0-5)[h]	10 (5-20)[i]	50 (20-85)[j]	.005

[a]BMI: body mass index.

[b]FEV_1: forced expiratory volume in 1 second.

[c]6MWD: 6-minute walking distance; 6MWD was missing in 2 patients.

[d]CAT: chronic obstructive pulmonary disease (COPD) assessment test.

[e]25th and 75th percentiles (P25-P75).

[f]QF: quadriceps force; QF was not measured in 2 centers and QF was missing in 27 patients.

[g]PA: physical activity.

[h]Indicates statistical significance (*P*<.05) between low versus medium contact time groups.

[i]Indicates statistical significance (*P*<.05) between medium versus high contact time groups.

[j]Indicates statistical significance (*P*<.05) between low versus high contact time groups.

Figure 5. Change in physical activity (PA; mean [SE]) across groups of patients according to total contact time; adjusted for age, baseline functional exercise capacity, baseline forced expiratory volume in 1 second, baseline symptom score and number of acute exacerbations in the previous 12 months. *P* value (P for trend) indicates difference in intervention effect between patients divided based on total contact time, after adjusting for the covariates. Data are based on Actigraph measurements and include 140 patients. Unadjusted scores were mean(SD) +1489 (SD 2310) steps per day, +907 (SD 2306) steps per day and +193 (SD 2375) steps per day in low, medium and high contact time groups, respectively.

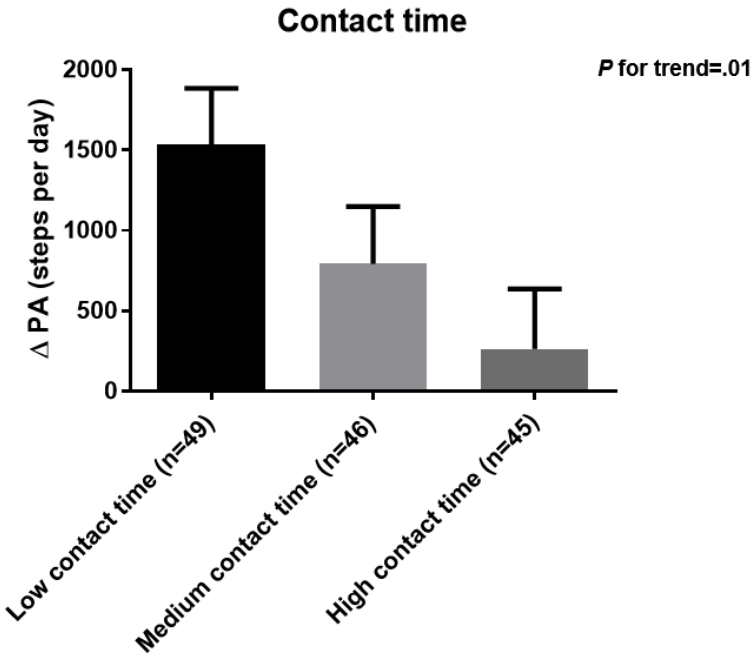


Table 4. Patient characteristics according to the total actual usage score (3 groups only including patients with valid physical activity measurement by actigraph, n=140); data are expressed as mean (SD) unless stated otherwise. *P* value indicates differences between the 3 actual usage groups.

Variables	Low actual usage, <47% of usage (n=47)	Medium actual usage, 47% to 75% of usage (n=46)	High actual usage, >75% of usage (n=47)	*P* value
Age in years, mean (SD)	66 (8)	66 (9)	65 (8)	.76
Gender (male), n (%)	31 (66)	29 (63)	29 (62)	.91
BMI[a] in kg per m^2, mean (SD)	27.5 (5.3)	27.6 (6.5)	26.0 (4.3)	.34
FEV_1[b] predicted percentage, mean (SD)	54.4 (20.3)	55.2 (19.5)	53.5 (21.6)	.92
6MWD[c] (m), mean (SD)	431 (106)	432 (105)	454 (107)	.50
6MWD predicted percentage, mean (SD)	69 (17)	69 (17)	72 (16)	.61
CAT[d] score, median (p25-p75)[e]	14 (7-19)	13 (6-19)	12 (7-21)	.94
QF[f] (kg), mean (SD)	32.0 (10.8)	30.0 (12.9)	31.1 (9.4)	.73
PA[g] (steps per day) median (p25-p75)	4369 (2868-5672)	3850 (2380-6108)	4540 (2940-6731)	.49

[a]BMI: body mass index.

[b]FEV_1: forced expiratory volume in 1 second.

[c]6MWD: 6-minute walking distance; 6MWD was missing in 2 patients.

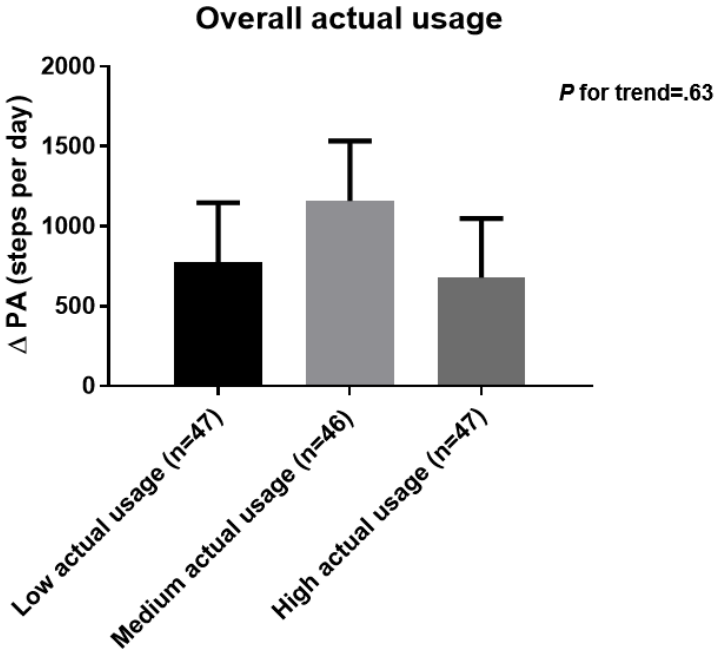
[d]CAT: chronic obstructive pulmonary disease (COPD) assessment test.

[e]25th and 75th percentiles (P25-P75).

[f]QF: quadriceps force; QF was not measured in 2 centers and QF was missing in 27 patients.

[g]PA: physical activity; valid PA measurements was present in 140 patients.

Figure 6. Change in physical activity (PA; mean [SE] across groups of patients according to overall actual usage score; adjusted for age, baseline functional exercise capacity, baseline forced expiratory volume in 1 second, baseline symptom score and number of acute exacerbations in the previous 12 months. *P* value (P for trend) indicates difference in intervention effect between patients divided based on the total actual usage score, after adjusting for the covariates. Data are based on Actigraph measurements and include 140 patients. Unadjusted scores were mean(SD) +777 (SD 2767) steps per day, +1159 (SD 2720) steps per day and +679 (SD 2075) steps per day in low, medium and high actual usage groups, respectively.



Discussion

Principal Findings

On the basis of the secondary analysis of the MrPAPP PA telecoaching trial in patients with COPD, this mixed-methods research design study shows that the intervention was feasible and well accepted by both patients and coaches. Given the design of the intervention (ie, patients were contacted when PA was not increasing), patients with high contact time with coaches had less PA improvements, suggesting that the high contact time resulted from either difficulty or reluctance to engage in PA. Furthermore, we observed that the overall level of actual usage with the program components in the entire group did not influence the intervention effect.

The intervention had good acceptability for patients who rated their satisfaction in line with previous PA telecoaching research in a mixed COPD and diabetes type-2 population [25]. Higher acceptability scores might result in a higher chance of patients having more actual usage of the intervention. This was the case for the high ratings of the step counter by the patients, which was translated into excellent actual usage of the step counter throughout the trial. These high actual usage scores are in line with previous studies [7,9,13,14]. As the step counter was used as the medium to coach patients in this trial, we chose steps per day a priori as primary outcome of the effectiveness of the intervention, which is in line with the initial trial report of the MrPAPP trial [6]. However, one should note that PA encompasses not only amount (eg, steps per day) but also intensity (eg, time spent in moderate to vigorous PA) and time spent in different postures.

The smartphone app was also well received by patients although to a lesser extent than the step counter. This was associated with a considerably lower actual usage score of patients for the smartphone intervention compared with the step counter. Several factors may explain this relatively lower actual usage. First, a proportion of patients with COPD who owned a smartphone before the study might have caused less fluency with the smartphone (*low smartphone literacy*), leading to technical problems and discouraging smartphone use. Unfortunately, we do not have information on smartphone literacy at baseline. Furthermore, the actual usage rate of the smartphone tasks decreased over time. This was against our expectations, as one would expect that patients who have low smartphone literacy at the start of the trial (mostly those without a smartphone of their own) would increase their actual usage over time as they learn to operate the smartphone better. The latter learning effect was often catalyzed through the help of patient's relative (eg, [grand] children or spouse) and through the study team as reported by patients during the interviews. Second, findings from the semistructured interview revealed that patients felt the interaction with the app was often hindered due to Bluetooth and internet connection issues. Especially, the process of sending the step count data with the smartphone was perceived to be time consuming. This might have caused frustrations among patients, which could have initiated a decline of actual usage of the smartphone. Third, findings from the focus group and patients interviews revealed that patients felt the content of the smartphone app lacked variation (eg, daily repetition of morning messages with the same weekly goal). It presents another probable reason on why actual usage of the smartphone app was rather low and decreased over the 3 months of the trial. This could perhaps be improved by implementing components of gamification [26].

In literature, mixed results and high heterogeneity are reported on the actual usage with PA coaching Web portals or smartphone apps. During a 4-month, internet-based PA telecoaching program, veterans with COPD logged into the website and uploaded their daily step counts for 5.7 days per month which decreased to 3.0 days per month over a follow-up of 12 months [7,27]. Of note, the Web portal in the latter trial was not intended for daily use with a recommended frequency of 4 log-ins per month. The low degree of actual usage over a longer follow-up time was confirmed by a 9-month home-based pilot study, in which a smartphone-based activity coach was rarely used (only for 29 days throughout the whole trial) [28]. However, no information was provided on the change in actual usage over time in the latter trial [28].

Components of the intervention that were not individually tailored (eg, educational activity tips and home exercise booklet) were rated as less useful. This confirms patients' self-reported actual usage of the home exercise booklet, which was low and is in line with findings from the focus group, in which PA coaches pointed out that the home exercise booklet was not individualized for each specific patient. This highlights the importance of introducing personalized components within PA telecoaching, which has also been suggested in patients with ischemic heart disease who participated in a mobile health cardiac rehabilitation intervention [29,30].

In line with the patients, the coaches expressed good acceptability of this PA telecoaching program. On future use of the intervention, coaches reached the following consensus:

1. "The goal of such a PA telecoaching intervention should be that patients are able to use this intervention quasi independently indefinitely. Every 6 months patients could come for a follow-up visit, synchronized with other planned health visits to the outpatient clinic." Interestingly, our data suggest that 3 months of coaching might be enough for patients to reach a plateau in PA increase (see Multimedia appendix 10).

2. "As their PA coach it is our task to provide further follow-up by giving them the step counter and occasional phone calls for follow-up." Such strategies merit further validation, but the statement strengthens the importance of acceptability, actual usage, and feasibility with long-term PA telecoaching programs in this patient population. In addition to the latter perspectives, the coaches highlighted that it is highly important that the preferences and experiences of the patients with the intervention are assessed and taken into account when looking at future implementation. Therefore, future (long-term) PA telecoaching interventions need to ensure whether enough variation within such apps is introduced in addition to those components deemed as the most essential to patients (ie, step counter and contact with the study team). Furthermore, such interventions need to take the occurrence of acute exacerbations into account and involve patients' relatives as these can play an important role

as social support in being physically active [31], which was supported by the analyses of the interview data. Focusing on introducing new daily PA routines can provide a good starting point for long-term PA improvement according to these interview data.

In terms of coach feasibility, the main paper of the MrPAPP trial revealed that patients were contacted for a median of 50 min throughout the 12 weeks intervention [6]. Translated into socioeconomic terms, this means that coaching 25 patients simultaneously corresponds to approximately 2 hours per week for 1 PA coach. This number might even decrease as the coach accumulates his or her expertise or problem-solving efficiency, resulting in a lower burden.

Literature about the relationship of both actual usage by the patients and coach feasibility (contact time) of the intervention with the change in PA in telecoaching trials is scarce. In this study, the degree of the overall actual usage score (including wearing the step counter and all the app tasks) was not associated with the effectiveness of the intervention. This is in contrast to a 4-week pilot (telecoaching) study which showed a positive relationship between the degree of actual usage of wearing a smartphone-based activity coach and the benefits from the intervention during the first 2 weeks albeit this association disappearing during the third week [13]. Next to actual usage of the intervention by patients, actual usage by coaches is also crucial to how the intervention is delivered. Despite a high degree of actual usage of the PA telecoaching program by patients in the trial by Vorrink et al (ie, 89% of the days used) [32], the program was not able to induce significant improvements in PA [9]. The latter might be partly explained by the lack of feasibility from the part of the coaches. Due to financial reasons and time constraints, there was a low degree of actual usage of the primary care physiotherapists in using the foreseen website to adjust the patients' PA goals and to send motivating messages to the patients. In our trial, actual usage of the coaches could not be assessed in depth as we did not have information on the exact timing when coaches solved the tasks. The latter could have influenced the effectiveness of the intervention. However, the automated goal calculation and adjustment in our intervention could have partly limited the impact on the effectiveness of the intervention in comparison to the trial of Vorrink et al. This highlights the importance of introducing automated or semiautomated components in such interventions.

In contrast to actual usage, the contact time between the coach and patients was associated with the effectiveness of the intervention, that is, a lower effect in those patients in need of more contact time. These patients were the more severe (ie, they have more severe airflow obstruction and tend to have a lower functional exercise capacity) and are more likely to experience exacerbations and therefore, have more chance of triggering coaching-related and/or health-related contacts with their coach. As contact time remained a significant, negative predictor of the change in PA, independent of the patient characteristics, this may point to the inability of some patients to work with the coaching app. This corroborates with the findings of the qualitative part of the study and should not be ignored as a reason for treatment failure. In clinical practice, we would therefore advocate flexible use of these interventions where patients are diverted to other interventions (eg, more supervised exercise programs such as pulmonary rehabilitation) if contact time accumulates. This is important for stratification in future trials.

Strengths and Limitations

To the best of our knowledge, this study is the first providing an in-depth analysis of the acceptability, actual usage, and feasibility with a PA telecoaching intervention developed for patients with COPD. Our study is unique as it allows us to investigate these aspects, relating them to physiological characteristics along with the level of response.

The results are based on a combination of quantitative and qualitative research, including information coming from patients as well as from coaches. In addition, the study is performed on the back of a properly powered randomized controlled trial, which was characterized by a comprehensive physiological assessment and objective assessment of PA. Furthermore, this PA telecoaching intervention consists of several behavioral principles (including but not limited to facilitating goal setting, action planning, feedback, and problem solving) which were based on the behavior change technique taxonomy of Michie et al [33]. Nevertheless, some limitations need to be considered.

First, we only included patients that completed the trial. This could have resulted in a selection bias. Coaches might have spent more time in those patients who subsequently dropped out during their intervention period. However, as only 7% (12/171) of patients discontinued, this is unlikely to have had a large impact on the results. Second, no multiple-comparison post hoc corrections were applied in the quantitative data analysis as these analyses should be regarded as exploratory and in need of independent confirmation. These results help to guide future research; however, they may not be taken as a final judgment and should be interpreted with caution due to the latter limitation. Third, only 1 focus group with a limited number of PA coaches was performed. Therefore, data saturation could not have been reached. Another focus group with participants with a broad background and experience would have been of great value for (1) external validity of findings and (2) to ensure data saturation. Nevertheless, coaches were asked during the focus group whether they had additional comments. In addition, a summary of the focus group was sent to the coaches who could not be present at the focus group for completion of the summary. New themes emerged, which allowed for more data capturing. Fourth, we did not specifically assess capabilities or history of patients with managing the smartphone device or their expectations. In hindsight, this might have provided even more detailed information to predict the therapeutic response to the PA telecoaching intervention. Fifth, for the assessment of acceptability of the intervention, we used a project-tailored questionnaire. In literature, several attempts have been made to measure the quality of mobile health apps; however, no measure from a user perspective has been widely accepted [34-36]. Incorporating methodologies as proposed within the human computer interaction research and tools such as the mobile app rating scale (MARS) and uMARS (user version) tools (which were not available at the time of trial initialization) would have

strengthened the development and validity of the acceptability assessments in this paper [37,38]. Nevertheless, the findings of our project-tailored questionnaire still provide interesting insights into the acceptability with these kinds of interventions. Sixth, as proposed by the Medical Research Council, a process evaluation incorporates 3 themes (ie, implementation, mechanisms of impact, and context) [5]. The concepts of *implementation* and *mechanisms of impact* are largely covered in this paper by the assessments of actual usage, feasibility, and acceptability as well by their association with the effectiveness of the intervention. However, we were not able to evaluate the *context* theme (ie, how external factors had an impact on our intervention) in depth in this study. Seventh, as the cutoffs for making tertiles for contact time and actual usage score were driven by the data collected in this trial, they should not be regarded as clinically important cutoff points despite the wide range of contact time and actual usage scores presented in this paper. Finally, future research should investigate whether a feature for social interactions among peers might further lower the burden on health care providers. Such peer support has also been integrated as a catalyst for behavior change in the taxonomy of Michie et al given that privacy of patients is not breached [26,33].

Clinical Importance

In line with general findings of the present behavioral modification program [6], this paper shows that PA telecoaching is not an intervention to which all patients respond, but it is feasible and well received by the vast majority of patients. The number of smartphone users is increasing worldwide [39]. Given that it requires only modest health care resources and is relatively less time-consuming compared with one-to-one PA counseling, PA telecoaching does have opportunities for future implementation. Furthermore, the use of an electronic communication strategy might lower the burden on both clinicians and patients as we found a relatively low contact time of 50 min over 3 months of coaching. Moreover, it offers the possibility of coaching people from a distance [3]. The theoretical framework and proven effectiveness of this intervention also provides opportunities for its use in other elderly populations who are in need of being coached toward a more active lifestyle. In addition, findings of this paper provide possible guidance for the selection of patients that will benefit the most from these types of interventions. Patients with very limited exercise capacity, more symptoms, GOLD quadrants C or D, and/or a high amount of contact time during the first 4 weeks of the program are less likely to improve [6]. In these patients, further coaching input may be futile and other more intensive face-to-face interventions should be considered.

Conclusions

This 12-week PA telecoaching intervention was well accepted and feasible for both patients with COPD and their coaches. Actual usage of the step counter was excellent, whereas actual usage of the smartphone tasks was lower and decreased over time. Overall actual usage was not associated with the effect of the intervention. The step counter and direct contact with the coach were perceived as the most useful components of the intervention by the patients. Patients with more need for contact had more severe airflow obstruction, tended to have more severely limited exercise capacity, and experienced less PA benefits. Alternative strategies (including more face-to-face contacts and offering pulmonary rehabilitation programs) might be more effective in these patients.

Acknowledgments

The authors would like to acknowledge Claudia Perez for providing data from the Linkcare app, Maarten Spruit (MS) and Ane Arbillaga-Etxarri (AAE) for their contribution with the data collection. The PROactive project is funded by the Innovative Medicines Initiative Joint Undertaking (IMU JU) #115011. The Leuven study group was supported by the Flemish Research Foundation (grant # G.0871.13). HD was the recipient of a joint ERS/SEPAR Fellowship (LTRF 2015) and is a postdoctoral fellow of the FWO-Flanders. ZL was the recipient of a European Respiratory Society Fellowship, grant number LTRF 2016-6686 and is a postdoctoral fellow of the FWO-Flanders (Fellowship number 12U5618N). FMR is funded by The National Council for Scientific and Technological Development (CNPq), Brazil (249579/2013-8). The Zurich study group was supported by an additional grant of the Lung League Aargau (nonprofit organization) as well as by Swisscom AG who provided 30 sim cards and data usage of up to 1 GB per month. MIP's contribution to this work was supported by the National Institute for Health Research (NIHR) Respiratory Biomedical Research Unit at the Royal Brompton and Harefield National Health Services (NHS) Foundation Trust and Imperial College, London UK who part fund his salary. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The sponsors did not have any influence on the design, conduct, and analysis of the study. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya.

Authors' Contributions

RR, HD, ZL, AF, CDJ, IV, JGA, MIP, Maarten Spruit (MS), RT, NH, and TT contributed to the study protocol and development of the intervention. ML, RAR, HD, ZL, RT, NR, AF, CDJ, SCB, GB, AS, and IV contributed to the data collection. ML, HD, and TT contributed to the data analyses, interpretation of the data, and the writing of the paper. RAR, HD, ZL, RT, NR, AF, CDJ, EGS, FMR, SB, NH, GB, AS, IS, IV, JGA, MP, and TT critically reviewed the paper. TT is the guarantor of the study. All authors had full access to all the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Home exercise booklet.
[PDF File (Adobe PDF File), 6068 KB - mhealth_v6i12e200_app1.pdf]

Multimedia Appendix 2
Linkcare application platform.
[PPTX File , 1196 KB - mhealth_v6i12e200_app2.pptx]

Multimedia Appendix 3
Project-tailored patient satisfaction form.
[PDF File (Adobe PDF File), 53 KB - mhealth_v6i12e200_app3.pdf]

Multimedia Appendix 4
Patient interview (discussion guide).
[PDF File (Adobe PDF File), 46 KB - mhealth_v6i12e200_app4.pdf]

Multimedia Appendix 5
Focus group (coaches).
[PDF File (Adobe PDF File), 48 KB - mhealth_v6i12e200_app5.pdf]

Multimedia Appendix 6
COREQ checklist.
[PDF File (Adobe PDF File), 62 KB - mhealth_v6i12e200_app6.pdf]

Multimedia Appendix 7 [PDF File (Adobe PDF File), 601 KB - mhealth_v6i12e200_app7.pdf]

Multimedia Appendix 8
Illustrative quotes interviews.
[PDF File (Adobe PDF File), 520 KB - mhealth_v6i12e200_app8.pdf]

Multimedia Appendix 9
Actual usage of the different intervention tasks according to gender and age.
[PDF File (Adobe PDF File), 27 KB - mhealth_v6i12e200_app9.pdf]

Multimedia Appendix 10
Overview of (mean ± SE) steps performed per week by all patients (step counter data).
[PNG File , 24 KB - mhealth_v6i12e200_app10.png]

References

1. Gimeno-Santos E, Frei A, Steurer-Stey C, de Batlle J, Rabinovich RA, Raste Y, PROactive consortium. Determinants and outcomes of physical activity in patients with COPD: a systematic review. Thorax 2014 Aug;69(8):731-739 [FREE Full text] [doi: 10.1136/thoraxjnl-2013-204763] [Medline: 24558112]
2. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. GOLD executive summary. Am J Respir Crit Care Med 2017 Mar 01;195(5):557-582. [doi: 10.1164/rccm.201701-0218PP] [Medline: 28128970]
3. Field MJ. Telemedicine: A Guide to Assessing Telecommunications in Health Care. Washington (DC): National Academies Press (US); Oct 1996:2-3.
4. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. Br Med J 2008;337:a1655 [FREE Full text] [Medline: 18824488]
5. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J 2015;350:h1258 [FREE Full text] [Medline: 25791983]
6. Demeyer H, Louvaris Z, Frei A, Rabinovich RA, de Jong C, Gimeno-Santos E, Mr Papp PROactive study group and the PROactive consortium. Physical activity is increased by a 12-week semiautomated telecoaching programme in patients

with COPD: a multicentre randomised controlled trial. Thorax 2017 May;72(5):415-423 [FREE Full text] [doi: 10.1136/thoraxjnl-2016-209026] [Medline: 28137918]
7. Moy ML, Collins RJ, Martinez CH, Kadri R, Roman P, Holleman RG, et al. An internet-mediated pedometer-based program improves health-related quality-of-life domains and daily step counts in COPD: a randomized controlled trial. Chest 2015 Jul;148(1):128-137. [doi: 10.1378/chest.14-1466] [Medline: 25811395]
8. Van Der Weegen S, Verwey R, Spreeuwenberg M, Tange H, Van Der Weijden T, de Witte L. It's LiFe! Mobile and web-based monitoring and feedback tool embedded in primary care increases physical activity: a cluster randomized controlled trial. J Med Internet Res 2015 Jul 24;17(7):e184 [FREE Full text] [doi: 10.2196/jmir.4579] [Medline: 26209025]
9. Vorrink SN, Kort HS, Troosters T, Zanen P, Lammers JJ. Efficacy of an mHealth intervention to stimulate physical activity in COPD patients after pulmonary rehabilitation. Eur Respir J 2016 Oct;48(4):1019-1029. [doi: 10.1183/13993003.00083-2016] [Medline: 27587557]
10. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
11. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
12. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to ehealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
13. Tabak M, Vollenbroek-Hutten MM, van Der Valk PD, van Der Palen J, Hermens HJ. A telerehabilitation intervention for patients with Chronic Obstructive Pulmonary Disease: a randomized controlled pilot trial. Clin Rehabil 2014 Jun;28(6):582-591. [doi: 10.1177/0269215513512495] [Medline: 24293120]
14. Mendoza L, Horta P, Espinoza J, Aguilera M, Balmaceda N, Castro A, et al. Pedometers to enhance physical activity in COPD: a randomised controlled trial. Eur Respir J 2015 Feb;45(2):347-354 [FREE Full text] [doi: 10.1183/09031936.00084514] [Medline: 25261324]
15. Karsh B. Beyond usability: designing effective technology implementation systems to promote patient safety. Qual Saf Health Care 2004 Oct;13(5):388-394 [FREE Full text] [doi: 10.1136/qhc.13.5.388] [Medline: 15465944]
16. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
17. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. Br Med J 2010;341:c4587. [Medline: 20851841]
18. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
19. Bryman A. Integrating quantitative and qualitative research: how is it done? Qual Res 2006;6(1):97-113. [doi: 10.1177/1468794106058877]
20. Rabinovich RA, Louvaris Z, Raste Y, Langer D, Van Remoortel H, Giavedoni S, PROactive Consortium. Validity of physical activity monitors during daily life in patients with COPD. Eur Respir J 2013 Nov;42(5):1205-1215 [FREE Full text] [doi: 10.1183/09031936.00134312] [Medline: 23397303]
21. Van Remoortel H, Raste Y, Louvaris Z, Giavedoni S, Burtin C, Langer D, PROactive consortium. Validity of six activity monitors in chronic obstructive pulmonary disease: a comparison with indirect calorimetry. PLoS One 2012;7(6):e39198 [FREE Full text] [doi: 10.1371/journal.pone.0039198] [Medline: 22745715]
22. Bree R, Gallagher G. Using Microsoft Excel to code and thematically analyse qualitative data: a simple, cost-effective approach. AISHE-J 2016;8(2) [FREE Full text]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 200;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Demeyer H, Burtin C, Hornikx M, Camillo CA, Van Remoortel H, Langer D, et al. The minimal important difference in physical activity in patients with COPD. PLoS One 2016;11(4):e0154587 [FREE Full text] [doi: 10.1371/journal.pone.0154587] [Medline: 27124297]
25. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. Process evaluation of physical activity counselling with and without the use of mobile technology: a mixed methods study. Int J Nurs Stud 2016 Jan;53:3-16. [doi: 10.1016/j.ijnurstu.2015.10.008] [Medline: 26518108]
26. Richards D, Caldwell P. Gamification to Improve Adherence to Clinical Treatment Advice: Improving Adherence to Clinical Treatment Internet. In: Novak D, Tulu B, Brendryen H, editors. Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice. Hershey, PA: IGI Global; 2016:41-77.
27. Moy ML, Martinez CH, Kadri R, Roman P, Holleman RG, Kim HM, et al. Long-term effects of an internet-mediated pedometer-based walking program for chronic obstructive pulmonary disease: randomized controlled trial. J Med Internet Res 2016 Aug 08;18(8):e215 [FREE Full text] [doi: 10.2196/jmir.5622] [Medline: 27502583]

28. Tabak M, Brusse-Keizer M, van der Valk P, Hermens H, Vollenbroek-Hutten M. A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: a pilot randomized controlled trial. Int J Chron Obstruct Pulmon Dis 2014;9:935-944 [FREE Full text] [doi: 10.2147/COPD.S60179] [Medline: 25246781]
29. Buys R, Claes J, Walsh D, Cornelis N, Moran K, Budts W, et al. Cardiac patients show high interest in technology enabled cardiovascular rehabilitation. BMC Med Inform Decis Mak 2016 Jul 19;16:95 [FREE Full text] [doi: 10.1186/s12911-016-0329-9] [Medline: 27431419]
30. Pfaeffli DL, Whittaker R, Dixon R, Stewart R, Jiang Y, Carter K, et al. Acceptability of a mobile health exercise-based cardiac rehabilitation intervention: a randomized trial. J Cardiopulm Rehabil Prev 2015;35(5):312-319. [doi: 10.1097/HCR.0000000000000125] [Medline: 26181037]
31. Mesquita R, Nakken N, Janssen DJ, van den Bogaart EH, Delbressine JM, Essers JM, et al. Activity levels and exercise motivation in patients with COPD and their resident loved ones. Chest 2017 Dec;151(5):1028-1038. [doi: 10.1016/j.chest.2016.12.021] [Medline: 28087303]
32. Vorrink S, Huisman C, Kort H, Troosters T, Lammers J. Perceptions of patients with chronic obstructive pulmonary disease and their physiotherapists regarding the use of an eHealth intervention. JMIR Hum Factors 2017 Sep 19;4(3):e20 [FREE Full text] [doi: 10.2196/humanfactors.7196] [Medline: 28928110]
33. 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]
34. 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]
35. Baptista S, Oldenburg B, O'Neil A. Response to “development and validation of the user version of the mobile application rating scale (uMARS)”. JMIR Mhealth Uhealth 2017 Jun 09;5(6):e16 [FREE Full text] [doi: 10.2196/mhealth.6419] [Medline: 28600277]
36. Powell AC, Torous J, Chan S, Raynor GS, Shwarts E, Shanahan M, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. JMIR Mhealth Uhealth 2016;4(1):e15 [FREE Full text] [doi: 10.2196/mhealth.5176] [Medline: 26863986]
37. 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]
38. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
39. eMarketer. Statista. New York, NY, US; 2016 Jun. Number of smartphone users worldwide from 2014 to 2020 URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [accessed 2017-10-15] [WebCite Cache ID 6ukYY3vl6]

Abbreviations

COPD: chronic obstructive pulmonary disease
COREQ: COnsolidated criteria for REporting Qualitative research
FEV_1: forced expiratory volume in 1 second
GOLD: Global Initiative for Chronic Obstructive Lung Disease
MARS: mobile app rating scale
MID: minimal important difference
MrPAPP: multicenter physical activity telecoaching trial
mMRC: modified Medical Research Council
PA: physical activity
6MWD: 6-minute walking distance

Edited by G Eysenbach; submitted 25.01.18; peer-reviewed by M Vollenbroek-Hutten, S Van den Bulck, Y Bartlett; comments to author 21.03.18; revised version received 30.05.18; accepted 24.09.18; published 21.12.18.

Please cite as:
Loeckx M, Rabinovich RA, Demeyer H, Louvaris Z, Tanner R, Rubio N, Frei A, De Jong C, Gimeno-Santos E, Rodrigues FM, Buttery SC, Hopkinson NS, Büsching G, Strassmann A, Serra I, Vogiatzis I, Garcia-Aymerich J, Polkey MI, Troosters T
Smartphone-Based Physical Activity Telecoaching in Chronic Obstructive Pulmonary Disease: Mixed-Methods Study on Patient Experiences and Lessons for Implementation
JMIR Mhealth Uhealth 2018;6(12):e200
URL: http://mhealth.jmir.org/2018/12/e200/
doi:10.2196/mhealth.9774
PMID:30578215

Original Paper

Assessing the Efficacy of an Educational Smartphone or Tablet App With Subdivided and Interactive Content to Increase Patients' Medical Knowledge: Randomized Controlled Trial

Thomas Timmers[1,2], MSc; Loes Janssen[3], PhD; Yvette Pronk[4], MSc; Babette C van der Zwaard[5], PhD; Sander Koëter[6], MD, PhD; Dirk van Oostveen[5], MD; Stefan de Boer[3], MD; Keetie Kremers[6], MPA; Sebastiaan Rutten[4], MD; Dirk Das[7], MD; Rutger CI van Geenen[8], MD, PhD; Koen LM Koenraadt[8], PhD; Rob Kusters[9], Prof Dr; Walter van der Weegen[7], PhD

[1]Interactive Studios, Rosmalen, Netherlands
[2]Radboud University Medical Center, IQ Healthcare, Nijmegen, Netherlands
[3]VieCuri Medical Centre, Venlo, Netherlands
[4]Kliniek ViaSana, Mill, Netherlands
[5]Jeroen Bosch Ziekenhuis, s-Hertogenbosch, Netherlands
[6]Canisius-Wilhelmina Hospital, Nijmegen, Netherlands
[7]Sint Anna Ziekenhuis, Geldrop, Netherlands
[8]Amphia Hospital, Breda, Netherlands
[9]Open University of The Netherlands, Heerlen, Netherlands

Corresponding Author:
Thomas Timmers, MSc
Interactive Studios
Huisbergenweg 6
Rosmalen, 5249 JR
Netherlands
Phone: 31 73 644 6069
Fax: 31 84 8710356
Email: thomas@interactivestudios.nl

Abstract

Background: Modern health care focuses on shared decision making (SDM) because of its positive effects on patient satisfaction, therapy compliance, and outcomes. Patients' knowledge about their illness and available treatment options, gained through medical education, is one of the key drivers for SDM. Current patient education relies heavily on medical consultation and is known to be ineffective.

Objective: This study aimed to determine whether providing patients with information in a subdivided, categorized, and interactive manner via an educational app for smartphone or tablet might increase the knowledge of their illness.

Methods: A surgeon-blinded randomized controlled trial was conducted with 213 patients who were referred to 1 of the 6 Dutch hospitals by their general practitioner owing to knee complaints that were indicative of knee osteoarthritis. An interactive app that, in addition to standard care, actively sends informative and pertinent content to patients about their illness on a daily basis by means of push notifications in the week before their consultation. The primary outcome was the level of perceived and actual knowledge that patients had about their knee complaints and the relevant treatment options after the intervention.

Results: In total, 122 patients were enrolled in the control group and 91 in the intervention group. After the intervention, the level of actual knowledge (measured on a 0-36 scale) was 52% higher in the app group (26.4 vs 17.4, $P<.001$). Moreover, within the app group, the level of perceived knowledge (measured on a 0-25 scale) increased by 22% during the week within the app group (from 13.5 to 16.5, $P<.001$), compared with no gain in the control group.

Conclusions: Actively offering patients information in a subdivided (per day), categorized (per theme), and interactive (video and quiz questions) manner significantly increases the level of perceived knowledge and demonstrates a higher level of actual knowledge, compared with standard care educational practices.

Trial Registration: International Standard Randomized Controlled Trial Number ISRCTN98629372; http://www.isrctn.com/ISRCTN98629372 (Archived by WebCite at http://www.webcitation.org/73F5trZbb)

(JMIR Mhealth Uhealth 2018;6(12):e10742) doi:10.2196/10742

KEYWORDS

patient education; shared decision making; smartphone; decision aid; orthopedics

Introduction

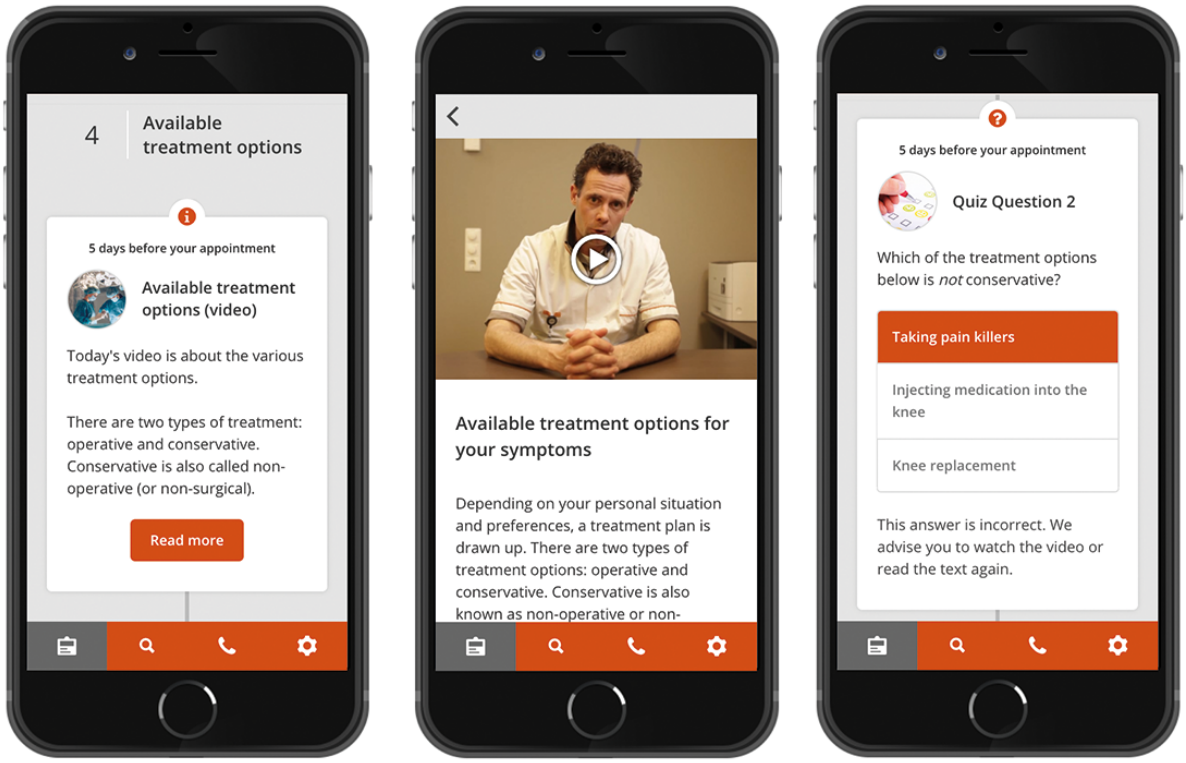
Background

Shared decision making (SDM) refers to the process that involves the participation of both the physician and patient to select the best suitable treatment, taking into account the clinical data and patients' preferences and expectations [1]. Modern health care increasingly focuses on SDM because of its positive effects on patient satisfaction, therapy compliance, and outcomes [2,3]. One of the key drivers of SDM is the patients' knowledge about the illness and treatment options available [4]. This knowledge is typically acquired through patient education, and that knowledge is currently primarily transmitted during medical consultations. Unfortunately, previous studies have shown that patients scarcely remember doctors' reports after their consultations and that their memory for medical information is substantially limited [5]. Indeed, recent research has indicated that, on average, approximately 40% to 80% of the information provided to patients by health care practitioners was immediately forgotten, and out of what the patients did recall, approximately half of the content was inaccurate [6].

Several factors, some of which are difficult to change, contribute to poor memory acquisition, including patient age [7-9], patient level of education [10-12], the fact that too much information is provided in too little time [13-15], and that is likely exasperated by doctors' busy schedules and their difficult use of language and jargon [5,16,17]. On the other hand, there are numerous factors that can positively influence patient recall, for example, subdividing the delivery of information [11,13-15] and the explicit categorization of content [18]. Furthermore, the usage of questions and feedback to test (and reflect) patients' knowledge [11,19] as well as the specific modality of information transmission contribute to patient recall of medical information. Indeed, people tend to remember 20% auditory information, 40% of read information, and up to 80% of information acquired from interactive education [20-22].

Technology for health (electronic health [eHealth] or mobile health) plays an increasingly important facilitating role in educating patients [23]. The information is always available, consistent, and complete [24]; the doctor is always welcoming; patients can determine their own pace [25]; and the information can even be tailored to patients' personal needs [26].

Figure 1. Examples of the interactive app. From left to right: the interactive timeline, information about a certain topic, and quiz-like questions with feedback.



The goal of this study was to examine the effectiveness of using an interactive app (Figure 1) in patients who were referred to the hospital by their general practitioner (GP) owing to knee complaints that indicated osteoarthritis (OA). Knee OA is a progressive condition that causes pain and functional limitations [27]. In the early phase of OA, it can be treated conservatively. End-stage OA, however, is most effectively treated with joint replacement surgery [28]. Knowledge about their condition and treatment options is important for these patients to make a well-considered decision between conservative or surgical treatment.

Objectives

We hypothesize that compared with the above-mentioned standard practices of educating patients, the use of an interactive app would lead to a higher level of knowledge (perceived and actual) about their illness and treatment options. This is the primary outcome of our study. With regard to the secondary outcomes, we hypothesize that there will be a relative increase in reported patient satisfaction concerning their level of knowledge and the amount of information available. In addition, it is hypothesized that using the interactive app would positively influence patients' general satisfaction with their consultation and their confidence in the treatment choice they made. All outcomes are measured by means of Web-based questionnaires.

Methods

Study Design

A total of 6 hospitals (4 nonacademic teaching hospitals, 1 general hospital, and 1 specialized orthopedic clinic) were selected. Between April and September 2017, patients with knee complaints due to OA were asked to participate in a surgeon-blinded randomized controlled trial. In this study, the effectiveness of an interactive app on patients' knowledge, satisfaction, and certainty about the treatment chosen was assessed. This was compared with standard education in a parallel group design with equal allocation ratio. No changes were made to the design after the study was commenced.

Informed Consent and Ethical Consideration

Patients were asked to consider participating in the study after scheduling an appointment at one of the recruiting hospitals. Patients who were willing received an email with all the necessary study information required for informed consent. Patients were offered at least 2 days to reflect on the information. In the case of any questions, patients were informed that they could contact the local research coordinator from each specific hospital by phone or email. Patients indicated their consent by signing an online informed consent form. Patients were also informed that their data would be kept confidential and protected. There were no indicators of substantial risk as a function of participating in this study. The study was registered at ISRCTN, with reference number ISRCTN98629372. Due to technical problems during the initial, prospective registration, the study was registered retrospectively. Registration took place after the study was completed, on May 13, 2018. The study was approved by the regional Medical Ethical Board of the Maxima Medisch Centrum (Eindhoven, The Netherlands), reference number N16.130, as well as at each of the participating sites. In addition, we attest that we have obtained appropriate permissions and paid any required fees for use of copyright-protected materials.

Participant Selection

The eligibility of patients was assessed during their first contact with the hospital to schedule their appointment with the orthopedic surgeon. Patients had to be older than 40 years and referred by their GP because of knee complaints indicating OA. Participants were required to be fluent in Dutch and in the possession of an email address and a smartphone or tablet. At least 10 days between scheduling the appointment and the hospital visit were required, to give patients in the app group the chance to experience the intervention.

Randomization

Patients who considered participation were controlled and randomly assigned by a computer to either a control or app group as soon as they were registered in the Web-based system by the hospital staff. Randomization was performed without block or stratification restrictions. Participants were not informed of which group they were assigned to, although both groups received an email with all the information about the study. From this email, patients who chose to participate in the study could directly give their Web-based informed consent and fill out the baseline questionnaire. Patients allocated to the app group received an additional email after completing the baseline questionnaire. This email contained download instructions for the app, a Google Play and Apple App Store download link, and the patients' personal code. Thereafter, both groups simultaneously received the same questionnaires again on 2 separate occasions: 2 days before the arranged consultation and 1 day after the consultation. Per questionnaire, a maximum of 2 email reminders was sent.

Intervention

In this study, the Patient Journey App (Interactive Studios, Rosmalen, The Netherlands) was used as the intervention. By using push notifications, we actively offered patients information about knee OA, (conservative and operative) treatment options, risks, rehabilitation, and expectancies in a subdivided (daily) and categorized (per theme) manner. Information was presented on an interactive timeline using text, photos, and video content (Figure 1). Interactive quiz-like questions were used to test their knowledge, providing direct feedback on the given answer.

The content for the app was compiled based on the input of 10 orthopedic surgeons from various hospitals, the Dutch option grid for knee OA [29], and information booklets from 3 participating hospitals. The 5 most important topics, as agreed upon by the surgeons, were (1) knee anatomy and the origin of the complaints, (2) different types of conservative and operative treatments, (3) risks of surgery, (4) rehabilitation after total knee replacement, and (5) expectations after total knee replacement. These topics also formed the base for the questionnaires addressing perceived and actual knowledge. Both the control group and the app group had access to standard education as offered by the hospitals, consisting of at least a website and an

information event. Only the app group received the app, protected with a personal code.

Patients used the app in the 7 days before the first consultation with their orthopedic surgeon. During the first 5 days, information concerning the 5 most important topics was provided, whereas on days 6 and 7, a summary as well as practical information on how to prepare for the consultation itself were provided. Patients received daily push notifications at 10:00 am. During the study, no changes or revisions to the app took place.

Study Outcomes

Study outcomes were measured at 3 moments in time: baseline, 2 days before consultation, and 1 day after consultation (Table 1). The baseline measurement commenced directly after patients were included in the study. Due to the timing of the intervention, the first follow-up measurement was scheduled 2 days before the consultation—enabling patients to complete the questionnaires before their hospital visit. To assure accurate recall of the consultation, the third measurement was scheduled 1 day after consultation.

All measurements were performed by using patient-reported questionnaires. Most questionnaires were developed especially for this study and can be found in Multimedia Appendix 1. All self-developed questionnaires were validated by surgeons and researchers of all participating hospitals. No additional validation was performed in a patient population.

The primary outcome measure was patient knowledge about the illness and the available treatment options. We divided knowledge into 2 concepts: perceived knowledge (ie, how much do patients think they know) and actual knowledge (ie, how much do the patients actually know). The questionnaires were based on the aforementioned 5 most important topics for the first consultation. In the perceived knowledge questionnaire, patients received 5 questions, with answers ranging from 1 (no knowledge at all) to 5 (best imaginable knowledge). The perceived knowledge questionnaire had a sum score ranging from 5 to 25. Perceived knowledge was measured at baseline and 2 days before the consultation. The actual knowledge questionnaire required patients to answer 12 questions, each ranging from 0 (incorrect answer) to 3 points (correct answer). The absolute knowledge questionnaire had a sum score range from 0 to 36. Actual knowledge was measured only once, 2 days before the consultation. The actual knowledge questionnaire was administered only once, as answering it could prime patients, which could influence their performance at future participation.

As secondary outcomes, we assessed patients' satisfaction with the provided information, satisfaction with their level of knowledge, and their need for more information. This questionnaire was developed for this trial. Numeric Rating Scale (NRS) scores were used to measure these outcomes. Questions concerning satisfaction had a range from 0 (not satisfied at all) to 10 (very satisfied). Questions concerning the need for more information had a range from 0 (no need at all) to 10 (very much in need of). Satisfaction and the need for more information were measured at baseline and 2 days before the consultation. Furthermore, 1 day after the consultation, we measured overall satisfaction with the consultation with an NRS score from 0 (not satisfied at all) to 10 (very satisfied). We also determined the level at which patients felt they had made a decision about their treatment together with their physician, with an NRS score ranging from 0 (strongly disagree) to 10 (strongly agree). Furthermore, the following items were measured: the type of treatment chosen (conservative, operative, and I don't know), how sure patients were about their choice (NRS, from 0 [not sure at all] to 10 [very sure]), and whether their complaints were, in the end, actually because of knee OA (yes, no, and I don't know). All tertiary outcomes were measured 1 day after the consultation. This questionnaire was developed for this trial.

As a fourth outcome, we assessed patients' mobile device proficiency. To measure it, the Mobile Device Proficiency Questionnaire-16 [30] was used. This questionnaire addresses 8 domains, ranging from "sending an email" and "downloading apps" to "privacy" and "update settings." Each domain is assessed by 2 questions about completing a task, measured on a Likert scale from 1 (never tried) to 5 (very easily), resulting in a sum score ranging from 16 to 80. As a Dutch version of this questionnaire was not available, it was translated from English to Dutch by 3 researchers independently. After reaching a consensus about the Dutch translation, it was translated back into English by a certified translation agency. No major differences to the original version were identified. This measure was performed at baseline.

Table 1. Overview of outcomes per measurement.

Baseline	2 days before consultation	1 day after consultation
Patient characteristics	Actual knowledge	Satisfaction with consultation
Perceived knowledge	Perceived knowledge	Type of treatment chosen
Satisfaction with information	Satisfaction with information	Certainty about the choice
Satisfaction with knowledge	Satisfaction with knowledge	—[a]
Need for more information	Need for more information	—
Mobile device proficiency	—	—

[a]Not applicable.

Sample Size and Statistical Analysis

A priori sample size calculation (alpha=.05, [1−beta]=.80) based on a reported difference of 11.7% in knowledge between orthopedic patients using a decision aid or not [31] resulted in a minimum requirement of 83 patients in each study arm. Our primary analysis was conducted using an intention-to-treat approach and therefore included all randomized patients. Normally distributed continuous variables were presented as a mean value with the SD and statistically compared between the groups using independent Student *t* tests. Non-normally distributed variables were presented as a median value with the interquartile range and statistically compared between the groups using the Mann-Whitney U tests. Categorical variables were presented as number and percentage and compared between groups using chi-square tests. Within-group differences were tested using paired Student *t* tests in the case of normally distributed data or Wilcoxon signed-rank tests in the case of nonparametric data. *P* values of ≤.05 were assumed to indicate a significant difference. A "per protocol" analysis was performed for all primary and secondary outcomes to also examine to the robustness of our main results. All data were analyzed using IBM SPSS Statistics for Macintosh, version 22.0, (Armonk, USA).

Results

Study Sample

Between May and August 2017, a total of 307 patients considered participation in the study. A total of 50 patients (16.2%, 50/307) withdrew from the study (for reasons unknown) without completing the baseline questionnaire, and 1 patient did not consent to participate. Moreover, 11 patients (3.5%, 11/307) could not be contacted because of incorrect email addresses. In addition, 32 patients (10.4%, 32/307) participated in the baseline questionnaire but did not respond to both follow-up questionnaires (Figure 2).

Patients who missed the baseline measurement or both the follow-up questionnaires were registered as lost to follow-up. Patients whose email address was incorrectly recorded, and therefore did not receive an invitation, were registered as "email address unknown or incorrect." Neither of them were included in the analysis. No significant differences were found between the baseline characteristics of the app group (45 male [49%], mean age=62.3 years, SD=8.3) and the control group (56 male [45%], mean age=61.8 years, SD=8.5). In addition, no significant differences were observed with respect to level of education, pain, symptoms, functional outcome, perceived knowledge, satisfaction with available knowledge, and the need for additional information (Table 2).

Patient Knowledge Acquisition

Two days before consultation, the app group had a 52% higher level of actual knowledge (app: mean 26.4 [SD 7.4], control: mean 17.4 [SD 6.8], *P*<.001; Figure 3 and Table 3). The level of perceived knowledge was 26% higher in the app group (app: mean 16.5 [SD 3.9], control: mean 13.0 [SD 4.1], *P*<.001).

Comparison within groups revealed an increase in perceived knowledge in the app group (baseline: mean 13.1 [SD 4.6], 2 days before consultation: mean 16.5 [SD 3.9], *P*<.001). This was not the case in the control group (baseline: mean 13.5 [SD 4.1], 2 days before consultation: mean 13.6 [SD 4.2], *P*=.78; Figure 4).

Patient Satisfaction

Patients' level of satisfaction with their knowledge was higher in the app group (app: mean 6.8 [SD 2.7], control: mean 5.4 [SD 2.5], *P*<.001). The level of satisfaction with the provided information was also higher in the app group (app: mean 7.0 [SD 2.3], control: mean 5.3 [SD 2.5], *P*<.001). The app group also had a lower need for additional information (app: median 7 [Q1-Q3 5-8], control: median 8 [Q1-Q3 6-9], *P*=.02).

Comparison within groups revealed an increase in patients' satisfaction with their knowledge in the app group (baseline: mean 5.18 [SD 2.8], 2 days before consultation: mean 6.8 [SD 2.7], *P*<.001). This was not the case in the control group (baseline: mean 5.3 [SD 2.4], 2 days before consultation: mean 5.4 [SD 2.5], *P*=.57).

Comparison within groups also revealed an increase in patients' satisfaction with the available information in the app group (baseline: mean 5.2 [SD 2.7], 2 days before consultation: mean 7.0 [SD 2.3], *P*<.001). This was not the case in the control group (baseline: mean 5.1 [SD 2.3], 2 days before consultation: mean 5.3 [SD 2.5], *P*=.56).

Consultation

Overall satisfaction with the consultation with the orthopedic surgeon showed no difference between groups (app: median 9 [Q1-Q3 8-9], control: median 9 [Q1-Q3 7-9], *P*=.32). The extent to which patients felt they chose their treatment together with the orthopedic surgeon also did not differ between groups (app: median 9 [Q1-Q3 7-9], control: median 8 [Q1-Q3 7-8], *P*=.25).

Treatment

Patients in the app group were more confident about their choice of treatment (app: median 8 [Q1-Q3 7-10], control: median 8 [Q1-Q3 5-8], *P*=.04). There was no difference between groups concerning the choice for conservative or operative treatment (*P*=.34). In the control group, more patients were uncertain of their choice of treatment (22.3% vs 10.1%, *P*=.03). In addition, the control group had significantly more patients that reported they had no idea whether their complaints were, in the end, caused by OA (26.3% vs 10.1%, *P*=.02).

Figure 2. Patient flow diagram.

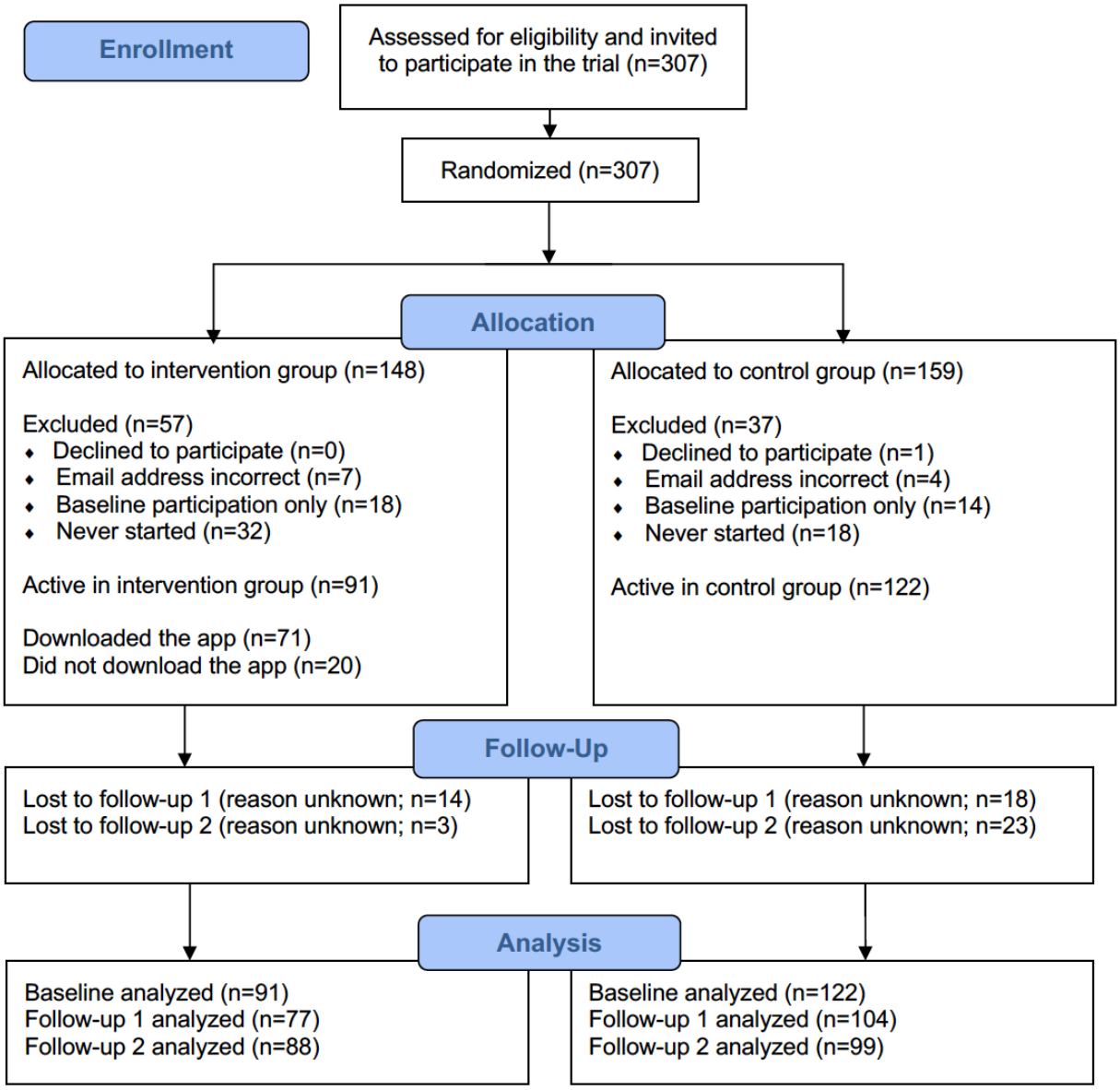


Table 2. Patient characteristics.

Characteristics	App group (n=91)	Control group (n=122)	*P* value
Sex, n (%)			
Male	45 (49)	56 (45.9)	—[a]
Female	46 (50)	66 (54.1)	.61
Education, n (%)[b]			
Group 1[c]	56 (62)	73 (64.0)	—
Group 2[d]	33 (37)	41 (36.0)	.87
Duration of complaints >6 months[e,f], n (%)	57 (62)	70 (57.4)	.44
Walking <30 min[f,g], n (%)	75 (82)	100 (82.0)	.93
Pain during the night[f,h], n (%)	21 (23)	28 (23.0)	.98
Age, mean (SD)	62.27 (8.32)	61.75 (8.54)	.66
KOOS PS[i,j], mean (SD)	22.96 (5.19)	22.73 (5.44)	.75
Perceived knowledge, mean (SD)	13.04 (4.41)	13.39 (4.14)	.56
Satisfaction with knowledge, mean (SD)	5.34 (2.77)	5.22 (2.46)	.72
Pain, mean (SD)			
At rest[j,k]	4.91 (2.61)	5.02 (2.58)	.77
During activity[j,k]	7.16 (1.91)	6.94 (2.23)	.45
Mobile Device Proficiency Questionnaire-16, mean (SD)	59.31 (19.73)	60.28 (18.77)	.97
Satisfaction with information, median (Q1-Q3)	6.00 (3-7)	5.00 (4-7)	.40
Need for more information, median (Q1-Q3)	9.00 (8-10)	8.00 (7-10)	.88

[a]Not applicable.

[b]Level of education has been split into 2 groups for the purpose of analysis.

[c]Educational levels in group 1: none, elementary school, and secondary (vocational) education.

[d]Educational levels in group 2: higher secondary education, pre-university education, and university (of applied science).

[e]Duration of complaints has been split into 2 groups for analysis purposes and was measured categorically (<3 months, 3-6 months, 6-12 months, and >12 months).

[f]Typical complaints for knee osteoarthritis patients, advised by participating orthopedic surgeons.

[g]Ability to walk for 30 min was measured as a dichotomous variable. Data represents patients who answered “no”.

[h]Pain at night was measured as a dichotomous variable. Data represents patients who answered “yes”.

[i]The Knee injury and Osteoarthritis Outcome Score-Physical Function Shortform (KOOS PS) [32] was used to assess functional outcome.

[j]Part of The Netherlands Orthopaedic Association guideline for knee osteoarthritis.

[k]Pain at rest and during activity was measured using a Numeric Rating Scale (NRS) from 0 (no pain) to 10 (severe pain).

Figure 3. Level of actual knowledge (2 days before consultation; error bars: 95% CI).

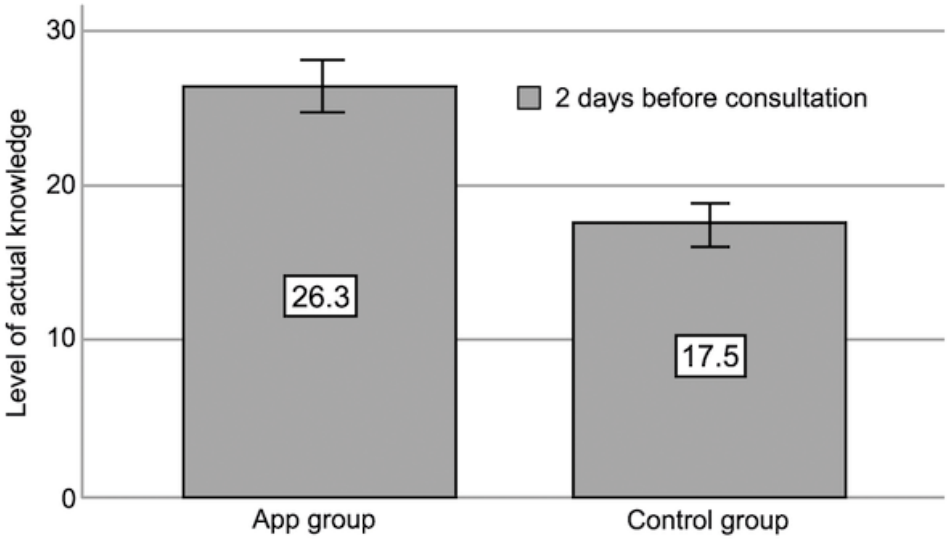


Table 3. Actual knowledge questionnaire and distribution of the correct answers.

Question	App, n (%)	Control, n (%)	*P* value
1. What is knee osteoarthritis? a. A type of pain relief; *b. Wear and tear of the knee joint*[a]; c. The conservative treatment of knee problems; d. I don't know	75 (96.2)	101 (96.2)	.99
2. In what way does osteoarthritis cause knee problems? a. Poor circulation in the leg; *b. Deterioration of the cartilage quality*; c. Injury to the knee caused by work or sport; d. I don't know	70 (89.7)	87 (82.9)	.19
3. Which of the following treatments is not a conservative treatment? a. Walking with crutches or a stick, possibly combined with physiotherapy; b. Injection in the knee; *c. Placement of a knee prosthesis*; d. I don't know	51 (65.4)	37 (35.2)	<.001
4. What is the average life of a knee prosthesis? a. An average of 5-10 years; b. An average of 10-15 years; *c. An average of 15-20 years*; d. I don't know	44 (56.4)	32 (30.5)	<.001
5. Which of the following preparations are important to reduce the risk of complications during an operation? More than one answer can be correct[b]; *a. Stop smoking*; *b. Certain physiotherapeutic exercises*>; c. Stop exercising; *d. Healthy eating*; e. Stop working; f. I don't know	43 (55.1)	31 (29.5)	<.001
6. How often does the knee prosthesis become infected so that it needs to be replaced? *a. In about 1 percent of cases*; b. In about 5 percent of cases; c. In about 10 percent of cases; d. I don't know	32 (41.0)	14 (13.3)	<.001
7. A possible complication of a knee prosthesis is thrombosis (blood clot) in the legs. How can this be avoided? a. Avoid overextending the operated leg for 4-6 weeks after the operation; b. Walk with crutches or a stick for 4-6 weeks after the operation; *c. Use blood thinners for 4-6 weeks after the operation*; d. I don't know	58 (74.4)	50 (47.6)	<.001
8. What is the duration of the average hospital stay for patients who have received a knee prosthesis? a. *1-3 days*; b. 4-7 days; c. 7-10 days; d. I don't know	67 (85.9)	63 (60.0)	<.001
9. How many months on average will you receive physiotherapy after you have had a knee prosthesis? a. Less than a month; b. 1-3 months; *c. 3-6 months*; d. I don't know	41 (56.6)	29 (27.6)	<.001
10. How long on average will it take until you have fully recovered after a knee prosthesis operation? a. 1-3 months; b. 3-6 months; *c. 6-12 months*; d. I don't know	53 (67.9)	36 (34.3)	<.001
11. Which of the following statements about a knee prosthesis are true? More than one answer can be correct[b]; *a. For many patients the pain will decrease, allowing them to move more easily*; b. It is safe to partake in activities such as basketball, football, and volley ball; *c. After 2-3 months, many patients are able to resume part of their daily activities*; *d. It is safe to partake in activities such as walking, swimming, and cycling*; e. I don't know	55 (71.4)	54 (51.4)	.01
12. What percentage of patients will be completely without pain after receiving a knee prosthesis? a. 65-70%; b. 75-80%; *c. 85-90%*; d. I don't know	51 (65.4)	17 (16.2)	<.001

[a]Italics indicate the correct answer.

[b]Only the combination of all 3 correct answers was indicated as "correct".

Figure 4. Level of perceived knowledge (baseline vs 2 days before the consultation; error bars: 95% CI).

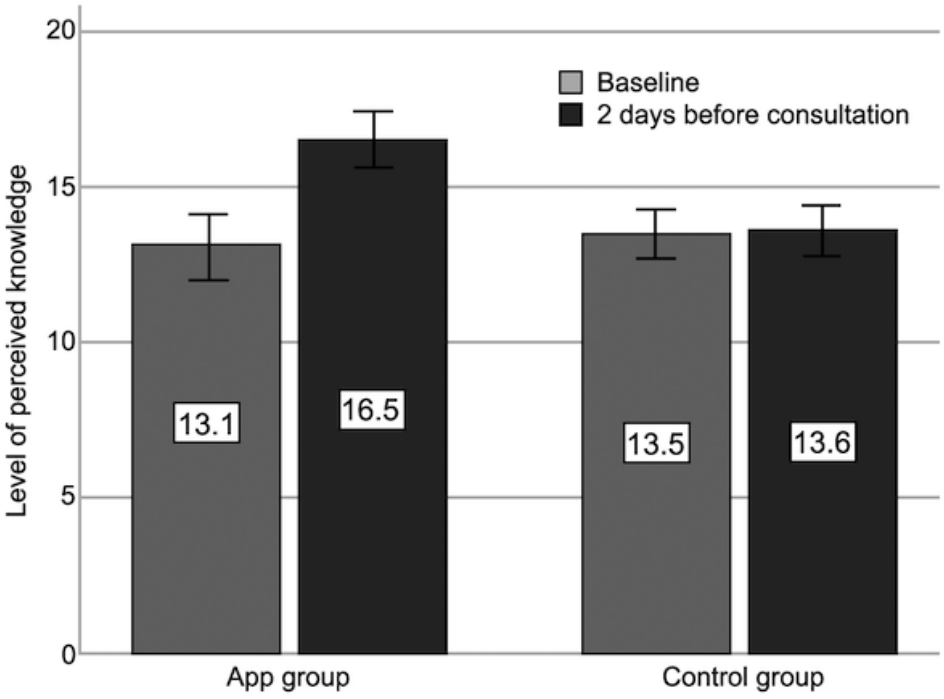


Table 4. Per protocol analysis of the main outcomes (2 days before consultation).

Outcome	Intention to treat			Per protocol		
	App (n=91), mean (SD)	Control (n=122), mean (SD)	*P* value	App (n=71), mean (SD)	Control (n=142), mean (SD)	*P* value
Actual knowledge	26.4 (7.4)	17.4 (6.8)	<.001	27.6 (7.3)	17.9 (6.8)	<.001
Perceived knowledge	16.5 (3.9)	13.0 (4.1)	<.001	17.0 (3.7)	13.6 (4.1)	<.001
Satisfaction with knowledge	6.8 (2.7)	5.4 (2.5)	<.001	6.8 (2.2)	5.5 (2.5)	<.001
Satisfaction with information	7.0 (2.3)	5.3 (2.5)	<.001	7.1 (3.9)	5.4 (4.1)	<.001

Mobile Device Proficiency of the Population

There was no difference in mobile device proficiency between groups at baseline (app: mean 59.3 [SD 19.73], control: mean 60.3 [SD 18.77]). The items most frequently referred to as "never tried" were the transferring of data to and from a mobile device, scheduling appointments in their agenda, playing games, and listening to music. All the items necessary for the use of the educational app were rated "easily" or "very easily" by >75% of patients. These items included using the device to find and start the app and using the keyboard. About 24% of the participants had never tried to search for an app in the App or Play Store.

Per Protocol Analysis

All presented results so far were analyzed using the intention-to-treat method. Analysis based on the per protocol method also resulted in the main outcomes being in favor of the app group, albeit somewhat more pronounced (Table 4).

Discussion

Principal Findings

In this study, we primarily investigated the possible effects of actively sending subdivided, categorized, and interactive information through an app on patients' knowledge of knee complaints and their treatment options. In addition, patient satisfaction with provided information, knowledge, and the consultation as well as patient self-reported confidence in treatment choice were measured. In comparison with patients who received standard care, the level of actual knowledge was 52% higher in patients who used the app designed for this study. This approach seems to be much more effective compared with the use of decision aids as described in the Cochrane 2017 systematic review, in which a total of 52 studies, calculating for 13,316 patients, were included, and knowledge increased by 13.27% [3].

Apart from actual knowledge, patients in the app group also experienced a significantly higher level of perceived knowledge about their illness and treatment options, whereas in the control group, there was hardly any change. The app group rated their level of perceived knowledge as 16.5 out of 20 (ie, 8.3 on a 0-10

scale). The control group rated their level of knowledge as 13.5 out of 20 (ie, 6.8 on a 0-10 scale). Their level of actual knowledge, however, demonstrated that they sometimes had little to very little knowledge about the topics (see Table 3). This *overestimation* of one's knowledge fosters their confidence of being capable of choosing the right treatment [33]. Choosing a treatment based on nonexistent or wrong information, however, is a known predictor for dissatisfaction [34-36].

Patients in the app group not only had more knowledge but were also more satisfied with their knowledge and the information they received and were to a lesser extent in need of more information. This is in line with earlier research, describing increased satisfaction when one is offered a decision aid [3]. Nonetheless, no differences were found between the groups in the way they experienced their consultation with their physician nor in the extent to which they felt they had made the choice for their treatment together with their physician.

On the basis of these results, one might question patients' need for knowledge. However, previous studies have demonstrated that when patients are more knowledgeable, they have less decisional conflict [3]. This is arguably supported by our finding that patients in the app group were more confident about their chosen treatment. Furthermore, in the days directly following the consultation, 90.2% of patients in the app group could remember the type of treatment chosen versus 78.3% in the control group (P=.02). In addition, the number of patients who did not know whether their complaints were caused by OA was smaller in the app group (10.3%) than in the control group (26.7%; P=.01).

Comparison With Prior Work

To try to explain the large difference in actual knowledge gain, we took a closer look at the Cochrane 2017 review. This review updated the Cochrane 2014 review on the use of decision aids, to which 18 new studies were added. We examined these 18 studies, assuming they would provide an up-to-date overview of the types of interventions used in recent years. All newly added studies were conducted between 2012 and 2015, except for one that was conducted in 2006. In these studies, booklets, DVDs, websites, one-on-one conversations, phone calls, and group sessions were used as decision aids. Decision aids were made available during consultations, between consultations, or after consultations with doctors. Decision aids ranged from 1-page instructions to 2-hour information sessions online or on-site.

In contrast to these studies, we used an app for smartphone or tablet as an information carrier in our study. One of the characteristics of these devices, especially smartphones, is that people often carry the device with them, lowering the barrier to use them. Within the app, we used a combination of known mechanisms on information retention: small bits of information [11,13-15], information about specific themes [6], multiple modes of information [37], and quiz-like questions with instant feedback to test (and reflect) patients' knowledge [11,19]. None of the studies in the Cochrane review used this combination of mechanisms. Type of information carrier was often the limiting factor for using different mechanisms, as you cannot, for instance, offer subdivided content on a piece of paper, small bits of information in a 2-hour group session, or multiple modes of information in a phone call.

Another important and distinctive factor that we believe contributed to the higher level of knowledge was the usage of (daily) push notifications—actively bringing information to patients, reminding them about the information in the app, and giving them the opportunity to directly access the information by clicking on the notification. In our study, the median number of times patients viewed the information in the week before the consultation was 25 (Q1-Q3: 12-39). Sending push notifications had a direct effect on usage of the app in terms of patients opening the app, viewing information, watching a video, or answering a quiz-like question.

Strengths and Limitations

To our knowledge, this is the first study that investigated patient education through an interactive app, while taking actual and perceived knowledge as well as satisfaction and confidence about treatment choice into account. Second, the study covers a relevant topic in modern health care: eHealth. The third strength is the design of the trial: multicenter (6 hospitals), randomized, controlled, and blinded for the treating physician. Finally, the content for the app was composed by using multiple sources of information, including orthopedic surgeons and current guidelines.

A limitation of our study is the fact that the level of actual knowledge was only measured once. Therefore, we could not perform a within-group comparison to assign the higher level of actual knowledge to the intervention. However, the randomization, the similarity of baseline characteristics between groups, and the significant increase in perceived knowledge only in the intervention group render it likely that the difference in actual knowledge between groups is because of the intervention.

Second, 22% of patients in the app group did not download the app, for unknown reasons. Perhaps, the instructions were too complex or the patients had trouble with the initial download from the App or Play Store. This demonstrates the necessity to offer support to patients for the initial usage of interventions such as these. Nevertheless, the level of adherence was high (70%). Moreover, even without correcting for this, our results show a clear advantage in the level of knowledge in the app group. This effect of the app became more pronounced when data were analyzed using the *per protocol* method. Third, we only included patients in possession of an email address and a smartphone or tablet. These criteria could limit the generalizability of the results. However, majority of the sample did use email (45-65 years: 92.9%; >65 years: 62.2%) and have a smart device (45-65 years: 92.4%; >65 years: 68.7%) [38], and this number will most probably only increase in the future. The fourth limitation is the use of self-created questionnaires to measure perceived and actual knowledge, as validated questionnaires covering our *top 5* important topics do not exist. To minimize this limitation, these questionnaires were developed in close cooperation with practicing clinicians and in line with Dutch guidelines and patient educational booklets. Finally, we did not consider individual preferences of patients regarding education on treatment options, nor their desire, or the lack of

it, to participate in SDM. Earlier research has shown that some older people feel that “the doctor knows best” and that “their own knowledge is superfluous in the decision-making process” [4].

Conclusions

Modern health care is more and more focused on patient involvement, in which knowledge about ones’ illness and treatment options is a must. Our results show that educating patients by using an interactive app with timely information could play an important role in increasing knowledge—one of the key drivers for SDM. It could benefit the consultation with the physician as well as result in long-term satisfaction with the treatment chosen.This study focused on patients being referred to an orthopedic surgeon by their GP with complaints indicating knee OA. We hypothesize, however, that neither the illness or the (phase of the) treatment limits the extent to which this type of intervention could be useful in improving patient education. Future research is needed to show the generalizability of using an app to actively offer patients subdivided and interactive information related to their specific illness or treatment. Furthermore, the (long-term) effects of this intervention on the final choice patients made regarding their treatment and the effect to which they were satisfied with this choice need to be demonstrated.

We found that, in comparison with standard educational tools, using an app to actively educate patients with subdivided, categorized, and interactive content significantly increased their level of perceived knowledge. Furthermore, a significantly higher level of actual knowledge was demonstrated in the intervention group. Patients in the app group were also more satisfied with the information they received and with their level of knowledge compared with the control group. Even though the intervention did not have an effect on their appreciation of the consultation with the doctor, patients in the app group were more confident and aware about their choice of treatment.

Acknowledgments

The authors would like to thank all the participating hospitals, their medical professionals, research coordinators, and research assistants for their contribution to the success of this study. Furthermore, they would like to thank the team of Interactive Studios for creating the app used as an intervention in this study. Finally, they would like to thank all patients who were willing to participate in the study. Interactive Studios offered the app used in this study free of charge.

Authors' Contributions

TT conceived the study and designed the trial. LJ, YP, BCvdZ, KLMK, KK, and WvdW supervised the data collection. TT managed the data. TT, LJ, BCvdZ, RK, and WvdW provided statistical analysis. SK, DvO, SdB, KK, SR, DD, and RCIvG provided the content used on the patients’ timeline. TT drafted the manuscript. All authors contributed to its revision.

Conflicts of Interest

The principal investigator, TT, is one of the cofounders of Interactive Studios. Interactive Studios is the company that developed the app used in this study. The coauthors declare that the research was conducted in the absence of any other commercial or financial relationships that could be construed as a potential conflict of interest. Moreover, all authors have completed the ICMJE uniform disclosure form and declare the following: no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years, and no other relationships or activities that could appear to have influenced the submitted work.

Editorial notice: This randomized study was only retrospectively registered, explained by authors with "technical problems during the initial, prospective registration". The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low and the study was considered formative. 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

Overview of the self-developed questionnaires used in this study.

[PDF File (Adobe PDF File), 36KB - mhealth_v6i12e10742_app1.pdf]

Multimedia Appendix 2

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 590KB - mhealth_v6i12e10742_app2.pdf]

References

1. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692. [Medline: 9032835]
2. Foster G, Taylor SJ, Eldridge SE, Ramsay J, Griffiths CJ. Self-management education programmes by lay leaders for people with chronic conditions. Cochrane Database Syst Rev 2007 Oct 17(4):CD005108. [doi: 10.1002/14651858.CD005108.pub2] [Medline: 17943839]
3. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Dec 12;4:CD001431. [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
4. Joseph-Williams N, Elwyn G, Edwards A. Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns 2014 Mar;94(3):291-309. [doi: 10.1016/j.pec.2013.10.031] [Medline: 24305642]
5. McGuire LC. Remembering what the doctor said: organization and adults' memory for medical information. Exp Aging Res 1996;22(4):403-428. [doi: 10.1080/03610739608254020] [Medline: 8968711]
6. Kessels RP. Patients' memory for medical information. J R Soc Med 2003 May;96(5):219-222 [FREE Full text] [Medline: 12724430]
7. Ekdahl AW, Andersson L, Friedrichsen M. "They do what they think is the best for me." Frail elderly patients' preferences for participation in their care during hospitalization. Patient Educ Couns 2010 Aug;80(2):233-240. [doi: 10.1016/j.pec.2009.10.026] [Medline: 19945814]
8. Henderson S. Influences on patient participation and decision-making in care. Prof Nurse 2002 May;17(9):521-525. [Medline: 12025012]
9. Kawabata H, Konishi K, Murakami M, Kisa K, Maezawa M. Factors affecting the physician-patient relationship regarding patient participation in medical encounters in primary care. Hokkaido Igaku Zasshi 2009 May;84(3):171-175. [Medline: 19507555]
10. Bastiaens H, Van Royen P, Pavlic DR, Raposo V, Baker R. Older people's preferences for involvement in their own care: a qualitative study in primary health care in 11 European countries. Patient Educ Couns 2007 Sep;68(1):33-42. [doi: 10.1016/j.pec.2007.03.025] [Medline: 17544239]
11. Jansen J, van Weert J, van der Meulen N, van Dulmen S, Heeren T, Bensing J. Recall in older cancer patients: measuring memory for medical information. Gerontologist 2008 Apr;48(2):149-157. [Medline: 18483427]
12. Marks JR, Schectman JM, Groninger H, Plews-Ogan ML. The association of health literacy and socio-demographic factors with medication knowledge. Patient Educ Couns 2010 Mar;78(3):372-376. [doi: 10.1016/j.pec.2009.06.017] [Medline: 19773144]
13. Ferguson M, Brandreth M, Brassington W, Wharrad H. Information retention and overload in first-time hearing aid users: an interactive multimedia educational solution. Am J Audiol 2015 Sep;24(3):329-332. [doi: 10.1044/2015_AJA-14-0088] [Medline: 26649541]
14. Latorre-Postigo JM, Ros-Segura L, Navarro-Bravo B, Ricarte-Trives JJ, Serrano-Selva JP, López-Torres-Hidalgo J. Older adults' memory for medical information, effect of number and mode of presentation: an experimental study. Patient Educ Couns 2017 Jan;100(1):160-166. [doi: 10.1016/j.pec.2016.08.001] [Medline: 27567498]
15. Selic P, Svab I, Repolusk M, Gucek NK. What factors affect patients' recall of general practitioners' advice? BMC Fam Pract 2011 Dec 28;12:141 [FREE Full text] [doi: 10.1186/1471-2296-12-141] [Medline: 22204743]
16. Say R, Murtagh M, Thomson R. Patients' preference for involvement in medical decision making: a narrative review. Patient Educ Couns 2006 Feb;60(2):102-114. [doi: 10.1016/j.pec.2005.02.003] [Medline: 16442453]
17. Tait AR, Voepel-Lewis T, Malviya S, Philipson SJ. Improving the readability and processability of a pediatric informed consent document: effects on parents' understanding. Arch Pediatr Adolesc Med 2005 Apr;159(4):347-352. [doi: 10.1001/archpedi.159.4.347] [Medline: 15809387]
18. Hess TM, Tate CS. Adult age differences in explanations and memory for behavioral information. Psychol Aging 1991 Mar;6(1):86-92. [Medline: 2029372]
19. Tait AR, Teig MK, Voepel-Lewis T. Informed consent for anesthesia: a review of practice and strategies for optimizing the consent process. Can J Anaesth 2014 Sep;61(9):832-842. [doi: 10.1007/s12630-014-0188-8] [Medline: 24898765]
20. 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]
21. 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]
22. 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]
23. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol 2015;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
24. Adsit KI. Multimedia in nursing and patient education. Orthop Nurs 1996;15(4):59-63. [Medline: 8868790]

25. Kahn G. Computer-based patient education: a progress report. MD Comput 1993;10(2):93-99. [Medline: 8469102]
26. Bental DS, Cawsey A, Jones R. Patient information systems that tailor to the individual. Patient Educ Couns 1999 Feb;36(2):171-180. [Medline: 10223021]
27. Liikavainio T, Lyytinen T, Tyrväinen E, Sipilä S, Arokoski JP. Physical function and properties of quadriceps femoris muscle in men with knee osteoarthritis. Arch Phys Med Rehabil 2008 Nov;89(11):2185-2194. [doi: 10.1016/j.apmr.2008.04.012] [Medline: 18996249]
28. Crawford DC, Miller LE, Block JE. Conservative management of symptomatic knee osteoarthritis: a flawed strategy? Orthop Rev (Pavia) 2013 Feb 22;5(1):e2 [FREE Full text] [doi: 10.4081/or.2013.e2] [Medline: 23705060]
29. Patientenfederatie Nederland (Dutch Patient Federation), Nederlandse Orthopedische Vereniging (Dutch Orthopedic Federation), Kennisinstituut Federatie Medisch Specialisten (Knowledge center federation medical specialist). Consult Kaart. ARTROSE IN DE KNIE: behandelmogelijkheden URL: http://consultkaart.nl/wp-content/uploads/2018/01/FMS_ck_artrose-knie_2016.02.pdf [accessed 2018-03-26] [WebCite Cache ID 6yCbkBryp]
30. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults. J Appl Gerontol 2018 Feb;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
31. Sepucha K, Atlas SJ, Chang Y, Dorrwachter J, Freiberg A, Mangla M, et al. atient decision aids improve decision quality and patient experience and reduce surgical rates in routine orthopaedic care: a prospective cohort study. J Bone Joint Surg Am 2017 Aug 02;99(15):1253-1260. [doi: 10.2106/JBJS.16.01045] [Medline: 28763411]
32. 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]
33. Henderson S. Power imbalance between nurses and patients: a potential inhibitor of partnership in care. J Clin Nurs 2003 Jul;12(4):501-508. [Medline: 12790863]
34. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010 Jan;468(1):57-63 [FREE Full text] [doi: 10.1007/s11999-009-1119-9] [Medline: 19844772]
35. Dunbar MJ, Richardson G, Robertsson O. I can't get no satisfaction after my total knee replacement: rhymes and reasons. Bone Joint J 2013 Nov;95-B(11 Suppl A):148-152. [doi: 10.1302/0301-620X.95B11.32767] [Medline: 24187375]
36. Mancuso CA, Graziano S, Briskie LM, Peterson MG, Pellicci PM, Salvati EA, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. Clin Orthop Relat Res 2008 Feb;466(2):424-431 [FREE Full text] [doi: 10.1007/s11999-007-0052-z] [Medline: 18196427]
37. Murphy KR. Computer-based patient education. Otolaryngol Clin North Am 1998 Apr;31(2):309-317. [Medline: 9518439]
38. Centraal Bureau voor de Statistiek (CBS). StatLine. Internet; access, use and facilities URL: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=83429ned&D1=0-69&D2=0,3-6&D3=0&D4=a&VW=T [WebCite Cache ID 6yBnUJla9]

Abbreviations

eHealth: electronic health
GP: general practitioner
NRS: Numeric Rating Scale
OA: osteoarthritis
SDM: shared decision making

Edited by G Eysenbach; submitted 13.04.18; peer-reviewed by D Nault; comments to author 09.10.18; revised version received 15.10.18; accepted 17.10.18; published 21.12.18.

Please cite as:
Timmers T, Janssen L, Pronk Y, van der Zwaard BC, Koëter S, van Oostveen D, de Boer S, Kremers K, Rutten S, Das D, van Geenen RCI, Koenraadt KLM, Kusters R, van der Weegen W
Assessing the Efficacy of an Educational Smartphone or Tablet App With Subdivided and Interactive Content to Increase Patients' Medical Knowledge: Randomized Controlled Trial
JMIR Mhealth Uhealth 2018;6(12):e10742
URL: http://mhealth.jmir.org/2018/12/e10742/
doi:10.2196/10742
PMID:30578185

Original Paper

An Analytics Platform to Evaluate Effective Engagement With Pediatric Mobile Health Apps: Design, Development, and Formative Evaluation

Quynh Pham[1,2], MSc; Gary Graham[2], MSc; Chitra Lalloo[1,3], PhD; Plinio P Morita[1,4], PEng, PhD; Emily Seto[1,2], PEng, PhD; Jennifer N Stinson[1,3,5,6], RN, PhD, CPNP; Joseph A Cafazzo[1,2,7], PEng, PhD

[1]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[2]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada
[3]Child Health Evaluative Sciences Research Institute, The Hospital for Sick Children, Toronto, ON, Canada
[4]School of Public Health and Health Systems, Faculty of Applied Health Sciences, University of Waterloo, Waterloo, ON, Canada
[5]Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, ON, Canada
[6]Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
[7]Institute of Biomaterials and Biomedical Engineering, Faculty of Applied Science and Engineering, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Quynh Pham, MSc
Institute of Health Policy, Management and Evaluation
Dalla Lana School of Public Health
University of Toronto
Health Sciences Building, Suite 425
155 College Street
Toronto, ON, M5T 3M6
Canada
Phone: 1 (416) 340 4800 ext 4765
Email: q.pham@mail.utoronto.ca

Abstract

Background: Mobile health (mHealth) apps for pediatric chronic conditions are growing in availability and challenge investigators to conduct rigorous evaluations that keep pace with mHealth innovation. Traditional research methods are poorly suited to operationalize the agile, iterative trials required to evidence and optimize these digitally mediated interventions.

Objective: We sought to contribute a resource to support the quantification, analysis, and visualization of analytic indicators of effective engagement with mHealth apps for chronic conditions.

Methods: We applied user-centered design methods to design and develop an Analytics Platform to Evaluate Effective Engagement (APEEE) with consumer mHealth apps for chronic conditions and implemented the platform to analyze both retrospective and prospective data generated from a smartphone-based pain self-management app called *iCanCope* for young people with chronic pain.

Results: Through APEEE, we were able to automate the process of defining, operationalizing, and evaluating effective engagement with *iCanCope*. Configuring the platform to integrate with the app was feasible and provided investigators with a resource to consolidate, analyze, and visualize engagement data generated by participants in real time. Preliminary efforts to evaluate APEEE showed that investigators perceived the platform to be an acceptable evaluative resource and were satisfied with its design, functionality, and performance. Investigators saw potential in APEEE to accelerate and augment evidence generation and expressed enthusiasm for adopting the platform to support their evaluative practice once fully implemented.

Conclusions: Dynamic, real-time analytic platforms may provide investigators with a powerful means to characterize the breadth and depth of mHealth app engagement required to achieve intended health outcomes. Successful implementation of APEEE into evaluative practice may contribute to the realization of effective and evidence-based mHealth care.

(JMIR Mhealth Uhealth 2018;6(12):e11447) doi:10.2196/11447

KEYWORDS

analytics; engagement; log data; mobile health; mobile apps; chronic disease

Introduction

Background

The emergence of consumer mobile health (mHealth) apps for chronic disease self-management presents new opportunities and challenges for evidencing these novel interventions. Most consumer mHealth apps have not been evaluated for effectiveness on health outcomes [1]. This trend is particularly evident within the field of pediatrics, where recent reviews have revealed a paucity of evidence-based apps for young people across chronic conditions [2-4]. In spite of this, apps for pediatric chronic conditions are growing in availability [4] and challenge investigators to conduct rigorous evaluations that keep pace with mHealth innovations. Traditional research methods are poorly suited to operationalize the agile, iterative trials required to evidence and optimize these digitally mediated interventions [5,6].

In recent years, digital health researchers have called for novel methods to study engagement with digital health interventions [5]. They propose that engagement with an intervention is a precondition for effectiveness and warrants careful study to understand its relationship with the desired behavior change (eg, pain self-management) [7]. Yardley et al have furthered this focus on evaluating engagement by arguing that it may be more valuable to identify the mechanisms that underlie *effective engagement*, defined as "sufficient engagement with an intervention to achieve intended outcomes" [8]. They recommend the following 6 distinct methods to assess different aspects of effective engagement: (1) self-report interviews or observational sessions, (2) self-report questionnaires, (3) ecological momentary assessments, (4) system usage logs, (5) sensor data, and (6) psychophysiological measures.

On reviewing these methods, we noted that the majority can be delivered or collected from data generated by users directly engaging with a digital health intervention. These multilevel, temporally dense datasets may be sufficiently large to reliably model and experimentally test mediation of outcomes by engagement with particular intervention features and functionality, while statistically controlling for confounding moderator effects, such as baseline pain levels [9]. However, these data can also be complex [10,11], making it difficult to discern signal from noise [12,13]. Investigators may struggle to efficiently distill thousands of data points into meaningful insights that relate digitally mediated engagement with changes in health outcomes [14]. Realizing a method to cull through large mHealth app datasets and identify meaningful patterns of digitally mediated behavior change may promote a data-driven understanding of their impact on disease self-management.

Objectives

Motivated by this understanding of the barriers to interpreting data from measures of effective engagement, we sought to contribute a resource to support the quantification, analysis, and visualization of analytic indicators of effective engagement with mHealth apps for chronic conditions. Specifically, we designed and developed an Analytics Platform to Evaluate Effective Engagement (APEEE) with consumer mHealth apps for chronic conditions and implemented the platform to analyze both retrospective and prospective data generated from a smartphone-based pain self-management app called *iCanCope* for young people with chronic pain [15]. Our intent was for APEEE to broadly enable investigators to query data being generated by users engaging with their mHealth apps in real time and specifically support the identification of mediating mechanisms that motivate effective engagement. This research assessed the feasibility of configuring APEEE for use in a pediatric research environment and its preliminary acceptability by mHealth investigators to inform evaluative practice. Specifically, (1) can the process of defining, operationalizing, and evaluating effective engagement with *iCanCope* be automated through APEEE? and (2) what are investigators' perceptions regarding acceptability and satisfaction with APEEE?

This paper is organized as follows: first, we present the user-centered design (UCD) framework used to support the design and development of APEEE and review the features and functionality of the minimum viable product build of the platform. Second, we define the analytic indicators of effective engagement with *iCanCope* for inclusion in APEEE. Third, we review the technical and architectural considerations for modeling *iCanCope* engagement data and representing it on the platform. Finally, we describe the prototypical integration of APEEE with *iCanCope* to support a pilot randomized controlled trial (RCT) evidencing the intervention for young people with chronic pain.

Methods

The iCanCope App for Young People With Chronic Pain

Before initiating work on APEEE, we chose to identify a typical mHealth app to define our scope of work. Our rationale for establishing a single use case to guide the platform's features and functionality was as follows: we wanted (1) a testing environment to experiment with various data integration and visualization methods, (2) to refine the platform's computational capacity for modeling and managing dynamic data, (3) to validate data generated by the platform against data already being generated as part of an ongoing evaluation (eg, number of users, number of log-ins, and session duration), and (4) a direct route to implementation following development to trial our platform in evaluative practice. To meet these needs, we selected *iCanCope*, a smartphone-based pain self-management mHealth app tailored for adolescents and young adults aged 12 to 25 years with chronic pain [15].

The *iCanCope* project was conceived by the Improving Outcomes in Child Health through Technology (iOuch) research group, based out of the Hospital for Sick Children in Toronto, Canada [16]. The iOuch research group aims to improve the

lives of children and adolescents through the use of innovative information and communication technologies. Research personnel includes a principal investigator, a research associate, 2 clinical research managers, 2 clinical research coordinators, 5 clinical research assistants, and a rotating roster of 5 to 7 research students and fellows. The group conducts research to conceptualize, design, and evidence digital health interventions such as *iCanCope* and outsources the development of the interventions to external research groups or software development studios. Moreover, 5 members of the research group are dedicated staff on the *iCanCope* project.

iCanCope was an appropriate match to inform our work because the app was already collecting data from participants in a pilot RCT to evaluate its preliminary efficacy on improving pain outcomes. Furthermore, our research group is the development partner on the *iCanCope* project, thereby ensuring ethical and direct access to both app data and the core research group evaluating the app. The main *iCanCope* features are (1) symptom tracking for pain intensity, pain interference, sleep, mood, energy, and physical activity in the form of daily *check-in* reports, (2) structured goal setting to improve pain and function, (3) an interactive toolbox of pain coping strategies, and (4) peer-based social support [15]. The app was developed natively for iOS and Android smartphone platforms. It was deployed in March 2017 for evaluation in a pilot RCT and had generated a significant amount of data before integration with APEEE in April 2018. We wish to note that although *iCanCope* features heavily in the conceptual narrative of APEEE, this research focuses on the platform as a proof-of-concept resource for pediatric mHealth app evaluation and as such does not constitute a study of *iCanCope* as an intervention for pediatric chronic pain.

The User-Centered Design and Development of Analytics Platform to Evaluate Effective Engagement

The design and development of APEEE were guided by the UCD framework, which has been endorsed by the World Health Organization as a systematic approach to considering the needs of end users throughout all stages of the design life cycle [17]. As a design philosophy, the UCD framework endorses creating technology that users can, want, or need to use, rather than forcing users to change their behavior to accommodate the technology [18,19]. Starting with the *concept generation and ideation* processes in phase 1, user needs are identified to inform the intended goal of the digital health intervention. In phase 2, the *prototype design and system development* process is initiated, whereby identified user needs are translated into a set of functional requirements and design guidelines. Prototypes are created using these guidelines and refined through cycles of iterative design, often with real-time feedback elicited from end users. Phase 3 is the *evaluation* component of the process and ensures that the application can be implemented effectively in practice. Once these 3 phases are completed, the application is deployed to users. We initiated phase 1 of the UCD process in March 2018, progressed to phase 2 in May 2018, commenced phase 3 in June 2018, and advanced to a field study of APEEE in October 2018.

Results

Concept Generation and Ideation

We initiated phase 1 of the UCD framework by conducting a needs assessment session with 5 members of the iOuch research group to inform a baseline understanding of (1) their experience with the evaluation process, (2) their perception of the barriers and facilitators to evaluating the intervention, and (3) their definition of what constitutes effective engagement with the intervention. Investigators were prompted to speak about their specific evaluation questions, what measures were used to answer the evaluation questions, what data were required to operationalize those measures, and how that data had been collected. In parallel, we conducted a scoping review to identify and validate the terminology, definitions, and taxonomy of analytic indicators being used to measure effective engagement with mHealth apps for chronic conditions. Preliminary findings from the review informed the creation of a library of analytic indicators, which we referenced to define a shortlist of analytic indicators specific to *iCanCope*. Finally, we reviewed the existing *iCanCope* system architecture and data model to assess the feasibility of implementing the proposed shortlist of analytic indicators. We presented our recommendations to the iOuch research group for review and collaboratively finalized a list of 25 analytic indicators to represent on APEEE. Table 1 presents all analytic indicators, each expressed as a research question, and their corresponding definition.

Prototype Design and System Development

To execute phase 2, we determined the design and development specifications required to represent each analytic indicator on APEEE. These specifications subsequently guided the selection of products to build out the platform as well as platform features and functionality. APEEE was developed using a collection of 3 open-source products: *Logstash*, *Elasticsearch*, and *Kibana* [20]. *Logstash* is a server-side data processing pipeline that ingests data from various sources simultaneously, executes different transformations, and exports the data to various targets. Given that data can be siloed across systems in different formats, Logstash supports data from logs, metrics, Web apps, data stores, and cloud computing services. As data travel from source to store, Logstash filters parse each event, identify named fields to build structure, and transform them to converge on a common format for analysis. *Elasticsearch* is a search engine based on the Lucene information retrieval software library. It provides a distributed, multitenant-capable, full-text search engine with a Web interface and schema-free JavaScript Object Notation documents. Elasticsearch allows users to perform and combine many types of searches, such as structured, unstructured, geographical, and metric. *Kibana* is an analytics and visualization plugin to Elasticsearch. Users can interface with Kibana to search, view, and interact with data stored in Elasticsearch indices. They can also perform advanced time-series analyses and visualize data in a range of charts, tables, and maps. Kibana facilitates the analysis of large volumes of data and also enables the creation of dynamic dashboards that display data queries in real time.

Table 1. Analytic indicators of effective engagement with *iCanCope*.

Analytic indicator	Definition
Health status	
How are users doing on pain-related outcomes?	Raw and mean pain intensity, pain interference, sleep, mood, energy, and physical activity check-in scores generated over time
Are users recording positive or negative check-in trends?	Number of positive and negative trends triggered
Which pain-related outcome scores are users reporting the most?	Number of scores reported per check-in score response
Which pain-related outcome scores are most users reporting?	Number of users reporting scores per check-in score response
Check- ins	
How many check-ins are being completed daily?	Number of check-ins completed every day
How many check-ins have been completed since study launch?	Number of check-ins completed in the last 90 days
How many users have completed at least one check-in a day, every day, over the last 7 days?	Number of users with ≥1 check-in completed in a day, every day
How many users have not completed a check-in for 7 consecutive days?	No check-ins logged for 7 consecutive days
How long did it take for users to complete their first check-in?	Time between account creation and first check-in completed
Which 10 users have completed the most check-ins?	Identity of user and number of check-ins completed
How many check-ins were completed this week versus last week?	Number of check-ins completed this week and number of check-ins completed last week
Goals	
Are users completing set goals?	Number of goals set and completed
What types of goals are users setting the most?	Number of activity, sleep, energy, mood, and social goals set
What types of goals are most users setting?	Number of users setting activity, sleep, energy, mood, and social goals
How long did it take for users to complete their first goal?	Time between account creation and first goal created
Community	
How many users have engaged with the community features?	Number of users who liked or made a post on the community feature
What were the top 5 community questions with the most responses?	Content of community questions and number of responses
Library	
What are the top 10 most popular library articles?	Content of library articles and number of reads
History	
How many users accessed the history feature at least once?	Number of users who clicked on the history feature
What symptoms are users reviewing in the history feature?	Contents and number of history pages clicked
Other	
How many users have activated an *iCanCope* account?	Number of users registered on the study server
How many users have logged any activity in the last 7 days?	Number of users who generated ≥1 event on the study server in the last 7 days
How many users have logged any activity in the last 24 hours?	Number of users who generated ≥1 event on the study server in the last 24 hours
Where in the world are users accessing the app?	Geolocation of user internet protocol addresses
How far have users progressed in the study?	Numbers of days elapsed since account creation

Dashboards can be shared with a broader sample of users through URL weblinks, and analytic reports can be exported to comma-separated value (CSV) and PDF formats. To summarize, Logstash collects and parses log data, Elasticsearch indexes and stores the data, and Kibana visualizes the data to provide actionable insights. Together, these 3 open-source products are designed for use as an integrated solution, commonly referred to as the *Elastic Stack*.

This research primarily focused on configuring the Elastic Stack to conform with APEEE product specifications. Our decision to forego writing proprietary code in favor of adopting an open-source solution was dually motivated: (1) we wanted to leverage the Elastic Stack features, functionality, and

documentation built by a community of 100,000 developers over 6 years [21], and (2) we are proponents of the open-source software development methodology [22]. Using Elastic Stack capabilities, we developed a prototype of APEEE that enabled investigators to (1) visualize a library of effective engagement analytic indicators extracted from *iCanCope* data; (2) build filters to segment the study population into cohorts for comparative analyses; (3) monitor the status of informed electronic consent, study progression, and fidelity of intended engagement by young people with chronic pain; (4) conduct basic statistical analyses on a dynamic engagement dataset; and (5) generate individual- and aggregate-level analytic insights in real time. The principal feature of APEEE is the *APEEE Dashboard*, which is an interface for investigators to view analytic indicator trends for immediate research-to-action application (eg, inform the need to modify an intervention feature because of poor engagement). To support platform functionality, we built a Personal Health Information Protection Act–compliant *APEEE Engine* to serve as the foundational information management and data infrastructure required to integrate and store engagement data and support data mining and export for advanced statistical analyses.

For APEEE to produce meaningful insights on engagement with *iCanCope*, we first needed to aggregate and store events generated from both client devices and the application server. To realize this work, we used an *event streaming* architecture, which lends itself well to analyzing temporally dense data. Furthermore, it provides us with the ability to see data as they were at any given point in time; this property is useful for conducting time-series analyses. At the most basic level, a unit of data is an *event* and contains 2 pieces of information: identification and payload. The former is used for aggregatory purposes and cohort tracking, and the latter is the actual event that occurred, along with any useful metadata. To illustrate the passage of our data from *iCanCope* to APEEE, we will use a simple event in which a participant signs into the app on their device. First, an event called *user_signed_in* is generated by the device, along with a time stamp and user identifier. This event is sent to our application server, which stores it to a local database and then emits it to a local log file. However, these data cannot be analyzed on the application server and must be forwarded to a destination that can effectively process it. There is a lightweight daemon (ie, a computer program that runs as a background process) called *Filebeat* running on the application server, which detects any new events in log files and forwards them over a private network to our instance of the Elastic Stack. On arrival at the Elastic Stack, the event is ingested via Logstash, and metadata are pulled out from the event to increase the overall amount of data we can later query and visualize. From there, the event is tagged as an analytic event and is sent to Elasticsearch for indexing and storage. Elasticsearch not only performs minor analysis on the incoming event but also provides it with durability and ease of lookup by duplicating it over replica shards. Once the data arrive at Elasticsearch, it can now be queried in Kibana, which supports our various visualizations and dashboards. Figure 1 relates this use case as a system architecture diagram of APEEE.

Following platform configuration, we initiated the process of translating all analytic indicators into visualizations on the APEEE dashboard. This process involved selecting the appropriate graphic for each indicator (eg, line graph, pie chart, heat map, and data table), defining the appropriate data fields and parameters, and adjusting graphic assets (eg, axis values, table headers, and color schema) to represent the indicator as a dynamic visualization. We sought feedback from our internal team of human factors specialists and designers to ensure appropriate alignment between data and visualization. We also engaged in a rapid-cycle iterative prototyping process with the iOuch research group, where we sent over dashboard prototypes for review on a near-daily basis. This constant communication and collaboration with our end users allowed us to recalibrate our prototypes with emerging needs, which led to timely adjustments and improvements to the dashboard. Figure 2 presents the APEEE dashboard with a subset of finalized analytic indicators. To provide a comprehensive and instructive description of platform features and functionality, we have chosen to present 3 indicators in detail: (1) where in the world are users accessing *iCanCope*?, (2) what types of goals are users setting?, and (3) how many check-ins are being completed daily?

Where in the World Are Users Accessing iCanCope?

As a participant in the *iCanCope* pilot RCT, young people aged 12 to 25 years with chronic pain were instructed to download the app onto their personal device, create an account, and use the app as needed over the 8-week study period. Given that participants were free to engage with the app wherever they wanted, they subsequently generated an internet protocol (IP) address trail that we were able to access through analyzing their device log data. The opportunity to determine and map the physical location of a user's IP address in real time was a novel challenge for our research group and a feature we wanted to trial in APEEE. We used Logstash's *GeoIP* database to convert IP addresses into latitude and longitude coordinate pairs, which were then stored in Elasticsearch as *geo_point* fields and converted into *geohash* strings. Kibana was then used to read the *geohash* strings and draw them as points on a global map.

Figure 3 presents this analytic indicator, visualized through APEEE as a choropleth map covering 5 continents, with scaled circle markers representing the number of unique IP addresses logged by participants accessing *iCanCope*. Higher intensity colored circles indicate a greater concentration of addresses in a particular region. Investigators can click on a specific region of interest to view a narrowed spread of addresses; Figure 4 presents the view generated from repeatedly clicking on the large red circle in Figure 3 to plot participants across the greater Toronto area in Ontario, Canada. This geographical insight allows the iOuch research group to (1) validate whether participants are accessing the intervention in the community, (2) measure the geographical scale and spread of intervention access, and (3) observe shifts in access and engagement patterns over time.

Figure 1. APEEE system architecture. APEEE: Analytics Platform to Evaluate Effective Engagement.

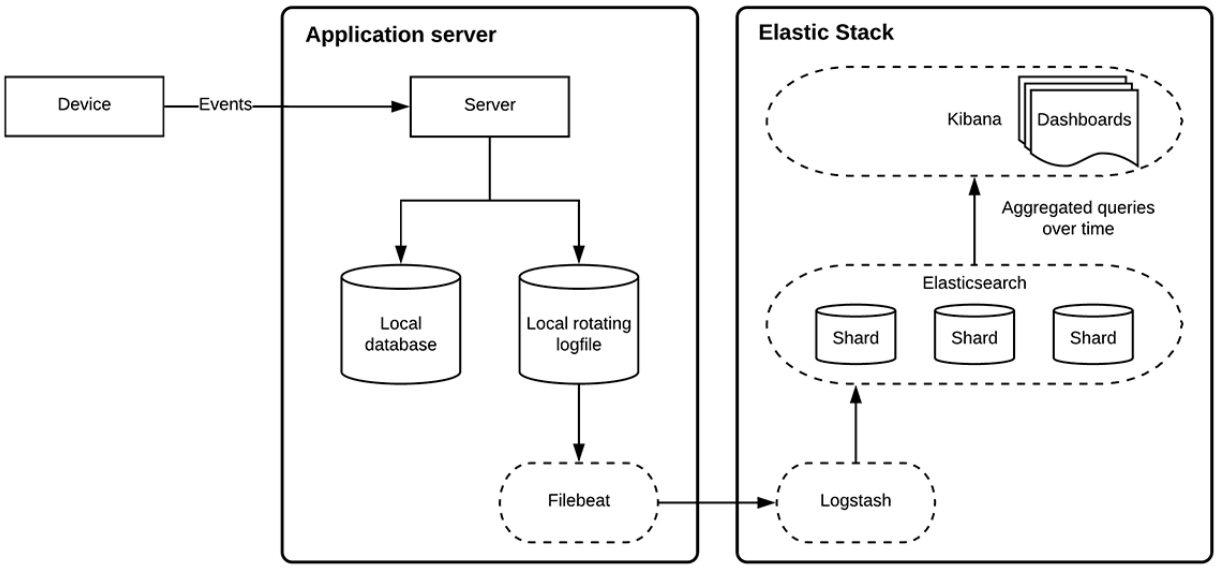


Figure 2. APEEE dashboard with a subset of analytic indicators of effective engagement with iCanCope. APEEE: Analytics Platform to Evaluate Effective Engagement.

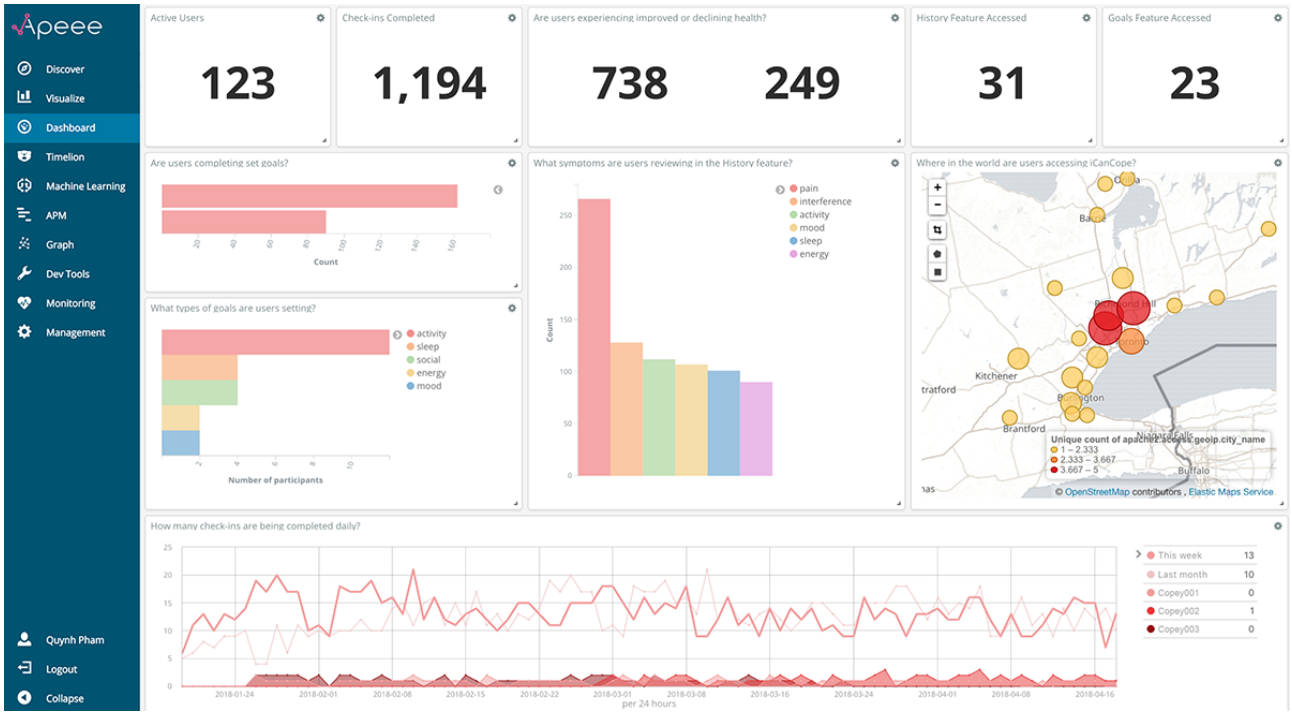


Figure 3. The analytic indicator for "where in the world are users accessing iCanCope?," visualized through APEEE as a choropleth map covering 5 continents. APEEE: Analytics Platform to Evaluate Effective Engagement.

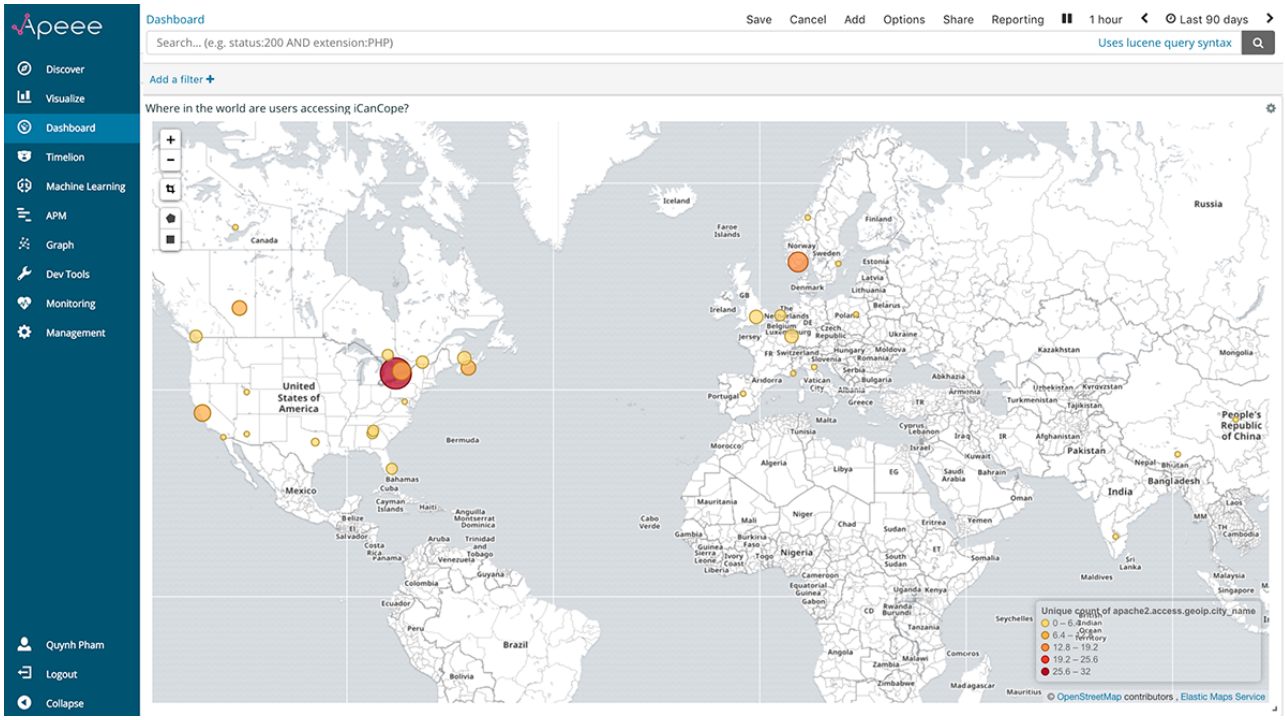
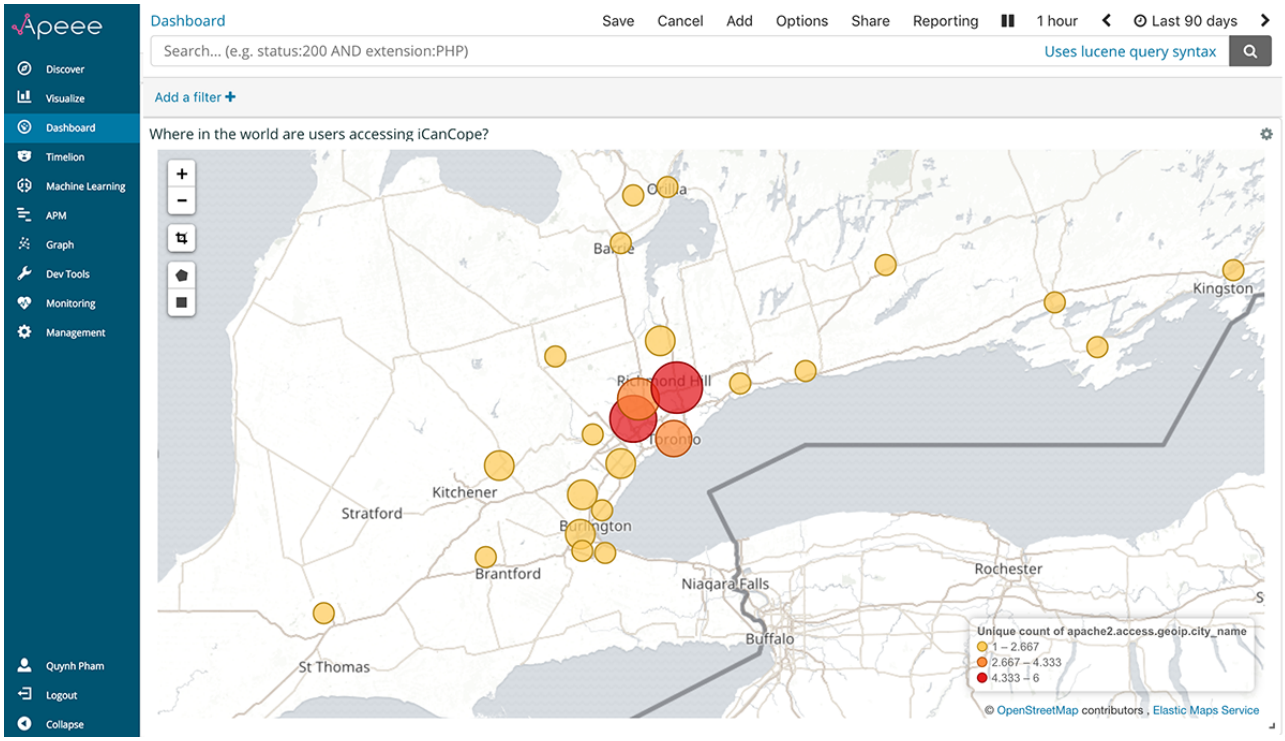


Figure 4. The analytic indicator for "where in the world are users accessing iCanCope?," visualized through APEEE as a choropleth map covering the greater Toronto area in Ontario, Canada. APEEE: Analytics Platform to Evaluate Effective Engagement.



What Types of Goals Are Users Setting?

A core component of *iCanCope* is the Goals feature, where young people are guided in setting structured goals aimed at improving their pain and function. Goals can be categorized across 5 domains: sleep, mood, energy, physical activity, and social activity. Given the hypothesized importance of this feature in promoting positive behavior change, we wanted to explore what types of goals participants were setting to understand what aspects of their behavior were amenable to improvement. We used Elasticsearch's aggregations framework to build a summary of all goals set by participants throughout the trial. An aggregation can be considered a unit of work that builds analytic information over a set of data. For this work, we specifically applied 6 *bucketing* aggregations to our full set of study data: 1 parent aggregation for all goals completed and 5 nested subaggregations for each goal domain. When bucketing

aggregations are executed in Elasticsearch, criteria for each bucket are evaluated against all data in a given set; if a criterion matches, the data *fall into* the relevant bucket.

Figure 5 presents this analytic indicator, visualized through APEEE as a horizontal bar chart, with the y-axis representing goal domains and the x-axis representing the number of goals set. A color-coded legend on the right side of the chart identifies the domain for each bar. This graph indicates that participants are setting more physical activity goals than other goal types as a group. However, to ensure that findings were not being skewed by a small number of users setting a large number of physical activity goals, we accessed Kibana settings and changed the x-axis to represent the number of participants setting goals for each domain. This axis change and the consequent graph generated (Figure 6) were implemented in under a minute and allowed us to instantly corroborate both user-level and event-level insights. With this knowledge, investigators might, for example, design more physical activity goals for participants to browse and set.

How Many Check-Ins Are Being Completed Daily?

As part of the *iCanCope* trial, participants were asked to adhere to a symptom tracking protocol, aimed at helping them to recognize and understand patterns in their pain and functioning, and better communicate their symptoms with health care providers. This protocol was delivered through the check-in feature of the app, which prompted participants to complete a check-in a day for 56 consecutive days (ie, the duration of the trial). Participants tracked symptoms across 6 domains: pain intensity, pain interference, sleep, mood, energy, and physical activity. At the time of app integration with APEEE, more than 50 participants were enrolled in the trial and had collectively logged more than 3000 data points across all symptom domains. This temporally dense ecological momentary assessment dataset enabled us to develop time-series data parsing, analysis, and visualization functionality into APEEE. To realize this feature, we implemented Elastic's aggregations framework and applied 2 bucketing aggregations to our data: (1) aggregating all check-ins logged by participants since study launch and (2) aggregating the number of daily check-ins over time. We then applied Kibana's *time series visual builder* filter over our data to visualize insights.

Figure 7 presents this analytic indicator, visualized through APEEE as a histogram with 3 layered graphs. The y-axis represents the total number of check-ins completed, and the x-axis represents time; the selected time range is 90 days. The first bolded line graph denotes the total number of check-ins completed per day. The second thin line graph also denotes the total number of check-ins completed per day but offset by 4 weeks. The third vertical bar chart with 3 superimposed bars denotes the 3 participants who have logged the most check-ins over the selected time range; participants were identified through a real-time count of check-ins conducted on the back end of the platform. Participants' usernames are presented in the legend but have been changed to maintain confidentiality. This layering of analytic insights might allow investigators to understand, for example, if there is a widening gap between daily check-in counts this week versus 4 weeks ago or the extent of check-in contribution from highly engaged participants.

In summary, these 3 functional use cases serve to illustrate the potential for APEEE to support investigators in their evaluative practice. We aim for the real-time analysis and visualization of analytic indicators through APEEE to provide investigators with timely and meaningful insights, which can then be further investigated outside of the platform using qualitative measures of effective engagement [8].

Evaluation

To operationalize phase 3 of the UCD framework, we conducted (1) 2 iterative cycles of evaluation on APEEE; the first with 2 members of the iOuch research group and the second with 7 members and (2) a between-cycle round of design and development. The first evaluation cycle was intended to assess the usability and acceptability of the platform and identify critical design and development requirements to be addressed and validated in the second evaluation cycle.

We first conducted a 1-day on-site observation session with 2 members of the iOuch research group to evaluate their initial use of APEEE. We were unable to provide investigators with direct access to the platform from their own devices because of the ongoing work at the time on the APEEE Engine. Instead, 1 member of our research group installed an instance of the platform onto a laptop, traveled to the evaluation site, connected to the APEEE Engine through a virtual private network, and launched the platform for use by the research group. Investigators were first provided with an overview of platform features and functionality and then presented the APEEE dashboard with visualizations for all *iCanCope* analytic indicators. They were then encouraged to explore each visualization and *think aloud* about the data representation and design specifications. Investigators were also asked to work independently through the following tasks while simultaneously verbalizing any difficulties encountered: (1) filtering visualizations by time range, (2) expanding a visualization to see more granular data points, (3) rearranging visualizations on the dashboard, (4) sorting numerical and string data table visualizations, and (5) exporting data table visualizations for download as CSV files. Field notes were taken during the session to record any technical difficulties encountered, ease of use, and learnings as well as nonverbal behaviors related to acceptability. Suggestions made by investigators on platform features or functionality that were not identified during the *concept generation and ideation* phase were considered for incorporation into the platform.

Figure 5. The analytic indicator for "what types of goals are users setting the most?," visualized through APEEE as a horizontal bar chart. APEEE: Analytics Platform to Evaluate Effective Engagement.

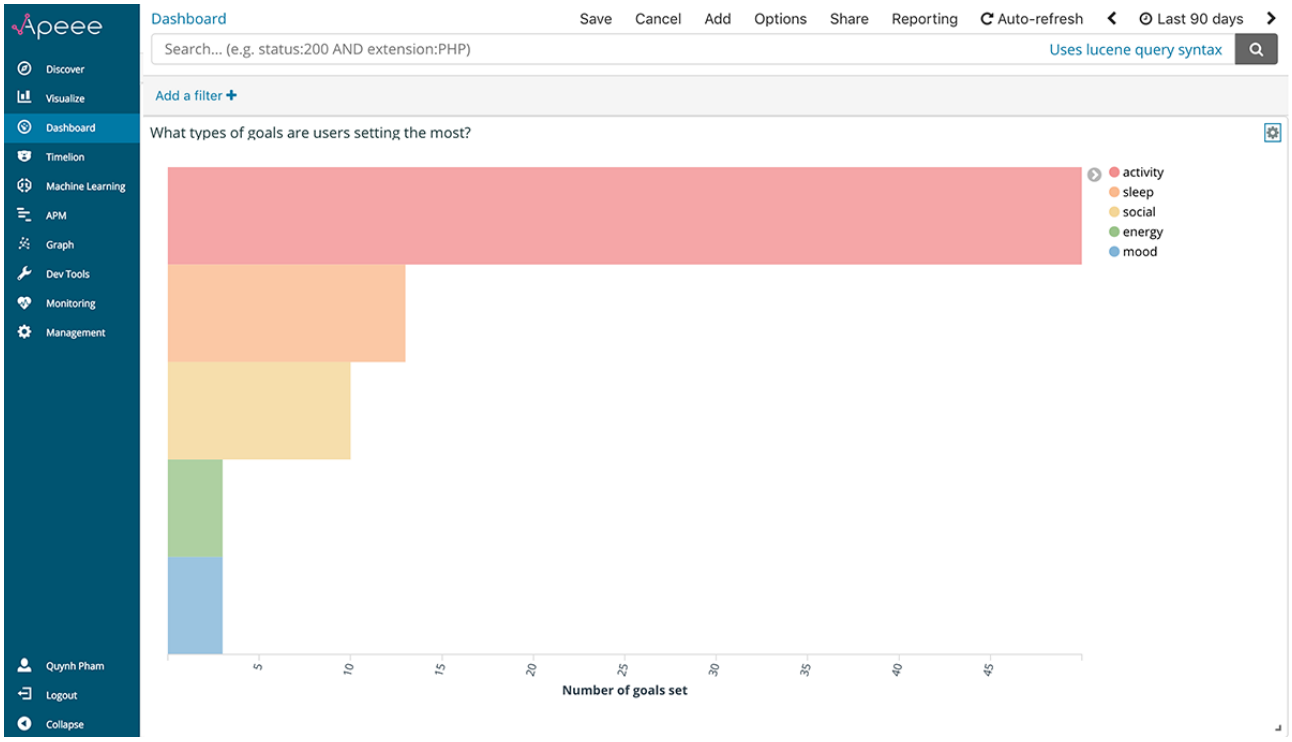
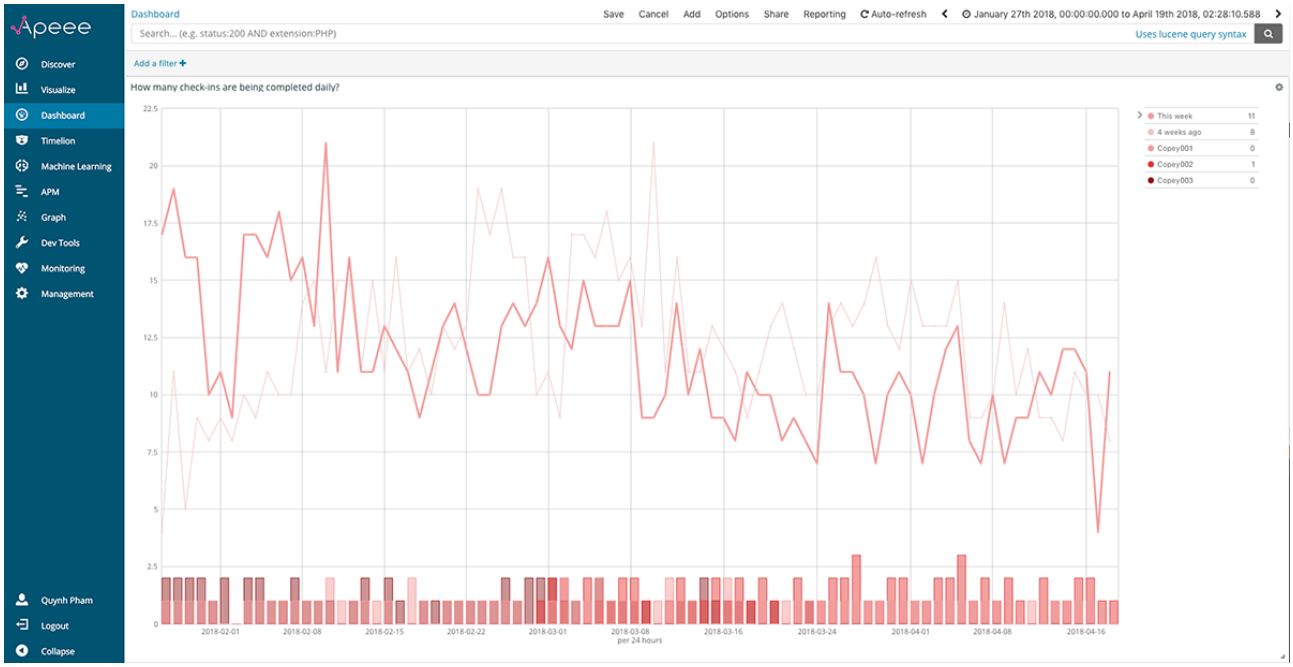


Figure 6. The analytic indicator for "what types of goals are most users setting?," visualized through APEEE as a horizontal bar chart. APEEE: Analytics Platform to Evaluate Effective Engagement.

Figure 7. The analytic indicator for "how many check-ins are being completed daily?," visualized through APEEE as a histogram with 3 layered graphs. APEEE: Analytics Platform to Evaluate Effective Engagement.



Overall, investigators found APEEE to be an acceptable resource to support their evaluation of *iCanCope*. They were able to independently work through all tasks with minimal guidance and sought clarification out of curiosity rather than necessity. There were no software bugs detected or system error codes returned during the session. The platform was explored with relative ease; however, some minor difficulties encountered included (1) inexperience with platform navigation, resulting in redundant actions to perform a task; (2) confusion regarding variable names, which retained server nomenclature and were sometimes difficult to interpret (eg, *clientCreated* to represent "participant"); and (3) unfamiliarity with performing Boolean searches using the Lucene query syntax, which is the default search syntax in APEEE. To alleviate these issues, we encouraged investigators to practice navigating the platform interface and repeating tasks until they felt intuitive and also provided them with a copy of the *iCanCope* data dictionary and a link to the Lucene query syntax as reference documentation. Investigators were satisfied with these additional resources and were able to complete tasks independent of them by session end. They saw potential in APEEE to accelerate and augment evidence generation both during and after trial conduct and expressed enthusiasm for adopting the platform as part of their evaluative practice. Suggestions to improve platform features and functionality included (1) partitioning the main dashboard into multiple subdashboards, each relating to a different feature in *iCanCope*; (2) supporting visualizations of events over *relative* time (eg, number of users who completed a check-in as a function of time elapsed in the study); (3) computing advanced predictive statistical analyses (eg, linear regressions); and (4) enabling remote access to APEEE. These requirements were feasible in scope and served as motivation to further develop the platform before full deployment.

Partition Dashboard and Enable Remote Access

Following this observation session, we initiated a new round of iterative design and development informed by the identified requirements. We were able to apply Kibana functionality and partition the main dashboard into multiple subdashboards. We also leveraged this functionality to build out custom dashboards for 5 members of the iOuch research group. Screenshots of these dashboards were presented to their intended users for review and found to be more useful than a single generic dashboard. To enable access to these dashboards for further testing and also address the remote access requirement, we activated Elasticsearch's *Security* module and configured the *Authorization* functionality. Authorization in APEEE is the process of determining whether the user behind an incoming request is allowed to execute it. APEEE manages the privileges of users through *roles*. A role has a unique name and identifies a set of permissions that translate to privileges on secured resource. For example, we defined the *iCanCope research analyst* role on APEEE to have *read* privileges on all documents that match the query *action: checkin_completed*. This role is limited to only viewing check-in data, as opposed to the *iCanCope research coordinator* role that has to *manage* privileges on the *iCanCope* cluster and can view, edit, and delete all documents.

Once we had defined a series of roles that aligned with the management structure of the iOuch research group, we assigned them to the 5 users for whom custom dashboards had been built out. We added username and password functionality for all user accounts and then sent each user a Secure Sockets Layer encrypted link to their custom APEEE dashboard for testing. All 5 users were able to remotely access APEEE, log into the platform, and view their custom dashboard. We asked users to remotely access APEEE 3 more times throughout the day and then concluded testing by changing their passwords to withdraw access to the platform.

Visualize Engagement Outcomes Over Relative Time

The requirement for APEEE to support visualizations of events over relative time was of high priority for us to build out. We recognized the significant value that this functionality would add to APEEE, specifically in a research context where events are typically analyzed as a function of time elapsed in a study. To address this requirement, we sought to modify our existing analytic indicator for "how many check-ins are being completed daily" to have the x-axis represent time elapsed in the study. Visualizing check-in completion over time elapsed in the study supports determining effective engagement with *iCanCope* because the behavior of checking into the app and reporting symptoms is theorized to mediate improved pain-related outcomes [15]. We initiated work on this visualization by reviewing the *iCanCope* data model to determine the exact event that signified a user enrolling into the pilot RCT. Following discussions with the iOuch research group to clarify the enrollment protocol and validate the order of operations, we selected the first time a user logged into *iCanCope* as the *genesis event* from which to start recording time elapsed in the study. We then modified the *iCanCope* data model to generate a *daysSinceGenesis* metadata tag on every event logged, thereby enabling events to be positioned along the study timeline. Once this new metadata tag was deployed and tested, we sought to visualize the number of users who completed a check-in over time elapsed in the study. To create this visualization, we implemented Elasticsearch's aggregations framework and applied 3 bucketing aggregations to our data: (1) aggregating all check-ins logged by participants since study launch, (2) aggregating the number of daily check-ins over time elapsed in the study (eg, days 0-56), and (3) aggregating check-ins by study allocation, which was a metadata tag that was already exposed on all *iCanCope* log data. We then applied Kibana's *line graph* filter over our data to visualize insights.

Figure 8 presents the visualization for the number of users who completed a check-in over time elapsed in the study, visualized through APEEE as a line graph, with the y-axis representing the number of users who completed a check-in and the x-axis representing the number of days elapsed in the study; the selected time range is 2 years. A color-coded legend on the right side of the chart identifies the study allocation for each line. With this relative time functionality, APEEE can support investigators to (1) monitor engagement outcomes in real time and (2) assess emerging outcome patterns and shifts across study groups over time.

Visualize Clinical Outcomes Over Relative Time

Equipped with the ability to create relative time visualizations, we endeavored to trial a final visualization before concluding our development cycle: a line graph of pain-related outcome scores reported by users over time elapsed in the study. An advantage to *iCanCope* was the in-app collection of clinical outcomes through the check-in feature. These data were stored on our servers as Fast Healthcare Interoperability Resources (FHIR), which is a data format that cannot be visualized on APEEE. We resolved this interoperability issue by transforming the FHIR data into log data through parsing out the outcome scores, injecting related user-level metadata, and then ingesting these data into Elasticsearch using Logstash. Once ingested, we applied the same Elasticsearch framework and aggregations for querying check-in data over time elapsed in the study.

Figure 8. The analytic indicator for "are users adhering to the check-in protocol?," visualized through APEEE as a line graph. APEEE: Analytics Platform to Evaluate Effective Engagement.

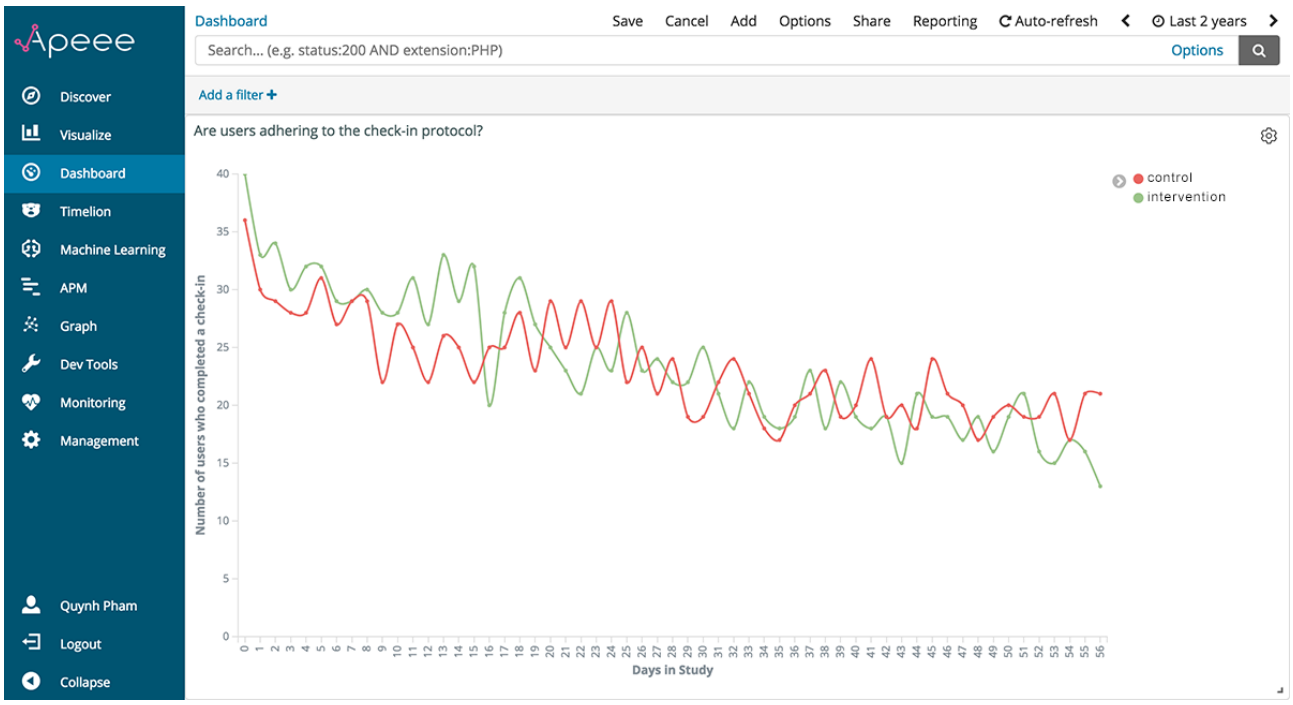


Figure 9. The analytic indicator for "Are intervention and control users reporting different pain scores?," visualized through APEEE as a line graph. APEEE: Analytics Platform to Evaluate Effective Engagement.

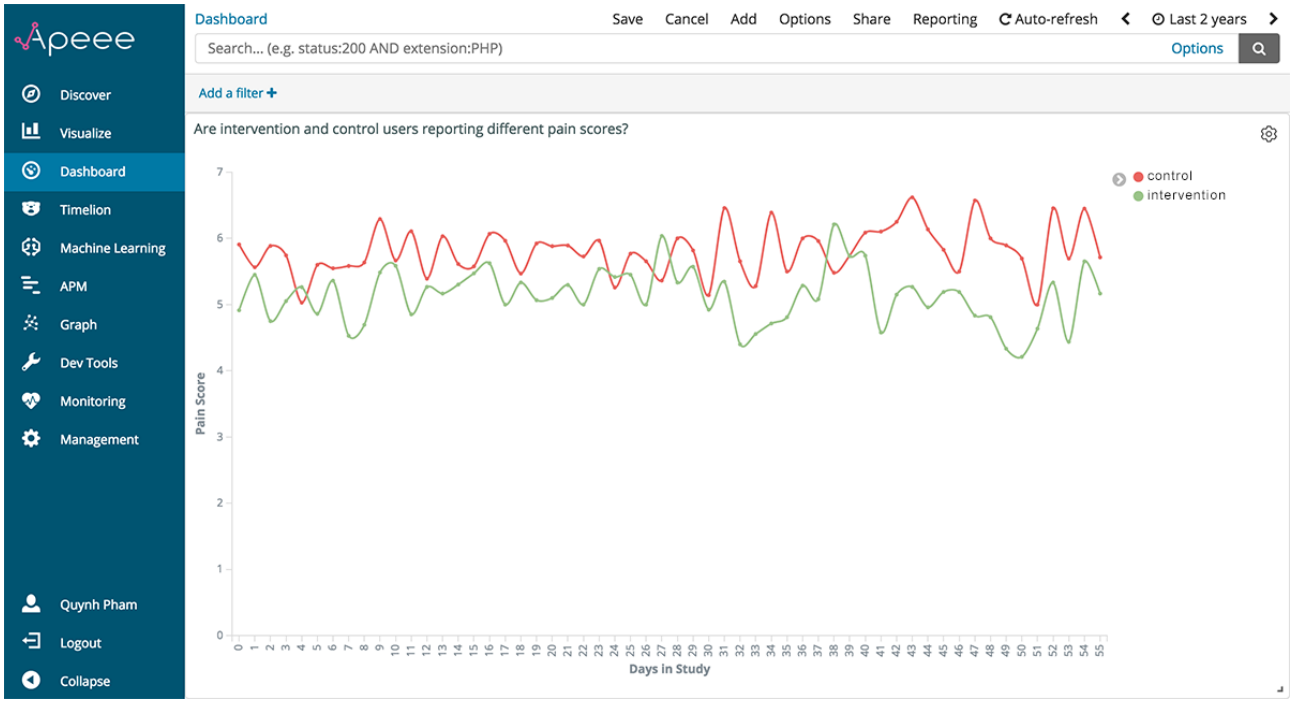


Figure 9 presents the visualization for pain scores reported by users over time elapsed in the study, visualized through APEEE as a line graph, with the y-axis representing pain scores and the x-axis representing the number of days elapsed in the study; the selected time range is 2 years. A color-coded legend on the right side of the chart identifies the study allocation for each line. The ability to monitor real-time changes to clinical outcomes over the course of a study may encourage investigators to adopt innovative methodologies in their mHealth evaluations [23,24].

A live demo of the updated APEEE platform, including both relative time visualizations, was presented to 7 members of the iOuch research group during a weekly laboratory meeting at the evaluation site. The visualizations were well received, and the represented data were perceived to be significantly more useful when graphed over relative time. Investigators were particularly surprised by the sustained pattern of adherence to the check-in protocol by users in the control group and engaged in a spirited discussion of the potential motivations for this behavior. Overall, investigators found the updated build of APEEE to better meet their evaluative needs. A collective decision was made to proceed with a full deployment of the platform to the iOuch research group as part of a field study in October 2018.

Discussion

Principal Findings

At a time of rapid advancement in the mHealth field, evaluations of pediatric mHealth apps for chronic conditions must keep pace to increase the volume of evidenced apps made available to young people [25]. A shift toward adopting data-driven research methods would mark a significant development for the field, which has historically been "data-rich but evidence-poor" [26,27]. We posit that the adaptation of pediatric mHealth apps at the right time and under the right circumstances can accelerate evaluative practice and improve health outcomes. In this paper, we have shown that the process of defining, operationalizing, and evaluating effective engagement with *iCanCope* can be automated through APEEE. To our knowledge, APEEE is the first application of the Elastic Stack in a digital health context to support mHealth evidence generation. Configuring the platform to integrate with the app was feasible and provided investigators with a resource to consolidate, analyze, and visualize engagement data generated by participants in real time. Preliminary efforts to evaluate APEEE showed that investigators perceived the platform to be an acceptable evaluative resource and were satisfied with its design, functionality, and performance. Furthermore, they expressed enthusiasm for adopting the platform to support their evaluative practice once fully implemented. Future research is required to formally evaluate the impact of the platform on evaluative practice and mHealth app effectiveness.

Limitations

Some methodological and functional limitations of our research warrant discussion. First, having a small number of members from a single research group participate in our evaluation was a major limitation and may have introduced bias, given the likelihood of shared perspectives. Second, our decision to build APEEE using the Elastic Stack exposes the platform to open-source updates made by the community of Elastic developers. This effectively means that changes may be pushed to APEEE's features and functionality and implemented with little warning. We perceive this risk to be minimal and acceptable for the following reasons: (1) since initiating work on APEEE, all updates to the Elastic Stack have added value to the platform (eg, faster Elasticsearch queries, streamlined Kibana visualization builder) at no cost to our project, and (2) we are able to overwrite undesirable changes through branching

the Elastic Stack source code and writing a version of the code for APEEE. Third, we were not able to address the suggestion for APEEE to compute advanced predictive statistical analyses in time for validation during our second evaluation cycle. We have since been able to graphically represent a series of probability distributions (eg, box plots and scatter plots) and interval estimations (eg, CIs and error bars) on APEEE using the Vega visualization grammar, which is a declarative language for building interactive graphs [28,29]. We will continue these preliminary explorations into VEGA-enabled predictive modeling and aim to validate this functionality in the field study of APEEE. Finally, although we were able to connect *iCanCope* to APEEE with relative ease, this process may not be indicative of the work effort required to connect a third-party mHealth app that we did not develop. APEEE benefits from the extensive Elastic Stack documentation and community resources (eg, blogs, YouTube videos, and forums) that detail the technical work effort required for connection [30]. However, the service design considerations for this connection are consequential [31] and may include (1) obtaining research ethics approval, (2) drafting data sharing agreements, and (3) reaching a shared understanding of what constitutes effective engagement and how to interpret analytic insights.

Conclusion

Dynamic, real-time analytic dashboards such as the one discussed in this paper provide investigators with a powerful means to characterize the breadth and depth of mHealth app engagement required to achieve intended health outcomes. Through APEEE, participant engagement with *iCanCope* can be modeled with pain-related outcomes data to provide data-driven and actionable feedback. For example, daily check-in frequency can be analyzed against pain severity to inform a contextualized interpretation of app effectiveness. Using this information, the evaluative approach to evidencing *iCanCope* and its modular features can be optimized. Indeed, APEEE may enable the identification of digital biomarkers across chronic conditions for use in developing predictive engagement algorithms to tailor the content and timing of mHealth intervention delivery. In this way, the platform may contribute to the realization of effective and evidence-based mHealth care.

Conflicts of Interest

None declared.

References

1. Pham Q, Wiljer D, Cafazzo J. Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. JMIR Mhealth Uhealth 2016 Sep 09;4(3):e107 [FREE Full text] [doi: 10.2196/mhealth.5720] [Medline: 27613084]
2. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
3. Lalloo C, Jibb L, Rivera J, Agarwal A, Stinson J. "There's a Pain App for That": review of patient-targeted smartphone applications for pain management. Clin J Pain 2015 Jun;31(6):557-563. [doi: 10.1097/AJP.0000000000000171] [Medline: 25370138]
4. Smith K, Iversen C, Kossowsky J, O'Dell S, Gambhir R, Coakley R. Apple apps for the management of pediatric pain and pain-related stress. Clin Pract Pediatr Psychol 2015 Sep 01;3(2):93-107 [FREE Full text] [doi: 10.1037/cpp0000092]
5. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
6. Noser A, Cushing C, McGrady M, Amaro C, Huffhines L. Adaptive intervention designs in pediatric psychology: the promise of sequential multiple assignment randomized trials of pediatric interventions. Clin Pract Pediatr Psychol 2017 Jun;5(2):170-179. [doi: 10.1037/cpp0000185]
7. Yardley L, Choudhury T, Patrick K, Michie S. Current issues and future directions for research into digital behavior change interventions. Am J Prev Med 2016 Dec;51(5):814-815. [doi: 10.1016/j.amepre.2016.07.019] [Medline: 27745680]
8. 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 Dec;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
9. Riley WT, Glasgow RE, Etheredge L, Abernethy AP. Rapid, responsive, relevant (R3) research: a call for a rapid learning health research enterprise. Clin Transl Med 2013 May 10;2(1):10 [FREE Full text] [doi: 10.1186/2001-1326-2-10] [Medline: 23663660]
10. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]
11. Chan Y, Wang P, Rogers L, Tignor N, Zweig M, Hershman S, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol 2017 Apr;35(4):354-362 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]

12. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 07;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
13. 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;8.
14. Tignor N, Wang P, Genes N, Rogers L, Hershman SG, Scott ER, et al. Methods for clustering time series data acquired from mobile health apps. Pac Symp Biocomput 2017 Jan;22:300-311 [FREE Full text] [doi: 10.1142/9789813207813_0029] [Medline: 27896984]
15. Stinson JN, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. iCanCope with Pain™: user-centred design of a web- and mobile-based self-management program for youth with chronic pain based on identified health care needs. Pain Res Manag 2014;19(5):257-265 [FREE Full text] [Medline: 25000507]
16. iOuch Pain Lab Internet. Improving Outcomes in Child Health through Technology URL: http://lab.research.sickkids.ca/iouch/ [accessed 2018-11-21] [WebCite Cache ID 746Il8pEu]
17. Ryu S. Book review: mHealth: new horizons for health through mobile technologies: based on the findings of the second global survey on eHealth (Global Observatory for eHealth Series, Volume 3). Healthc Inform Res 2012;18(3):231. [doi: 10.4258/hir.2012.18.3.231]
18. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
19. Morita P, Cafazzo J. Challenges and paradoxes of human factors in health technology design. JMIR Hum Factors 2016 Mar 01;3(1):e11 [FREE Full text] [doi: 10.2196/humanfactors.4653] [Medline: 27025862]
20. Elastic Stack 6.5. URL: https://www.elastic.co/products [accessed 2018-04-19] [WebCite Cache ID 70dQo7UxY]
21. Thakker D, Nguyen-Huu D. Forbes Internet. Another Open-Source IPO Shows the Market Power of Free Software URL: https://www.forbes.com/sites/dharmeshthakker/2018/09/12/another-open-source-ipo-shows-the-market-power-of-free-software/ [accessed 2018-11-21] [WebCite Cache ID 746JApUOJ]
22. Elastic. Why Open Source? URL: https://www.elastic.co/about/why-open-source [accessed 2018-11-21] [WebCite Cache ID 746J2lz8I]
23. Canadian Institutes of Health Research. Innovative Clinical Trials Initiative Internet URL: http://www.cihr-irsc.gc.ca/e/49773.html [accessed 2018-11-21] [WebCite Cache ID 746J5eCh4]
24. Klasnja P, Hekler E, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34S:1220-1228 [FREE Full text] [doi: 10.1037/hea0000305] [Medline: 26651463]
25. Clough B, Casey L. Smart designs for smart technologies: research challenges and emerging solutions for scientist-practitioners within e-mental health. Prof Psychol Res Pract 2015;46(6):429-436. [doi: 10.1037/pro0000053]
26. Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. Am J Prev Med 2016 Dec;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]
27. Riley WT. A new era of clinical research methods in a data-rich environment. In: Oncology Informatics. United States: Academic Press; 2016:55.
28. Github. Vega - A Visualization Grammar URL: https://vega.github.io/vega/ [accessed 2018-11-21] [WebCite Cache ID 746Jg1A6y]
29. Satyanarayan A, Moritz D, Wongsuphasawat K, Heer J. IEEE. 2017. Vega-Lite: a grammar of interactive graphics URL: https://idl.cs.washington.edu/files/2017-VegaLite-InfoVis.pdf [accessed 2018-11-23] [WebCite Cache ID 748bY3ACZ]
30. Elastic. Elastic Stack Product Documentation URL: https://www.elastic.co/guide/index.html [accessed 2018-11-21] [WebCite Cache ID 746JJF3kI]
31. Shaw J, Agarwal P, Desveaux L, Palma D, Stamenova V, Jamieson T. Beyond “implementation”: digital health innovation and service design. NPJ Digit Med 2018;1 [FREE Full text]

Abbreviations

APEEE: Analytics Platform to Evaluate Effective Engagement
CSV: comma-separated value
FHIR: Fast Healthcare Interoperability Resources
iOuch: Improving Outcomes in Child Health through Technology
IP: internet protocol
mHealth: mobile health
RCT: randomized controlled trial
UCD: user-centered design

Edited by G Eysenbach; submitted 03.07.18; peer-reviewed by J Pollak, A Pichon, S Davis; comments to author 07.10.18; revised version received 12.10.18; accepted 29.10.18; published 21.12.18.

Please cite as:
Pham Q, Graham G, Lalloo C, Morita PP, Seto E, Stinson JN, Cafazzo JA
An Analytics Platform to Evaluate Effective Engagement With Pediatric Mobile Health Apps: Design, Development, and Formative Evaluation
JMIR Mhealth Uhealth 2018;6(12):e11447
URL: http://mhealth.jmir.org/2018/12/e11447/
doi:10.2196/11447
PMID:30578179

Original Paper

Data Integrity–Based Methodology and Checklist for Identifying Implementation Risks of Physiological Sensing in Mobile Health Projects: Quantitative and Qualitative Analysis

Jia Zhang[1], MSc; Laura Tüshaus[1], PhD; Néstor Nuño Martínez[2,3], MSc; Monica Moreo[1], MSc; Hector Verastegui[4], Lic; Stella M Hartinger[2,3,4], PhD; Daniel Mäusezahl[2,3], PhD; Walter Karlen[1], Prof Dr

[1]Mobile Health Systems Lab, Institute of Robotics and Intelligent Systems, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

[2]Department of Epidemiology & Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland

[3]University of Basel, Basel, Switzerland

[4]Universidad Peruana Cayetano Heredia, Lima, Peru

Corresponding Author:
Walter Karlen, Prof Dr
Mobile Health Systems Lab, Institute of Robotics and Intelligent Systems
Department of Health Sciences and Technology
ETH Zurich
Building BAA
Lengghalde 5
Zurich, 8092
Switzerland
Phone: 41 44 63 3 77 54
Email: walter.karlen@ieee.org

Abstract

Background: Mobile health (mHealth) technologies have the potential to bring health care closer to people with otherwise limited access to adequate health care. However, physiological monitoring using mobile medical sensors is not yet widely used as adding biomedical sensors to mHealth projects inherently introduces new challenges. Thus far, no methodology exists to systematically evaluate these implementation challenges and identify the related risks.

Objective: This study aimed to facilitate the implementation of mHealth initiatives with mobile physiological sensing in constrained health systems by developing a methodology to systematically evaluate potential challenges and implementation risks.

Methods: We performed a quantitative analysis of physiological data obtained from a randomized household intervention trial that implemented sensor-based mHealth tools (pulse oximetry combined with a respiratory rate assessment app) to monitor health outcomes of 317 children (aged 6-36 months) that were visited weekly by 1 of 9 field workers in a rural Peruvian setting. The analysis focused on data integrity such as data completeness and signal quality. In addition, we performed a qualitative analysis of pretrial usability and semistructured posttrial interviews with a subset of app users (7 field workers and 7 health care center staff members) focusing on data integrity and reasons for loss thereof. Common themes were identified using a content analysis approach. Risk factors of each theme were detailed and then generalized and expanded into a checklist by reviewing 8 mHealth projects from the literature. An expert panel evaluated the checklist during 2 iterations until agreement between the 5 experts was achieved.

Results: Pulse oximetry signals were recorded in 78.36% (12,098/15,439) of subject visits where tablets were used. Signal quality decreased for 1 and increased for 7 field workers over time (1 excluded). Usability issues were addressed and the workflow was improved. Users considered the app easy and logical to use. In the qualitative analysis, we constructed a thematic map with the causes of low data integrity. We sorted them into 5 main challenge categories: environment, technology, user skills, user motivation, and subject engagement. The obtained categories were translated into detailed risk factors and presented in the form of an actionable checklist to evaluate possible implementation risks. By visually inspecting the checklist, open issues and sources for potential risks can be easily identified.

Conclusions: We developed a data integrity–based methodology to assess the potential challenges and risks of sensor-based mHealth projects. Aiming at improving data integrity, implementers can focus on the evaluation of environment, technology, user skills, user motivation, and subject engagement challenges. We provide a checklist to assist mHealth implementers with a structured evaluation protocol when planning and preparing projects.

(JMIR Mhealth Uhealth 2018;6(12):e11896) doi:10.2196/11896

KEYWORDS

physiological monitoring; data completeness; data quality; signal quality; medical sensors; implementation research; content analysis; mHealth; digital health

Introduction

Background

Limited access to adequate health care is a major burden in low- and middle-income countries and affects the poor most [1]. Centralized and outreach health care facilities are often sparsely available, difficult to reach, and overloaded. In addition, access to the health care centers can be costly, as patients often have to pay for transportation and compensate for the loss of income because of their absence from work [2]. Mobile health (mHealth) is a promising field that seeks to bring health care closer to the patient, thereby improving access and reducing costs because of its potential for a system-wide application [3]. We interpret mHealth as the use of mobile, digital communication technologies (eg, mobile phones) in medical and public health applications for effectively delivering health care and medical information [4]. Biomedical sensing using connected mobile sensors is an important but largely unexplored application in mHealth. It provides objective measurement of physiological parameters and facilitates more reliable diagnoses and assessments of patients. Physiological parameters that can currently be measured with mobile tools include blood pressure [5], respiratory rate (RR) [6], heart rate (HR) and electrocardiogram [7], peripheral capillary oxygen saturation (SpO_2) [8], and blood glucose levels [9].

The integration of additional medical sensors into mHealth projects increases the technological complexity. Furthermore, users require additional skills and medical knowledge, whereas systems need to be purchased and maintained. Thus, these additional challenges need to be considered during the implementation of physiological monitoring projects. The validated use of medical sensors depends on well-defined working conditions and the adherence to standards to ensure correct sensor function and data quality. Sensor failures and motion artifacts are possible intermittent issues encountered and, therefore, when operating in remote settings, a basic understanding of medical sensing mechanisms is required for safe application of sensors and to identify faulty or noisy data at the point of use. It can be challenging to address these issues when inexperienced community health care workers with little or no prior knowledge about interpreting physiological signals are operating the sensors. Numerous mHealth projects have implemented physiological sensors, for example, pulse oximeters, for measuring SpO_2 and HR, but none of them directly focused on evaluating the challenges associated with their implementation. Challenges were observed in clinical settings, that is, Hudson et al identified that the lack of training and nonfamiliarity with clinical alarms are barriers to apply pulse oximeters [10]. Furthermore, Spence et al identified different priorities across stakeholders [11], and English et al identified significant differences in observed errors between clinicians and nursing staff [12]. In summary, no research study has systematically examined the challenges of implementing physiological sensing and monitoring with mHealth tools.

As a consequence, no established methodology exists that would enable mHealth implementers to formally evaluate their projects and prevent implementation pitfalls with respect to physiological monitoring in low-resource settings. Although King et al organized focus group discussions with trained health care providers to identify challenges when managing pediatric pneumonia with pulse oximetry [13], their findings are country specific and limited to pulse oximeters. Wallis et al organized group discussions and proposed a roadmap for overcoming barriers of implementing image-based mHealth implementations [14], but their strategies are limited to image-based applications. On the other hand, Aranda-Jan et al applied the strengths, weaknesses, opportunities, and threats analysis method to review mHealth projects [15]. In addition, Eckman et al provided a conceptual strategy that involved all stakeholders into the design phase to assess the common failures of mHealth implementation [16]. However, both approaches did not explicitly address the challenges of physiological sensing and the specific risks associated to adding medical sensors to mHealth projects. The absence of a methodology or guideline during implementation could easily lead to overlooking domain-specific issues, evaluation errors, and the underestimation of risks and, therefore, prevent projects from achieving their goal of improving health outcomes.

We consider data integrity as the most important criterion for evaluating the risks of an mHealth project. Data integrity represents the faithfulness of information comprising criteria such as completeness, accuracy, relevance, consistency, usability, and reliability [17]. During unsupervised data collection, as it is frequently the case in mHealth, data completeness and consistency are critical quality metrics. Incomplete, poor, and missing data not only reduce the sample size but may also introduce bias or false conclusions. In clinical decision making, the signal quality and its reliability during physiological data collection using medical sensors are the most important factors [18]. Usability of a medical device is another component of data integrity that is associated with correct use and usage errors. International standards specify usability evaluation processes to reduce the risk of usability failures [19]. Poor usability can lead to the misuse of a medical device or a

reduction of user engagement, resulting in unusable or missing data.

Due to the decentralized nature of mHealth, the assurance of data integrity is challenging [20]. High measurement uncertainty because of the lack of a controlled environment, unknown training status of the user, and higher risks for misuse of the technology require special attention. Although the goal of any mHealth implementation is to provide access to health services and, consequently, improve health outcomes, obtaining good data integrity with the provided technology is essential to positively influence these outcomes. Therefore, evaluating data integrity should not only be part of the evaluation of implementation success at the end of an mHealth study but considered and assessed already early in the preparation phase. Consequently, data integrity could serve as the central theme when framing a methodology for evaluating implementation challenges.

Objectives

Our goal was to develop a methodology to systematically evaluate general risks and challenges of sensor-based physiological monitoring in mHealth and to avoid pitfalls before and during its implementation. Our specific aims in developing such a methodology were to (1) identify sources of low data integrity with a special focus on implementations that occur in remote or low-resource settings, (2) derive generalized risk factors that could guide a pre-implementation evaluation, and (3) provide an actionable tool to conduct such evaluation. The results could support implementers in evaluating their projects with regard to hidden risks and facilitate quality control early in the design and implementation of advanced mHealth tools.

Methods

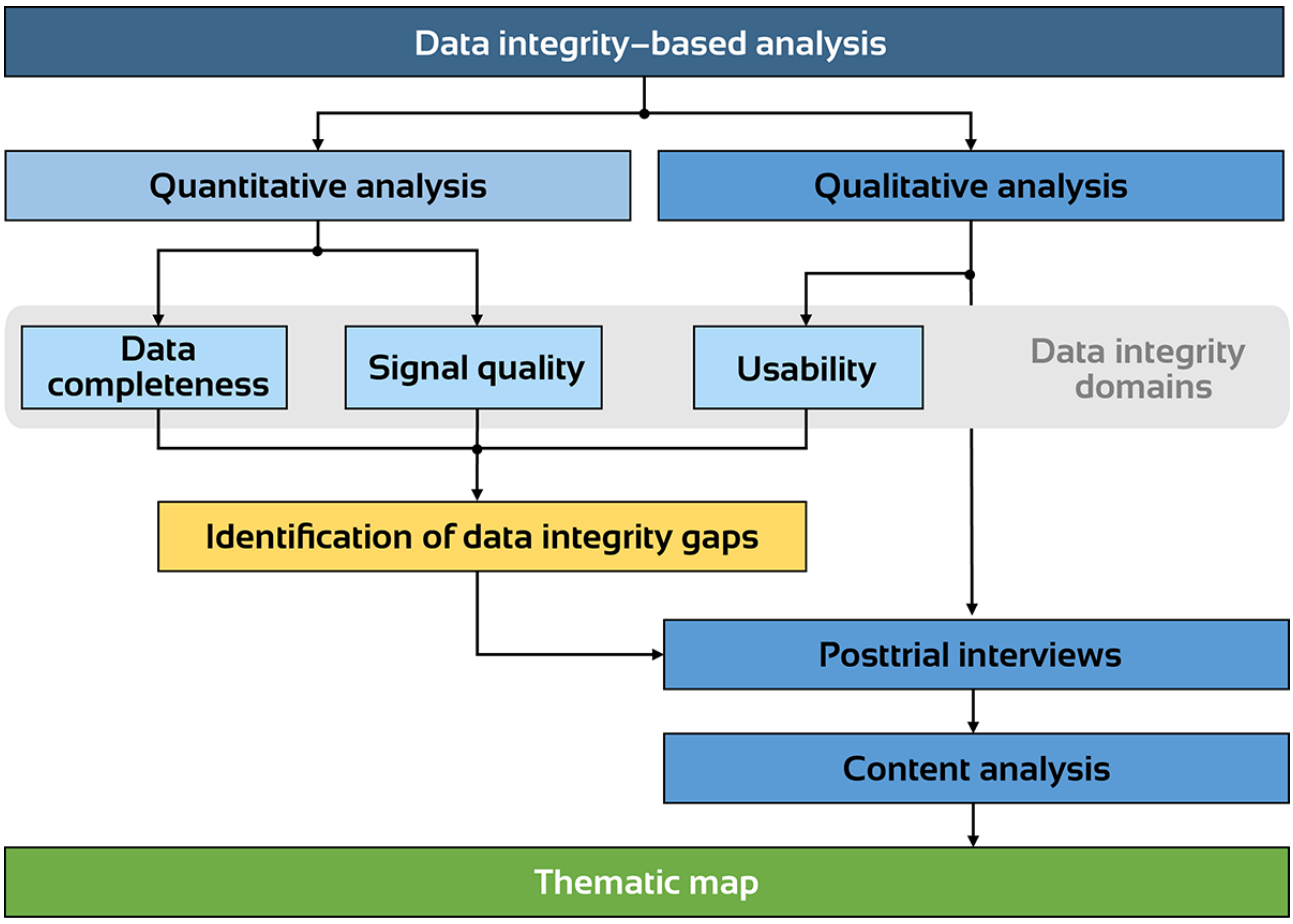
Overview

To identify sources of data integrity loss, we retrospectively analyzed physiological data collected from a randomized controlled trial that implemented sensor-based mHealth tools to assess health outcomes in a rural setting [21]. After the analysis of the data integrity gaps in the recorded data, we identified possible causes that could have led to these gaps from both the paper-based trial case report forms (CRFs) and through qualitative data obtained from posttrial semistructured interviews with the app users conducted on site after the trial. The method development process is shown in Figure 1.

Data Collection

We retrospectively analyzed physiological data and paper-based CRFs collected during a randomized controlled trial conducted in 82 Peruvian rural communities [21]. The trial was approved by the Universidad Peruana Cayetano Heredia ethical review board and the Cajamarca Regional Health Authority. The trial was registered on the ISRCTN registry (ISRCTN26548981).

Figure 1. The flowchart demonstrates the data integrity–based analysis for identifying the source of data integrity loss.



A total of 317 children aged between 6 and 36 months were enrolled, and informed written consent was obtained from the children's guardians. A total of 9 field workers (FWs) were trained to visit the children on 7 fixed geographical routes. Children were preassigned to these routes and visited in parallel by FWs once a week during the course of 60 weeks (6 weeks pilot, followed by a 54-weeks trial from February 2016 to May 2017, excluding 4 weeks of public holidays). To reduce the possibility of a courtesy bias, the routes of the FWs were rotated every 2 months.

During the weekly visits, FWs filled out a CRF and recorded physiological measurements with an mHealth app developed with LambdaNative (University of British Columbia, Canada) [22]. The app was installed on a tablet (Lenovo TAB 2 A7-10, Lenovo, CN) and connected to an external pulse oximetry sensor ($iSpO_2$ Rx, Masimo International, Neuchatel, CH). FWs placed the multisite Y probe on the child's thumb, index finger, or sole of a foot for the measurement of photoplethymogram (PPG), HR, and SpO_2. The FW also measured RR with the same app by tapping on the touch sensitive screen of the tablet with each inhalation phase of breathing while observing the child's bared belly [6]. All data collection procedures and interactions with the guardians and the child were subject of the informed consent and were approved by the ethics board.

The global positioning system sensor of the tablet registered the location where the visits took place (usually at the subject's home). The assigned identification codes for children and FWs, date, and time were recorded with the app and the CRF. Furthermore, the health status of the child in the preceding week (maximum 2-week recall), the availability of the child (eg, absent from home), and unexpected sensor- or app-related technical problems during the visit were annotated in the CRF. Field coordinators conversed daily with FWs whether any problems occurred during the day, downloaded data, tested the sensors, and charged the tablets for the following day.

In addition to the assessments by the FWs, health personnel from 22 health care centers in the trial's catchment area used the same tablets and software to collect physiological data in their consultations. The FWs received a 5-day initial training for tablet and CRF data collection with monthly retraining sessions of 2 hours. The health care center personnel were initially trained in 2 group sessions. Due to frequent changes of personnel in health care centers, new staff was retrained individually on site and physiological data were downloaded on a monthly basis.

Quantitative Data Analysis

We quantitatively evaluated data completeness and signal quality of the physiological data and CRFs completed by FWs (N=9) with Matlab (R2016b, MathWorks Inc, Natick, Massachusetts, USA).

Data Completeness

We analyzed the completeness of home visit data and explored reasons for missing data. For this assessment, we considered a child no longer contributing to our data integrity analysis if there were no visits available for more than 8 consecutive weeks during the main trial period. We analyzed the tablet and CRF data separately. We considered the visit as missed if there were no tablet or CRF entries during a given week. We compared the data completeness between the pilot trial and main trial to assess training effects and improvements because of feedback from the pilot period. In the case of missing visits, we consulted with the field coordinator that was responsible for the FWs route planning for possible reasons. In addition, we reviewed the CRFs for potential explanations for the missing visits or recordings. For health care center recordings, we investigated barriers of using the tablet from interviews with the staff members.

Signal Quality

We evaluated the signal quality of the waveform obtained from the pulse oximeter. We calculated a signal quality index (SQI) using the established cross-validation based on morphological features and short-term variations [23]. We classified the PPG signal into 2 quality categories. We defined PPG signals that had sufficient quality to extract SpO_2 values as "sufficient" (time series with high SQI for consecutive 8 seconds or longer) and the remaining as "insufficient". To evaluate the performance across FWs over time, we evaluated the PPG signal quality for each FW separately. We calculated a "sufficient" PPG ratio over the total number of PPG signals within a sliding window of 40 recordings and a step size of 8 recordings. We chose these specific numbers because ideally each of the FWs should have obtained approximately 40 recordings per week and 8 recordings per day.

Qualitative Data Analysis

We conducted semistructured posttrial interviews with the 7 FWs who were last to complete the children's visits to assess their routines, experiences, and problems encountered during data collection. In addition, we conducted interviews with 7 health care center staff members (nurses or technicians) who were trained to use the tablet and worked at 7 different health care centers. These 7 health care centers were selected because of their varied geographical distribution, infrastructure, and load of patients. We assessed the frequencies and difficulties of using the tablet (see Multimedia Appendix 1). JZ and MM conducted the face-to-face interviews. Questions were asked in English and translated into Spanish during the interviews. All interviews were recorded with written notes and later digitalized by JZ and MM. Spontaneous follow-up questions and answers were also included in the analysis. Furthermore, we investigated potential usability issues that were not identified during the app development and trial pilot phase as well as whether the users had any difficulties using the tablet.

We conducted a content analysis [24] on the qualitative data, resulting in predominant themes around potential reasons that could affect the 3 main sources for data integrity (data completeness, signal quality, and usability). JZ collected and familiarized with the data, coded the reasons, and searched for themes. The final themes were discussed with LT and WK and reviewed by DM and WK. JZ then created a thematic map of potential reasons that could cause insufficient data integrity by identifying commonalities among all codes.

Generalizing Risk Factors and Checklist Development

We systematically evaluated the obtained challenge categories to derive a methodology that could guide the pre-implementation evaluation of risks for general physiological sensing projects. We interpreted the main themes generated from the thematic map as challenges to be assessed and detailed each of them into specific risk factors based on the observed experiences during the trial. The risk factors were aggregated by JZ into a checklist draft.

To generalize the risk factors in this pulse oximetry–based checklist draft to other physiological sensing approaches, we selected 8 studies [25-32] that we considered representative of medical sensors–based mHealth projects (details are listed in Multimedia Appendix 2). A total of 4 graduate students (JB, SH, MM, and NN) with experience in conducting projects in low-resource settings reviewed and evaluated 2 selected projects each and applied the checklist to the selected projects. The list of risk factors was expanded with factors that were missing, either identified by the authors of the reviewed projects or from the reviewers' own experiences. The wording and usability issues of the checklist were improved based on the feedback from the reviewers.

A total of 5 researchers (AA, DC, KK, WK, and BP) with proven practical experience in global mHealth implementation were invited by email to join an expert panel, assess the checklist, and provide feedback in 2 evaluation rounds. The first round was conducted via email to collect individual feedback on the checklist and suggestions for change from each expert. JZ aggregated all feedback into a point-by-point list of change recommendations and distributed it to all experts for review before the second round. The second round consisted of a group discussion that was conducted via videoconference. The list was presented point-by-point to the experts (JZ) and in case of disagreements between experts, discussed until a final agreement was reached. JZ translated decisions on changes into the checklist, which was then distributed to experts for final approval.

Results

Quantitative Data Analysis

Data from 300 out of the 317 recruited children met the inclusion criteria for the quantitative data analysis. A total of 15,757 home visits were made to these children during the trial and 1589 during the pilot period (Figure 2). We observed a higher percentage of visits entered through CRFs during the trial (15,322/15,757, 97.27%) compared with the pilot (13,910/15,757, 88.28%; Table 1). FWs encountered the children at home in 13,802 (13,802/15,757, 87.59%) cases. In 1953 cases (1953/15,757, 12.39%), children were absent from home, and hence, no data could be collected.

Figure 2. Visits obtained from tablet and case report form (CRF) entries over study weeks. The pixels in the order of legend sequence denote Both: visits registered both with the tablet and in the case report form, Tablet: visits only entered in the tablet, CRF: visits only entered in the CRF, and Missing: no visit recorded with either tablet or CRF. The continuous black lines indicate 4 full weeks of public holidays in the trial region. Missing visits at week 60 were because of Easter vacation.

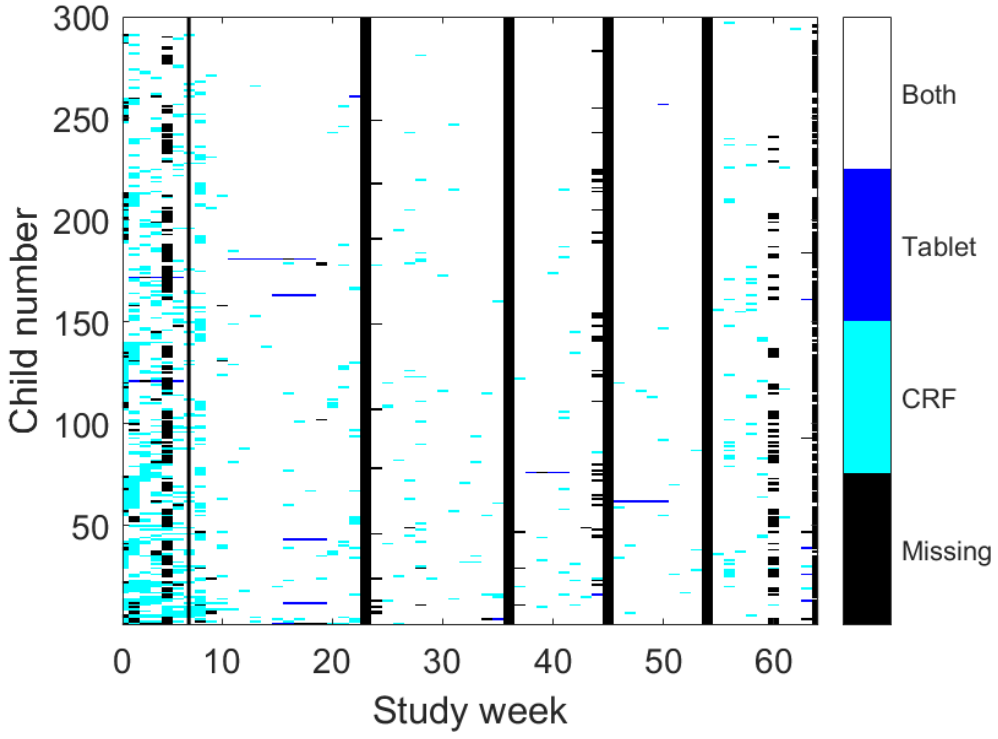


Table 1. Overview of the quantitative data from the 300 included children with respect to data completeness and photoplethymogram signal quality collected in the field during the pilot and trial.

Visits	Pilot (N=1800[a]), n (%)	Trial (N=16,200[a]), n (%)
Actual visits (tablet)	**n=1182 (65.67)**	**n=15,439 (95.30)**
Total PPG[b]	977 (82.66)	12,098 (78.36)
Sufficient quality PPG	368 (37.67)	7653 (63.26)
Insufficient quality PPG	609 (62.33)	4445 (36.74)
Actual visits (CRF[c])	**n=1589 (88.28)**	**n=15,757 (97.27)**
Successful visits	1212 (76.27)	13,802 (87.59)
Unsuccessful visits	377 (23.73)	1953 (12.39)
Unlabeled visits	0 (0.00)	2 (0.01)

[a]N values based on scheduled visits.

[b]PPG: photoplethymogram.

[c]CRF: case report form.

Overall, 2 FWs left the study team during the trial. FW 5 left because of personal reasons after 6 months and was replaced by FW 6. FW 9 left already after 218 recordings that were insufficient for estimating a signal quality trend and, consequently, was excluded from the comparison. In total, the remaining 8 FWs recorded 82.66% (977/1182) PPG measurements during the pilot and 78.36% (12,098/15,439) PPG measurements during the trial (Table 1). For the trial, we classified 7653 (7653/12,098, 63.26%) PPG signals as "sufficient" and 4445 (4445/12,098, 36.74%) as "insufficient". Of the 8 FWs, 7 increased their "sufficient" PPG ratio over time with a mean slope of 0.1226 (SD 0.0512; Figure 3).

Qualitative Analysis

After the interviews with 7 FWs and 7 health care center staff members, we identified sources of low data integrity in 3 data integrity domains: (1) reasons regarding incomplete data, (2) low signal quality, and (3) usability issues.

Data Completeness

FWs encountered difficulties to find the correct routes to the family homes at the beginning of the pilot because of long distances and rough roads. To arrange efficient routes for each FW, the field coordinators evaluated the number of children per route, the actual duration to complete each route, and a rotation of FWs to share extra workload for routes to remote communities or hardship during difficult weather conditions. The pilot enabled to adjust the routes and refine data collection tools and protocols. After adaptation, we observed that a higher percentage of children were visited during the trial compared with the pilot. Furthermore, FWs noticed that if children were absent from their homes during the scheduled visits, it was mainly because the guardians had taken them to the fields as most of them were farmers.

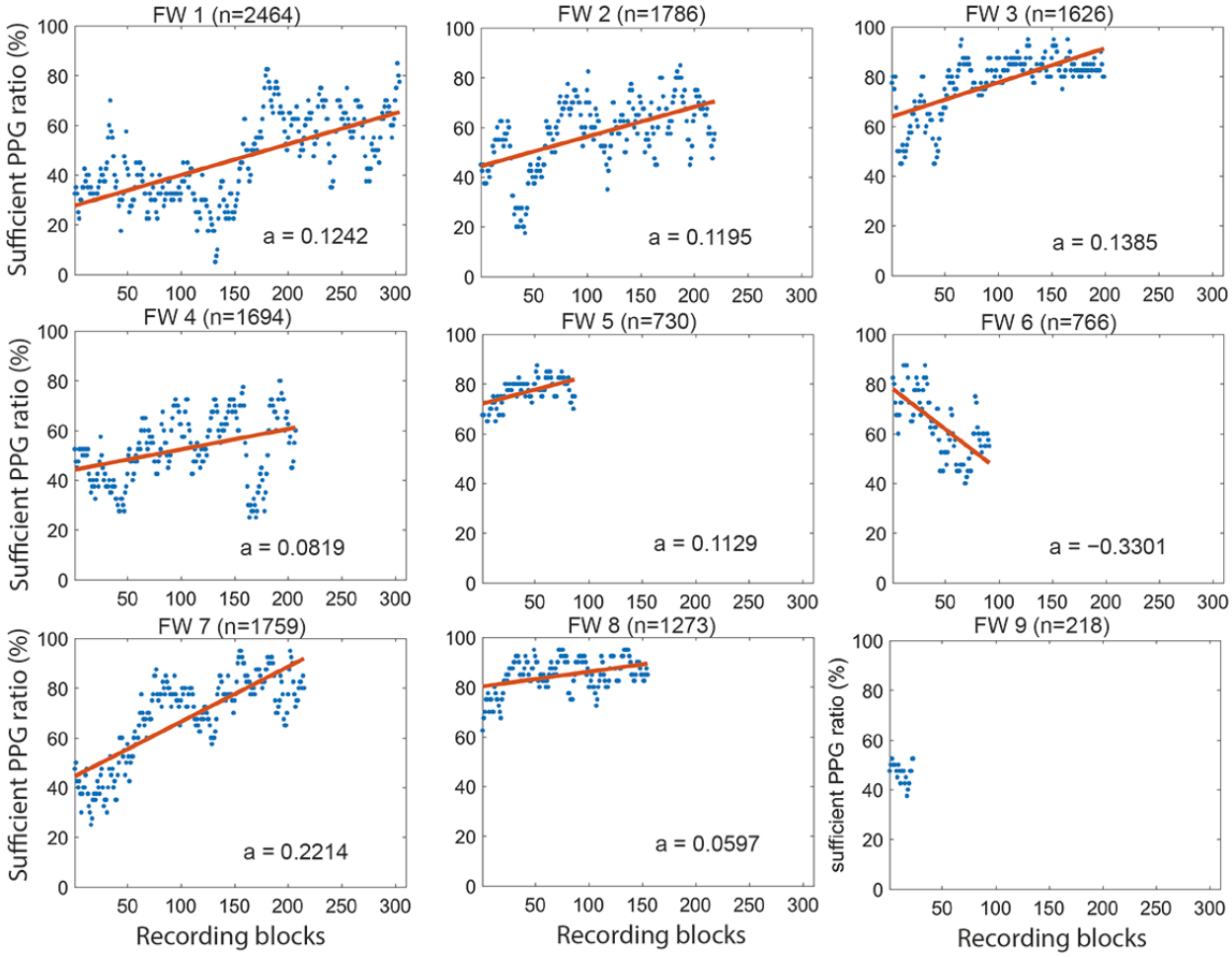
FWs reported issues with the tablets and sensors, specifically the freezing of the app during measurements (3 FWs), that no pulse oximeter connection could be established (3 FWs) or unexpected insufficient tablet battery levels (1 FW) to perform all measurements as planned. All these issues were addressed by reporting to and solved daily by research assistants when FWs returned to the research station.

Another factor that hindered the measurement was guardians' concern and preference not to let FWs interact with the children when the children resisted cooperating, when they were sick, or were sleeping. According to the CRFs, mothers did not allow measurements of the child in 305 cases. FWs also reported that when a child was sick, the mothers did not allow baring the child's chest and abdomen to measure respiratory movement.

Most FWs (5 out of 7) perceived the lack of rapport with the child as a hindering factor at the beginning of the trial and after route rotations. They reported that the child was agitated and nervous and, therefore, resistant to interact. This problem was eventually solved and the trust between children and FWs built up over time.

In general, health care center staff were eager to use the tablet to measure the 3 parameters (HR, SpO_2, and RR) using a single system. However, staff changes and extra workload were reasons for the low usage of the tablet. In 4 out of 7 health care centers where the interviews took place, the trained health care center staff member quit their job with the health service provider unexpectedly before a new staff member could be instructed to use the tablet. In addition, 1 health care center staff member indicated that health care center staff members were unable to spend extra time to collect measurements with the tablet because they had to complete their routine paper registrations and measurements for visiting patients with their regular medical devices.

Figure 3. "Sufficient" photoplethymogram quality ratio over recording blocks for all 9 field workers (field worker 1-9 [n=number of photoplethymogram recordings performed]) during the trial. The blue dots depict the ratios between number of "sufficient" photoplethymogram signals and total number of photoplethymogram signals within each recording block (40 consecutive recordings with a step size of 8) by each field worker and the red trend lines are the linear fit of the ratios estimating the trend of recording quality (a=slope of trend line). Field worker 9 did not produce sufficient recordings for meaningful trend estimation and was not included in the signal quality analysis. FW: field worker; PPG: photoplethymogram.



Signal Quality

The FWs reported that cold fingers and movements of the children led to poor signal quality. For most of the visits when ambient temperatures were low, the pulse oximeter was not able to acquire a signal and the app indicated insufficient perfusion. With the progression of the study, FWs addressed this problem by warming the child's finger before the measurement. The FWs also indicated that children tended to move after 10 seconds of measurements, leading to movement artifacts. In addition, children became nervous after approximately 3 unsuccessful measurement attempts and became less compliant.

Usability

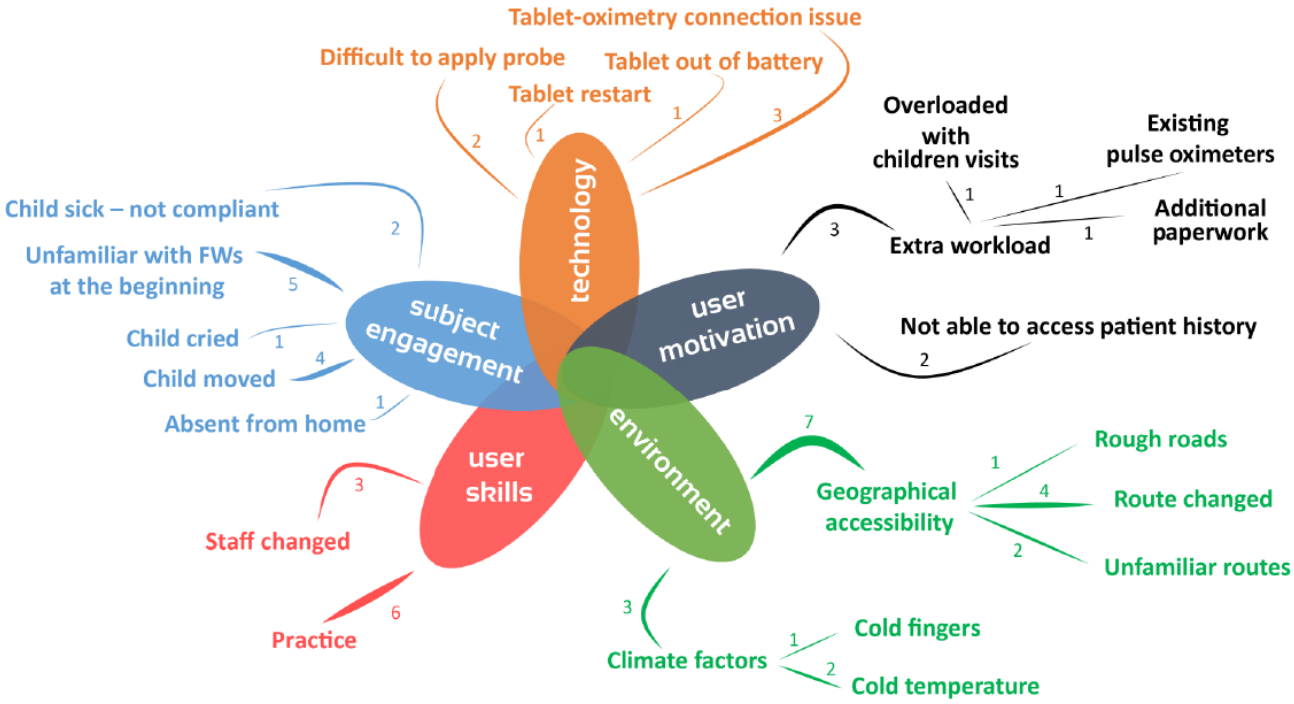
Usability was primarily assessed in the pilot phase where the app was iteratively improved day-by-day in close interaction with the FWs. Workflow issues were addressed and data entry speed optimized. Translations of instructions from English to Spanish were confusing and, consequently, simplified.

A single FW reported that the font size of the selection list for demographic information (eg, the child's communities and child's identifier) was too small and the selection lists were too long to go through. The remaining FWs considered the app easy to use with a logical workflow.

Thematic Map

From the coded reasons for loss of data integrity in the 3 studied data integrity domains (data completeness, signal quality, and usability), we obtained 5 clusters: (1) environment, (2) technology, (3) user skills, (4) user motivation, and (5) subject engagement, which were represented in a thematic map (Figure 4). The strength of connections between codes denotes the frequency of occurrence of the codes and, therefore, illustrates the importance of a code within the cluster. We identified these 5 clusters as main challenge categories for the implementation of mHealth physiological monitoring in low-resource settings.

Figure 4. Thematic map showing 5 main challenge categories (technology, environment, user skills, user motivation, and subject engagement) generated by the related codes. The numbers along the connection lines indicate the frequency of occurrence of the codes, therefore, indicating the weight each code contributes to the main challenge. FW: field worker.



Generalized Risk Factors and Evaluation Tool

The risk factors generated from the above-mentioned 5 challenge categories were expanded through further mHealth projects evaluation and expert reviews and were consolidated into a checklist. The checklist was divided into 5 sections that relate to the main challenge categories obtained from the thematic map and serves as an actionable evaluation tool. “Technology” considers technical aspects of the system, mobile devices, measurement devices or medical sensors, data management, software, and technical support. “Environment” takes into consideration the risks from climate, geography, culture, and society that can influence the quality of data collection and technology performance. “User skills” considers literacy, training, feedback, and retraining of the users. “User motivation” considers user availability and monitoring strategies to encourage user performance. “Subject engagement” focuses on the availability of the subject to be measured. Each section of the questionnaire features questions that can be answered with either a “yes,” “no,” “in progress,” or “not applicable (N/A).” By inspecting the “no” column of the checklist, the open issues and sources for potential risks can be visually identified. The checklist is available under a CreativeCommons NonCommercial ShareAlike licence as a printable PDF and an interactive Web-based form [32].

Discussion

Principal Findings

In this study, we evaluated implementation challenges of physiological monitoring with mobile sensors in low-resource settings and developed a data integrity–based methodology to evaluate the challenges according to the factors environment, technology, user skills, user motivation, and subject engagement. This methodology could, that is, in the form of the developed checklist, assist mHealth implementers to identify risks.

Until now, methodologies for systematically assessing implementation challenges in physiological monitoring enabled by mHealth did not exist. Implementation challenges were reported only intermittently covering training [10], limited resources [34], motivational barriers [35], language and cultural barriers, weak health systems, and limited external financing schemes [36]. With our approach that focuses on the exploration of challenges based on data integrity, we provide central themes that implementers can systematically follow. By exploring the causality of data integrity loss, the methodology provides a broad coverage of risks.

Environment- and technology-related challenges are closely linked and should be evaluated with respect to the following aspects: weather, geography, population, and related difficulties that influence the access to subjects as well as the mHealth tool’s functionality. Unlike text or voice message–based mHealth projects, where the mobile communication infrastructure is the major bottleneck that influences study outcomes [36], environmentally induced barriers such as missing subject recordings because of inconvenient transportation have large impact on sensor-based mHealth projects. Those factors should be carefully considered and potential solutions tested and planned for.

In addition, implementers should plan for sufficient follow-up and technical support during the lifetime of a project. In our case, the cold climate made the children feel uncomfortable to bare their chest and abdomen and, in addition, cold fingers negatively influenced the signal strength. This problem could be addressed by considering whether the chosen sensing modalities are suitable for the local settings. Moreover, from

our experience, good preparation includes collaborating and exchanging information with all stakeholders (parents, caregivers, and health care center personnel) early in the process, which helps to evaluate the feasibility of the chosen system and methods before implementation [27].

To solve user skill–related challenges, sufficient training of the users, understanding their opinions and attitudes toward devices and systems, as well as assisting them in fostering a good relationship with the subjects are essential. The "sufficient" PPG ratio for all except 1 FW increased across the study period, indicating a positive correlation between the users' experience level and achieved signal quality. FW 6 who had a negative trend in signal quality was hired midtrial and was not part of the extensive training during the pilot study. Therefore, we cannot exclude the fact that the training provided at the appointment was insufficient. The posttrial interview with FW 6 did not reveal a clear reason why the decreasing trend could have happened. Therefore, further investigations will be needed. This was the first time that mHealth technology was introduced into the trial region. Although mobile phones were widely used in this area, the sensor-related mHealth tools were new to the users (FWs). We recommend training the users to apply the sensors within the target environment to ensure they are fully comfortable with the functionality and able to perform minor troubleshooting themselves as well as perform regular refresher trainings. In addition, implementers should develop evaluation methods to track and supervise the performance of the users during the project's lifetime and be prepared to receive feedback from users. This way, users can be trained and retrained based on specific issues encountered with the aim of increasing data quality and efficiency.

The subject engagement challenges relate to the level of cooperation between users and subjects. The positive engagement is one of the most important factors that contribute to data completeness. Moreover, medical sensors are sensitive to motion artifacts; therefore, collecting measurements from pediatric populations highly depends on their willingness to cooperate. First of all, the user should establish a good relationship with the subjects. We preemptively considered this as an important factor and conducted extensive pretrial training for FWs in 2 kindergartens and day cares to familiarize them with working with children. Our FWs tried to establish a friendly rapport and played games with the children to calm them before measurements. In general, users should practice measurements on the targeted subject population to optimally perform measurements while creating a conducive environment. In addition, communication with and gaining support from subjects' family members are essential. In our case, the parents' support in general was high. In this trial, no cultural groups rejected participation. For pediatric studies, parents should be encouraged to support the mHealth users in handling their children.

Although health care centers in low-resource settings are eager to use technological support to assist clinical measurements, the users faced motivational challenges. On the one hand, supervised training and observable benefits for the staff might increase their motivation to use the new technology. Haberer et al show that sufficient training and improved skills increase the motivation of users [37]. On the other hand, strong motivation also increases lay workers' performance. Mwendwa et al suggest that poor performance of the community health care workers cannot be solely solved by training skills but also by highlighting the consequences of the measurements and explaining the process of data collection [38]. A properly supervised training and explanations of the benefits of the mHealth tool have the potential to increase user motivation. However, as Graham et al identified in their recent study on implementation of handheld pulse oximetry in Nigerian hospitals, provision of equipment and training alone is not enough [39]. Reminders and encouragement of peers are needed as increased workload burden and technical difficulties were negatively influencing motivation to adopt pulse oximetry. Although these findings were not obtained from an mHealth implementation study, we have good reason to believe that this applies to technology implementation in general, including mHealth.

Checklists have been proven to raise awareness and prevent incidents of certain reoccurring issues. Pilots and aircrew perform preflight checklists to improve flight safety [40]. The World Health Organization suggests using a surgical safety checklist in operating room environments to reduce the number of surgical incidents and deaths [40]. Other health care–related checklists were developed such as assessing the scalability of pilot projects [41], reporting health interventions [42], checking mHealth solutions [43], and monitoring and evaluating outcomes of digital interventions [44]. However, the effectiveness of a checklist depends on the complete implementation of recommended actions. Van Klei et al showed that after applying the surgical safety checklist in operating rooms, the mortality rate only reduced significantly for those surgeons who fully completed the checklist [45]. Furthermore, to effectively distribute the checklist to targeted audiences as well as encourage its use is challenging. Therefore, we provide a tool online for easy and efficient assessment.

Historically, widespread adoption of mHealth tools is limited with too many proof-of-concept projects not achieving sustainable implementations and often lacking evidence to justify scaling [2]. The main challenge categories covered by our methodology coincide with the critical factors for success in scaling medical mobile technologies identified by Lundin and Dumont [46]. Besides understanding the needs from the local area, integrating the technology into the local health care systems, engaging end users, and involving all related stakeholders, other factors that are not driven by data integrity (eg, finance-related factors) can determine the scaling success of mHealth projects.

Limitations

Our methodology development is based on the physiological measurements performed in a single trial limited to pulse oximetry and RR measurements. Therefore, the 5 identified sources for loss of data integrity may not be equally weighted in other projects. For example, in a user self-management project, where the mHealth user is also the studied subject, the aspects of training and education become more important and, therefore, might require a stronger emphasis. Furthermore,

although the monitored trial implemented mHealth tools, it did not aim at scaling the usage of the tools. A scaling project could have, because of its extension to multiple geographical locations spanning over different health districts, slightly different aims and would have more sophisticated monitoring tools in place. Our methodology might have not comprehensively captured these aims. However, as our methodology is based on data integrity, the evaluation approach can be easily expanded to these differences.

We were not able to validate the effectiveness of the provided checklist prospectively on a large number of projects. We tested and expanded the checklist extensively by reviewing multiple published projects implementing medical sensors by using early drafts of the checklist and complementing missing aspects. Furthermore, an invited panel of experts evaluated and complemented the checklist with missing aspects based on their own diverse expertise. To promote adoption and collect feedback from early adopters, we have published the checklist online.

Outlook

To enable a dynamic growth of the checklist, we provide a digital form of the checklist online where anonymous usage of the checklist is tracked. We plan to use these data, together with direct feedback from implementers, to improve the checklist in regular intervals and redistribute updated versions through the same platform. As there is currently a lack of target product profiles for sensor-based mHealth systems in many disease management apps and our checklist is developed for implementers to reduce the risk of data integrity loss, we would like to explore the potential of the checklist to serve as a reference for building target product profiles that call for high degrees of data integrity.

Conclusions

Introducing physiological monitoring with mHealth tools into low-resource settings can deliver simple and effective sensing technologies to improve objectivity of health assessments but faces challenges on multiple levels. The target environment, appropriateness of the technology, the skills and motivation of the user, as well as the subject engagements influence the implementation of mHealth solutions alike. With our newly developed methodology and its derived checklist, we enable project implementers to follow a structured evaluation protocol, identify potential risks, and reevaluate challenges during implementation. Such a systematic evaluation of challenges could also be applied and adapted to other areas in the rapidly growing digital health field.

Acknowledgments

The authors would like to thank Matthias Hüser who designed the assessment app, translated usability issues into design improvements, and maintained the app throughout the study. Piotr Wilcynski prepared an earlier analysis of loss of data integrity based on real-time monitoring. The authors also thank the local health district *Red Salud* and health care centers, the families, and stakeholders, as well as the Swiss-Peruvian Health Research Platform and their staff in San Marcos, especially, Angelica Fernandez and Maria Luisa Huyalinos, for coordinating the project. The authors appreciate the great help from Raphaela Graf, Matias Finat, and Serena Haver for translations during interviews; Joanne Lim for proofreading the manuscript; Jenny Brown (JB) and Simon Hofstede (SH) for reviewing; and Dustin Dunsmuir for commenting on the checklist. Afua Adjekum, Dr Daniel Cobos, Dr Kristina Keitel, and Dr Beth Payne kindly accepted to join the expert panel and provided valuable time and feedback to improve the checklist. This study was supported through ETH Global seed funding, the Swiss National Science Foundation (150640), and the UBS Optimus Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured interview questions.

[PDF File (Adobe PDF File), 433KB - mhealth_v6i12e11896_app1.pdf]

Multimedia Appendix 2

Overview of the mobile health (mHealth) projects used for testing and reviewing the checklist.

[PDF File (Adobe PDF File), 341KB - mhealth_v6i12e11896_app2.pdf]

References

1. Peters DH, Garg A, Bloom G, Walker DG, Brieger WR, Rahman MH. Poverty and access to health care in developing countries. Ann N Y Acad Sci 2008;1136:161-171. [doi: 10.1196/annals.1425.011] [Medline: 17954679]
2. 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]

3. Istepanian R, Jovanov E, Zhang YT. Introduction to the special section on m-Health: beyond seamless mobility and global wireless health-care connectivity. IEEE Trans Inf Technol Biomed 2004 Dec;8(4):405-414. [doi: 10.1109/TITB.2004.840019] [Medline: 15615031]
4. World Health Organization. mHealth: New horizons for health through mobile technologies. 2011. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2018-12-04] [WebCite Cache ID 74PildG5j]
5. Poon CC, Zhang YT, Bao SD. A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-Health. IEEE Commun Mag 2006 Apr;44(4):73-81. [doi: 10.1109/MCOM.2006.1632652]
6. Karlen W, Gan H, Chiu M, Dunsmuir D, Zhou G, Dumont GA, et al. Improving the accuracy and efficiency of respiratory rate measurements in children using mobile devices. PLoS One 2014;9(6):e99266 [FREE Full text] [doi: 10.1371/journal.pone.0099266] [Medline: 24919062]
7. Rawstorn JC, Gant N, Warren I, Doughty RN, Lever N, Poppe KK, et al. Measurement and data transmission validity of a multi-biosensor system for real-time remote exercise monitoring among cardiac patients. JMIR Rehabil Assist Technol 2015 Mar 20;2(1):e2 [FREE Full text] [doi: 10.2196/rehab.3633] [Medline: 28582235]
8. Hardinge M, Rutter H, Velardo C, Shah SA, Williams V, Tarassenko L, et al. Using a mobile health application to support self-management in chronic obstructive pulmonary disease: a six-month cohort study. BMC Med Inform Decis Mak 2015 Jun 18;15:46 [FREE Full text] [doi: 10.1186/s12911-015-0171-5] [Medline: 26084626]
9. Coughlin SS. Mobile technology for self-monitoring of blood glucose among patients with type 2 diabetes mellitus. Mhealth 2017;3:47 [FREE Full text] [doi: 10.21037/mhealth.2017.10.03] [Medline: 29184899]
10. Hudson J, Nguku SM, Sleiman J, Karlen W, Dumont GA, Petersen CL, et al. Usability testing of a prototype Phone Oximeter with healthcare providers in high- and low-medical resource environments. Anaesthesia 2012 Sep;67(9):957-967 [FREE Full text] [doi: 10.1111/j.1365-2044.2012.07196.x] [Medline: 22861503]
11. Spence H, Baker K, Wharton-Smith A, Mucunguzi A, Matata L, Habte T, et al. Childhood pneumonia diagnostics: community health workers' and national stakeholders' differing perspectives of new and existing aids. Glob Health Action 2017;10(1):1290340 [FREE Full text] [doi: 10.1080/16549716.2017.1290340] [Medline: 28485694]
12. English LL, Dunsmuir D, Kumbakumba E, Ansermino JM, Larson CP, Lester R, et al. The PAediatric Risk Assessment (PARA) mobile app to reduce postdischarge child mortality: design, usability, and feasibility for health care workers in Uganda. JMIR Mhealth Uhealth 2016 Feb 15;4(1):e16 [FREE Full text] [doi: 10.2196/mhealth.5167] [Medline: 26879041]
13. King C, Boyd N, Walker I, Zadutsa B, Baqui AH, Ahmed S, et al. Opportunities and barriers in paediatric pulse oximetry for pneumonia in low-resource clinical settings: a qualitative evaluation from Malawi and Bangladesh. BMJ Open 2018 Dec 30;8(1):e019177 [FREE Full text] [doi: 10.1136/bmjopen-2017-019177] [Medline: 29382679]
14. Wallis L, Hasselberg M, Barkman C, Bogoch I, Broomhead S, Dumont G, et al. A roadmap for the implementation of mHealth innovations for image-based diagnostic support in clinical and public-health settings: a focus on front-line health workers and health-system organizations. Glob Health Action 2017 Jun;10(sup3):1340254 [FREE Full text] [doi: 10.1080/16549716.2017.1340254] [Medline: 28838310]
15. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
16. Eckman M, Gorski I, Mehta K. Leveraging design thinking to build sustainable mobile health systems. J Med Eng Technol 2016 Aug;40(7-8):422-430. [doi: 10.1080/03091902.2016.1218560] [Medline: 27535325]
17. Boritz JE. IS practitioners' views on core concepts of information integrity. Int J Account Inform Syst 2005 Dec;6(4):260-279. [doi: 10.1016/j.accinf.2005.07.001]
18. Orphanidou C. Signal Quality Assessment in Physiological Monitoring: State of the Art and Practical Considerations. Switzerland: Springer International Publishing; 2018.
19. International Standard Organisation. IEC 62366-1:2015 Medical devices -- Part 1: Application of usability engineering to medical devices. Geneva, Switzerland: IEC; 2015.
20. Karlen W. Data Quality in mHealth. 2015 Presented at: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 25-29, 2015; Milan. [doi: 10.3929/ethz-a-010566532]
21. Hartinger SM, Nuño N, Hattendorf J, Verastegui H, Ortiz M, Mäusezahl D. A factorial cluster-randomised controlled trial combining home-environmental and early child development interventions to improve child health and development: rationale, trial design and baseline findings. BioRxiv 2018; preprint 2018. [doi: 10.1101/465856]
22. Petersen CL, Gorges M, Dunsmuir D, Ansermino M, Dumont GA. Experience report: functional programming of mHealth applications. In: Proceedings of the 18th ACM SIGPLAN international conference on Functional programming. 2013 Presented at: ICFP'13; September 25-27, 2013; Boston, Massachusetts, USA p. 357-362. [doi: 10.1145/2500365.2500615]
23. Karlen W, Kobayashi K, Ansermino JM, Dumont GA. Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation. Physiol Meas 2012 Oct;33(10):1617-1629. [doi: 10.1088/0967-3334/33/10/1617] [Medline: 22986287]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

25. Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans Inf Technol Biomed 2006 Jan;10(1):156-167. [Medline: 16445260]
26. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215 [FREE Full text] [doi: 10.1177/1932296814526495] [Medline: 24876569]
27. Ettinger KM, Pharaoh H, Buckman RY, Conradie H, Karlen W. Building quality mHealth for low resource settings. J Med Eng Technol 2016;40(7-8):431-443. [doi: 10.1080/03091902.2016.1213906] [Medline: 27560713]
28. Dunsmuir DT, Payne BA, Cloete G, Petersen CL, Görges M, Lim J, et al. Development of mHealth applications for pre-eclampsia triage. IEEE J Biomed Health Inform 2014 Nov;18(6):1857-1864. [doi: 10.1109/JBHI.2014.2301156] [Medline: 25375683]
29. Lim J, Cloete G, Dunsmuir DT, Payne BA, Scheffer C, von Dadelszen P, et al. Usability and feasibility of PIERS on the Move: an mHealth app for pre-eclampsia triage. JMIR Mhealth Uhealth 2015;3(2):e37 [FREE Full text] [doi: 10.2196/mhealth.3942] [Medline: 25887292]
30. Dada OA, Odubena O, Adepoju AA, Vidler M, Orenuga EA, Osiberu B, et al. 3 Development and testing of pictograms for the symptoms of pre-eclampsia in Ogun State, Nigeria: preeclampsia in low and middle income countries. Pregnancy Hypertens 2016 Jul;6(3):179. [doi: 10.1016/j.preghy.2016.08.085]
31. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu LM, 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 09;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
32. Raihana S, Dunsmuir D, Huda T, Zhou G, Rahman QS, Garde A, et al. Development and internal validation of a predictive model including pulse oximetry for hospitalization of under-five children in Bangladesh. PLoS One 2015;10(11):e0143213 [FREE Full text] [doi: 10.1371/journal.pone.0143213] [Medline: 26580403]
33. Zhang J, Karlen W. ETH Zurich Sensor Based Digital Health Checklist. 2018. URL: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/302878/checklistV1.0.pdf?sequence=3&isAllowed=y
34. Prue CS, Shannon KL, Khyang J, Edwards LJ, Ahmed S, Ram M, et al. Mobile phones improve case detection and management of malaria in rural Bangladesh. Malar J 2013;12:48 [FREE Full text] [doi: 10.1186/1475-2875-12-48] [Medline: 23374585]
35. Medhanyie AA, Little A, Yebyo H, Spigt M, Tadesse K, Blanco R, et al. Health workers' experiences, barriers, preferences and motivating factors in using mHealth forms in Ethiopia. Hum Resour Health 2015 Jan 15;13:2 [FREE Full text] [doi: 10.1186/1478-4491-13-2] [Medline: 25588973]
36. 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]
37. 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]
38. Mwendwa P. What encourages community health workers to use mobile technologies for health interventions? Emerging lessons from rural Rwanda. Dev Policy Rev 2017;36(1):111-129. [doi: 10.1111/dpr.12275]
39. Graham HR, Bakare AA, Gray A, Ayede AI, Qazi S, McPake B, et al. Adoption of paediatric and neonatal pulse oximetry by 12 hospitals in Nigeria: a mixed-methods realist evaluation. BMJ Glob Health 2018 Jun;3(3):e000812 [FREE Full text] [doi: 10.1136/bmjgh-2018-000812] [Medline: 29989086]
40. Walker IA, Reshamwalla S, Wilson IH. Surgical safety checklists: do they improve outcomes? Br J Anaesth 2012 Jul;109(1):47-54 [FREE Full text] [doi: 10.1093/bja/aes175] [Medline: 22649183]
41. World Health Organization, Department of Reproductive Health and Research, ExpandNet. Beginning with the end in mind: planning pilot projects and other programmatic research for successful scaling up. Geneva, Switzerland: World Health Organization; 2011.
42. 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;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
43. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth assessment: conceptualization of a global framework. JMIR Mhealth Uhealth 2017 May 02;5(5):e60 [FREE Full text] [doi: 10.2196/mhealth.7291] [Medline: 28465282]
44. World Health Organization. Monitoring and Evaluating Digital Health Interventions: A practical guide to conducting research and assessment. 2016. URL: http://www.who.int/reproductivehealth/publications/mhealth/digital-health-interventions/en/ [accessed 2018-09-27] [WebCite Cache ID 72kQtDedn]
45. van Klei WA, Hoff RG, van Aarnhem EE, Simmermacher RK, Regli LP, Kappen TH, et al. Effects of the introduction of the WHO “surgical safety checklist” on in-hospital mortality: A cohort study. Ann Surg 2012 Jan;255(1):44-49. [doi: 10.1097/SLA.0b013e31823779ae] [Medline: 22123159]

46. Lundin J, Dumont G. Medical mobile technologies - what is needed for a sustainable and scalable implementation on a global scale? Glob Health Action 2017 Jun;10(sup3):1344046 [FREE Full text] [doi: 10.1080/16549716.2017.1344046] [Medline: 28838308]

Abbreviations

CRF: case report form
FW: field worker
HR: heart rate
mHealth: mobile health
PPG: photoplethymogram
RR: respiratory rate
SpO_2: peripheral capillary oxygen saturation
SQI: signal quality index

Edited by G Eysenbach; submitted 27.09.18; peer-reviewed by C King, S Cartledge, S Shah; comments to author 20.10.18; revised version received 14.11.18; accepted 22.11.18; published 14.12.18.

Please cite as:
Zhang J, Tüshaus L, Nuño Martínez N, Moreo M, Verastegui H, Hartinger SM, Mäusezahl D, Karlen W
Data Integrity–Based Methodology and Checklist for Identifying Implementation Risks of Physiological Sensing in Mobile Health Projects: Quantitative and Qualitative Analysis
JMIR Mhealth Uhealth 2018;6(12):e11896
URL: http://mhealth.jmir.org/2018/12/e11896/
doi:10.2196/11896
PMID:30552079

Original Paper

Technology Adoption, Motivational Aspects, and Privacy Concerns of Wearables in the German Running Community: Field Study

Martin Wiesner[1,2*], Dipl-Inform Med; Richard Zowalla[1,2], MSc; Julian Suleder[2,3], MSc; Maximilian Westers[1], MSc; Monika Pobiruchin[2,4*], Dr Sc Hum

[1]Department of Medical Informatics, Heilbronn University, Heilbronn, Germany

[2]Consumer Health Informatics Special Interest Group, German Association for Medical Informatics, Biometry and Epidemiology, Cologne, Germany

[3]ERNW Research GmbH, Heidelberg, Germany

[4]GECKO Institute for Medicine, Informatics and Economics, Heilbronn University, Heilbronn, Germany

[*]these authors contributed equally

Corresponding Author:
Martin Wiesner, Dipl-Inform Med
Department of Medical Informatics
Heilbronn University
Max-Planck-Straße 39
Heilbronn, D-74081
Germany
Phone: 49 71315046947
Email: martin.wiesner@hs-heilbronn.de

Abstract

Background: Despite the availability of a great variety of consumer-oriented wearable devices, perceived usefulness, user satisfaction, and privacy concerns have not been fully investigated in the field of wearable applications. It is not clear why healthy, active citizens equip themselves with wearable technology for running activities, and what privacy and data sharing features might influence their individual decisions.

Objective: The primary aim of the study was to shed light on motivational and privacy aspects of wearable technology used by healthy, active citizens. A secondary aim was to reevaluate smart technology adoption within the running community in Germany in 2017 and to compare it with the results of other studies and our own study from 2016.

Methods: A questionnaire was designed to assess what wearable technology is used by runners of different ages and sex. Data on motivational factors were also collected. The survey was conducted at a regional road race event in May 2017, paperless via a self-implemented app. The demographic parameters of the sample cohort were compared with the event's official starter list. In addition, the validation included comparison with demographic parameters of the largest German running events in Berlin, Hamburg, and Frankfurt/Main. Binary logistic regression analysis was used to investigate whether age, sex, or course distance were associated with device use. The same method was applied to analyze whether a runner's age was predictive of privacy concerns, openness to voluntary data sharing, and level of trust in one's own body for runners not using wearables (ie, technological assistance considered unnecessary in this group).

Results: A total of 845 questionnaires were collected. Use of technology for activity monitoring during events or training was prevalent (73.0%, 617/845) in this group. Male long-distance runners and runners in younger age groups (30-39 years: odds ratio [OR] 2.357, 95% CI 1.378-4.115; 40-49 years: OR 1.485, 95% CI 0.920-2.403) were more likely to use tracking devices, with ages 16 to 29 years as the reference group (OR 1). Where wearable technology was used, 42.0% (259/617) stated that they were not concerned if data might be shared by a device vendor without their consent. By contrast, 35.0% (216/617) of the participants would not accept this. In the case of voluntary sharing, runners preferred to exchange tracked data with friends (51.7%, 319/617), family members (43.4%, 268/617), or a physician (32.3%, 199/617). A large proportion (68.0%, 155/228) of runners not using technology stated that they preferred to trust what their own body was telling them rather than trust a device or an app (50-59 years: $P<.001$; 60-69 years: $P=.008$).

Conclusions: A total of 136 distinct devices by 23 vendors or manufacturers and 17 running apps were identified. Out of 4, 3 runners (76.8%, 474/617) always trusted in the data tracked by their personal device. Data privacy concerns do, however, exist

in the German running community, especially for older age groups (30-39 years: OR 1.041, 95% CI 0.371-0.905; 40-49 years: OR 1.421, 95% CI 0.813-2.506; 50-59 years: OR 2.076, 95% CI 1.813-3.686; 60-69 years: OR 2.394, 95% CI 0.957-6.183).

(JMIR Mhealth Uhealth 2018;6(12):e201) doi:10.2196/mhealth.9623

KEYWORDS

athlete; wearables; mobile phones; physical activity; activity monitoring

Introduction

Overview

Running has become one of the most popular exercise activities in western countries [1]. Technologically inclined runners find a great variety of wearable devices for the purpose of activity monitoring [2-5]. Research on device or app adoption and reasons for their use by runners seems to be underexplored in the literature, as reported by Evenson et al [6]. Motivation to either use or not use a wearable device for activity tracking is important, for example, for device vendors or health insurance companies to adjust their product strategies or incentive programs. For example, personal motivational factors can include "(a) Seeing if I met my goal (movement/sleep/calories count); (b) Look and feel good, improve mood and avoid sitting; and (c) Getting tips and recommendations" [7]. Other factors such as "technical failure or other technical problems, including empty batteries" can lead to negative experiences resulting in nonuse of wearables [8], and other reasons besides technology failure could also be a reason to abandon a particular device [9]. Despite the importance of understanding these factors, perceived usefulness, user satisfaction, and privacy concerns are under-investigated in the emerging field of consumer-centric mobile health (mHealth) applications. In this context, wearable technology such as Global Positioning System (GPS)–enabled sports watches and activity trackers is identified as a key trend in 2017 according to the worldwide survey of fitness trends [10,11]. However, little is known with respect to individual perceptions of the implications of activity data collection and possible sharing for a broad population of these technology users.

Related Work

A number of studies in a variety of different settings have been performed. These include several studies that have investigated the accuracy of commercially available wearable devices, mostly in laboratory settings, for example, treadmill experiments [12-14]. These studies mainly focused on technical features and capabilities of the devices and results from small sample sizes and homogenous cohorts, that is, younger and active males have been reported.

A study conducted by Kaewkannate and Kim compared "the accuracy of four wearable devices in conjunction with user friendliness and satisfaction" [15], using a small cohort size (n=7), including 6 healthy male participants, and all participants being graduate students.

By contrast, Mercer et al focused on older adults living with chronic illnesses [16]. They applied a mixed-methods approach to study the usability and usefulness of wearable activity trackers. The authors found "wearable activity trackers are perceived as useful and acceptable" for adults aged over 50 years. A different study examined the "Feasibility of Fitness Tracking with Urban Youth" with a body mass index of 23 or higher [17]. The findings indicate that "wearable devices alone are not sufficient to support significant changes in existing physical activity practices" for users (n=24) in younger age groups. Nevertheless, feasibility studies indicate that "monitor comfort and design and feedback features [are] important factors to children and adolescents" [18].

Another study assessed the "acceptance and usage of wearable activity trackers in Canadian community-dwelling older adults" in a crossover design study [19]. For 20 adults, aged 55 years and older (mean 64 years), 2 wearable devices were given to participants who then rated different aspects of the devices and their use after 21 days of use. The authors report that "privacy was less of concern for older adults, but it may have stemmed from a lack of understanding of the privacy risks and implications."

In other research, however, privacy seems to be an important aspect for users of wearable devices or apps [20]. Other researchers have found that "individuals' decisions to adopt healthcare wearable devices are determined by their risk-benefit analyses" [21]. The authors concluded that "individuals' perceived privacy risk is formed by health information sensitivity, personal innovativeness, legislative protection, and perceived prestige." Their findings suggest that consumers' motivations and buying decisions are "determined by [an individual] risk-benefit assessment."

A review paper on ethical implications of user perceptions concluded that "wearable device users are highly concerned regarding privacy issues and consider informed consent as 'very important' when sharing information with third parties." [22]. An explorative study including 82 participants investigated "privacy concerns and sensitivity regarding data gathered with wearables" [23]. The authors reported "that the participants would prefer to keep said data to themselves. Furthermore, user factors such as age, gender, and privacy behavior could not be identified as having an effect on sharing said data." Yet, it remains an open question whether these findings are applicable to a broad and heterogeneous population, for example, a running community at a road running event.

Alley et al determined "people's current use, interest and preferences for advanced [pedometer] trackers" via a cross-sectional Australia-wide telephone survey [24]. The authors found that 31% of the participants "considered counting steps the most important function and 30% regarded accuracy as the most important characteristic." About half of the participants were hesitant toward using current activity tracking devices or expressed individual skepticism. According to this

survey [24], the main reasons "for not wanting to use a tracker were, 'I don't think it would help me' (39%), and 'I don't want to increase my activity' (47%)." It is not clear whether these findings can be confirmed in similar study settings in other countries.

Aims of the Study

This study investigated several aspects of how citizens use smart technology for exercise activities. It was a follow-up study of previous research [4]. The primary study aim of the 2017 field study was to examine (a) reasons for use of wearable technology and (b) privacy concerns associated with the use of wearable technology. A secondary aim was (c) to study the current smart technology adoption within the running community in Germany and (d) to compare it with previous results from 2016.

This field study contributes to the mHealth field as it presents findings that originate from a real-world assessment and not from a potentially biased laboratory setting. In this context, the study cohort comprises participants from a public "Sport for All" road running event, that is, primarily physically active and healthy citizens of both sexes and all adult age groups (>16 years).

Methods

Study Design

The cross-sectional study consisted of 2 parts: a pre- and a postrace survey. The prerace survey aimed to answer research questions (a) to (c). It was scheduled for the registration period, that is, the day before and during the morning hours of the race day while runners picked up their number bibs and timing chips at the event site. For the postrace survey, the plan was to acquire data in the finisher area of the running event. Interviewer staff had the task of asking runners for individual step counts or the tracked distance in kilometers. As runners were quite exhausted after the race, no questions on motivational aspects, concerns, or willingness to share data with others could be posed at that time. At no point in time were individual, participant-related data, that is, name or address, collected.

Study Setting—Road Running Event

The study was conducted during the 17th Heilbronner Trollinger Marathon on 6th to 7th May, 2017. The Trollinger Marathon is an annual road running event located in southern Germany [25]. In 2017, according to the official starter list [26], 6397 runners lined up for 4 different running courses: (1) full marathon, that is, 42.195 km, (2) half-marathon, that is, 21.0975 km, (3) walking/Nordic walking course with a length of 14.4 km, and (4) a marathon relay of approximately 3×14 km.

The event is part of the German Road Races Society calendar. Full and half-marathon courses conform to Association of International Marathons and Road Races (AIMS) and International Association of Athletics Federation (IAAF) regulations as both event categories are precisely measured by an accredited AIMS/IAAF Grade A or B measurer.

The prerace survey took place on May 6 (11:45 am to 5:30 pm) and May 7 (7:00 am to 10:00 am). Study staff were divided into several shifts, and it was ensured that at least two interviewers were present during registration hours. Interviewers were instructed to select runners randomly. Only runners older than the minimum participation age (ie, 16 years) were included in the cohort. Participation in the study was voluntary, that is, registered runners were asked whether they wanted to participate in a survey on wearable technology.

Due to bad weather conditions and heavy rainfall on May 7, the authors decided to cancel the postrace survey in the finisher area of the *Heilbronn Frankenstadion*. The main reason was that runners left the stadium quickly after crossing the finish line to escape the weather conditions and thus chances of acquiring a reasonable amount of study data were low.

Questionnaire and Survey App

Participants were very focused on the registration and picking up their individual number bib, especially in the morning hours directly before the race. Therefore, questionnaires were designed in a compact and brief format.

Informed by the experiences from 2016, 2 questionnaires were designed:

1. Prerace questionnaire (Q_1) contained items on (1) tracking devices used, (2) demographic data, for example, age and sex, (3) the running course chosen, (4) reasons for device or nondevice usage (eg, trust in own body or technical barriers), (5) parameters checked, (6) validity of collected or displayed data, (7) concerns regarding data privacy, and (8) voluntary data sharing,
2. Postrace questionnaire (Q_2) to determine the accuracy of the tracked distances.

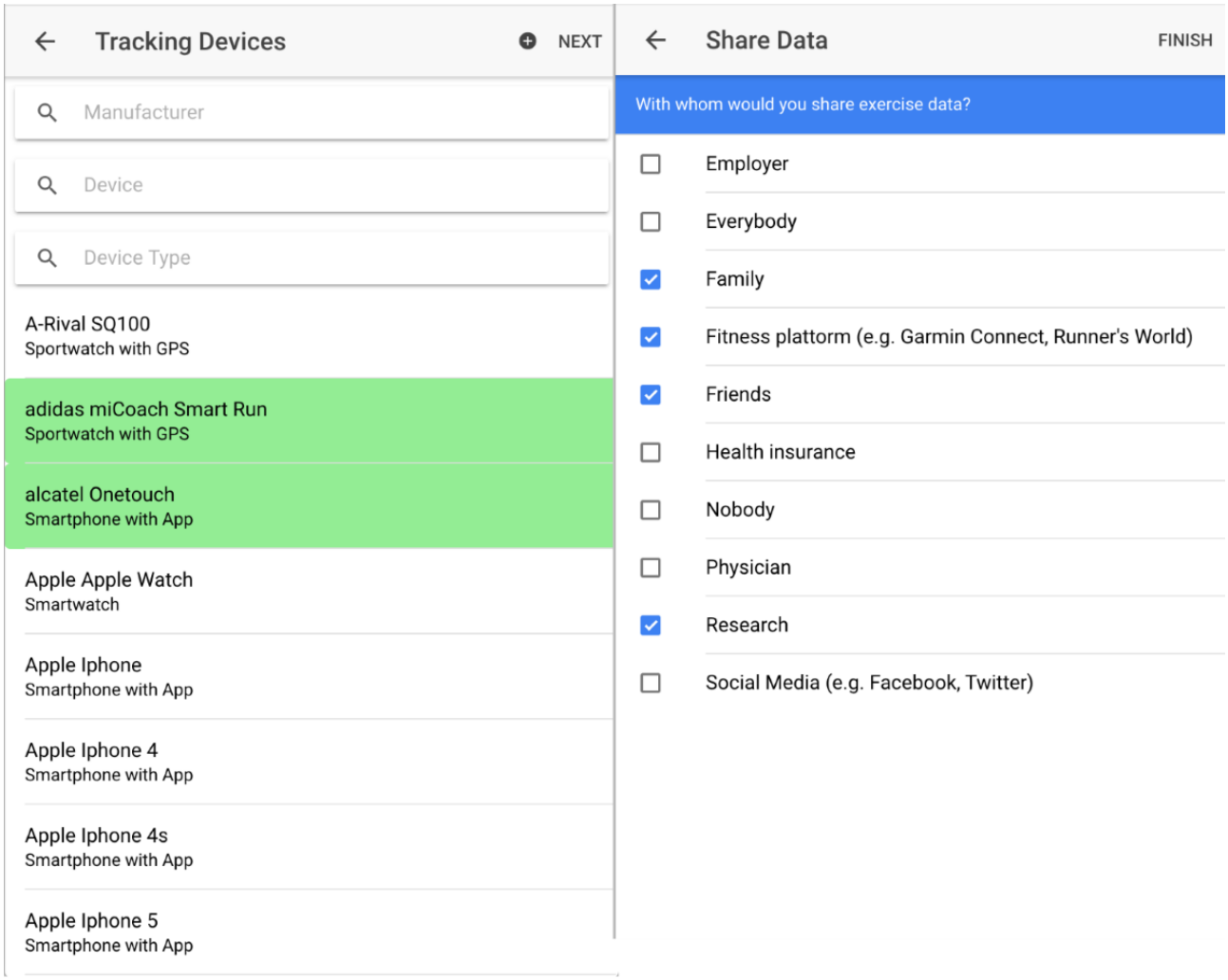
The items in Q_1:

- were derived from existing literature; for items (4) and (5), see [24]; for item (6), see [27],
- have been posed to runners in the previous edition in 2016; for items (1) to (3), see [4], or
- have been raised by some study participants themselves during the previous study.

The developed questionnaire was tested with several staff members of nonresearch departments in our institution. We thereby checked the understandability and if it was feasible to conduct interviews with it.

All questionnaire items in Q_2 were the same as in the previous study in 2016 to allow a comparison of the results in both years. English translations of Q_1 and Q_2 can be found in the Multimedia Appendices 1 and 2. An original German version of both questionnaires can be found as Multimedia Appendices 3 and 4.

Figure 1. Selected screenshots of the Trolli survey app. Left: device selection from the list of all available devices and companion apps; right: preferences on sharing exercise data.



A relational database schema (see Multimedia Appendix 5) was derived from each questionnaire. The survey database—collected in 2016—on wearable devices and apps (156 distinct devices, 25 running apps, and 36 different vendors) was updated before the actual running event, and comprised 199 devices, 35 apps, and 37 vendors ahead of completion of the 2017 survey. A related Trolli survey app was implemented with the Ionic framework (in version 2.2.2) [28], depicted in Figure 1.

The app was deployed to the mobile phones (iOS: n=5, Android: n=10) of the interviewer staff and to 2 extra Android tablets. The completed questionnaires were stored on the internal disk of the mobile device and then synchronized via a *RESTful* Web service connected to the study database. If interviewers identified a previously unknown device, the creation of an entry for a new device was possible. Thereby, other interviewers could make use of it during subsequent interviews.

Both questionnaires (Q_1 and Q_2) were implemented electronically within this survey app. Neither personal data nor contact details of survey participants were collected. Therefore, the resulting records can be considered as an anonymized dataset that does not conflict with European, national, or federal data privacy laws. In case of technical issues such as power loss, loss of network connectivity, etc, paper-based backup copies of Q_1 and Q_2 were available on-site.

In 2016, during the transcription of the paper-based format, several questionnaires had to be excluded from the evaluation. In such cases, either (1) interviewers forgot to complete the questionnaire, (2) the handwriting of an interviewer was illegible, or (3) questionnaires provided nonspecific vendor or device information, for example, “a sports watch I bought at a supermarket”, and thus had to be excluded, which corresponded to a dropout rate of 7.79% (98/1258).

By contrast, in 2017, with support of a survey app, no incomplete questionnaires were encountered. As a consequence, no data had to be excluded from the later analysis. This corresponds to a dropout rate of 0.0% (0/845). Moreover, 28 new devices were captured on-site during the interviews.

Statistical Analysis

The representativeness of the study cohort was analyzed via a comparison of age distributions of the full-marathon, half-marathon, and walking/Nordic walking against the official starter list—as provided by the event organizer in Heilbronn and the events in the German cities: Berlin, Frankfurt/Main, and Hamburg. No details on sex were available for relay runners

in the Heilbronn starter list. Therefore, this group (n=41) was excluded from the analysis of representativeness. Chi-square analysis was applied to test whether the given age distribution of the Heilbronn starter list matched that of the study cohort (H_0: distributions are equal, H_1: distributions differ). If $P \geq .05$, H_0 was accepted.

Important aspects for technology acceptance are trust in data and the protection of privacy [6,29,30]. Hence, these factors are of particular interest in the context of this study. Binary logistic regression was applied to analyze whether a runner's age is a predictor for 3 factors: (1) trust in one's own body, (2) privacy concerns, and (3) openness to sharing data. The same method was used to examine whether sex, age, or running course are predictors for wearable device usage.

Data were analyzed with the statistics software *R* [31] in version 3.3.3 (2017-03-06) on a Windows 10 Enterprise LTSB 2016/ 64-bit computer.

Results

Principal Findings

A total of 845 questionnaires were collected via our survey app and stored in the database. After verification of on-site data entries of previously unidentified manufacturers or devices (n=28), all 845 entries were included for further analyses. A comprehensive list of devices and manufacturers can be found in Multimedia Appendix 6.

Study Cohort

The official starter list of the marathon (full and half) and walking/Nordic walking course comprised 6327 men and women. Male runners dominated the starter field for both full- and half-marathon, especially in the full-marathon starter field (82.6%, 514/622). However, more female runners (73.1%, 742/1015) were registered for the walking/Nordic walking course, as listed in Table 1.

Our study cohort covered 13.2% (845/6397) of the registered runners. Likewise, the sample covered 13.0% (611/4689) of the registered half-marathon runners. Chi-square analysis revealed that age distributions are not similar for this subcohort ($P<.001$). Nearly a quarter of all marathon runners were interviewed (23.5%, 146/622), which was representative for this subgroup ($P=.55$). However, our study cohort underrepresented walkers ($P<.001$), for which only 4.6% (47/1015) were included.

The age distribution of the Heilbronner Trollinger Marathon resembles those of larger running events, for example, the Berlin marathon (n=33,248 finishers in 2017 [32]), Berlin half-marathon (n=23,957 finishers in 2016 [33]), Hamburg marathon (n=11,930 finishers in 2017 [34]), Hamburg half-marathon (n=8299 finishers in 2017 [35]), Frankfurt/Main marathon (n=11,121 finishers in 2017 [36]), and Frankfurt/Main half-marathon (n=4558 finishers in 2018 [37]). Table 2 compares the age distributions of the aforementioned events with the starter list in Heilbronn 2017 on the basis of the respective finisher lists.

Motivational Aspects

Usage and Nonusage

Runners who declared that they used one or more devices (73.0%, 617/845) for training or during running events were asked to select one or more reasons for doing so (see Multimedia Appendix 1). As presented in Table 3, nearly 9 out of 10 runners (89.8%, 554/617) used wearables as a tool for exercise control. For at least a third of the participants (34.0%, 210/617), technology was used for self-motivational reasons and as an enabler to *get more active* in general. Only a very small percentage (1.0%, 6/617) of the runners mentioned that they used a wearable device sponsored by their health insurance company/sickness fund or as recommended by their physician. The most common reason for technology use was to monitor exercise. This was consistent for runners across all age groups: 87.3% to 100% (see Table 3, Q_1, No. 7; Multimedia Appendix 1).

If runners used a wearable device, several activity parameters were checked (see Figure 2). On average, 4 parameters were checked. The most frequent parameters were distance covered, time, and average speed. This corresponds to the most common reasons for using a device, that is, monitoring exercise (compare answers for No. 7 in Table 3). Monitoring of the hydration parameter seemed negligible.

Nontechnology users were asked why they did not use wearable devices. More than two-thirds (68.0%, 155/228) answered that they listen to their own body' instead of technology. Technical barriers with wearables were reported by 12.7% (29/228); and 3.0% (7/228) had encountered bad experiences in the past.

As presented in Table 4, older age of nontechnology runners is associated with higher trust in one's own body as the individual's 'measuring instrument'. This is a statistically significant finding for older age groups, that is, 50-59 years ($P<.001$) and 60-69 years ($P=.008$).

Validity and Data Sharing

Validity of Collected or Displayed Data

We asked runners if they trusted the data captured by their wearable device. Three out of 4 participants (76.8%, 474/617) stated that they always trusted the data. A fifth of participants (127/617) considered the visualized data as *partly* valid, whereas 1.8% (11/617) did not trust the data gathered at all.

Table 1. Distributions of sex and age groups (Pr_{survey}) among runners for the full- and half-marathon and walking or Nordic walking. $Pr_{official}$ denotes the proportion as given in the official starter list for the respective subcohort. $Pr_{official}$ data were published by the event organizer of the Trollinger Marathon only as rounded percentage values, so precise n values for male and female age groups are unavailable. Furthermore, n=41 relay runners and runners with unknown course type were excluded.

Age groups per running course	Male (n)	Pr_{survey} (%)	$Pr_{official}$ (%)	Female (n)	Pr_{survey} (%)	$Pr_{official}$ (%)
Marathon	**N=121**			**N=25**		
16-29	12	9.9	8.4	5	20	13.0
30-39	17	14.1	20.2	7	28	17.6
40-49	42	34.7	30.5	7	28	29.6
50-59	37	30.6	29.6	4	16	31.5
60-69	12	9.9	10.1	2	8	8.3
70-79	1	0.8	1.2	0	0	0
80+	0	0	0	0	0	0
Unknown	0	0	0	0	0	0
Half-marathon	**N=400**			**N=211**		
16-29	72	18.0	21.7	64	30.3	29.8
30-39	89	22.3	27.5	43	20.4	28.1
40-49	93	23.3	24.0	50	23.7	21.6
50-59	113	28.3	20.4	41	19.4	16.8
60-69	26	6.5	5.5	13	6.2	3.7
70-79	6	1.5	0.8	0	0	0.1
80+	1	0.3	0.03	0	0	0
Unknown	0	0	0	0	0	0
Walking or Nordic walking	**N=12**			**N=35**		
16-29	0	0	24.6	2	6	21.1
30-39	0	0	25.0	3	9	23.3
40-49	2	17	18.8	11	31	30.5
50-59	5	42	16.9	11	31	16.2
60-69	1	8	11.4	7	20	3.4
70-79	4	33	3.3	1	3	0.5
80+	0	0	0	0	0	0
Unknown	0	0	0	0	0	0

Table 2. Age distributions of finishers at Berlin marathon 2017, Hamburg marathon 2017, Frankfurt/Main marathon 2017, Berlin half-marathon 2016, Hamburg half-marathon 2017, Frankfurt/Main half-marathon 2018 compared with registered runners in Heilbronn 2017.

Age groups per running course	Pr_{Berlin}, n (%)	$Pr_{Hamburg}$, n (%)	$Pr_{Frankfurt}$, n (%)	$Pr_{Heilbronn}$, n (%)
Marathon	**N=33,248**	**N=11,930**	**N=11,121**	**N=622**
16-29	3950 (11.88)	1439 (12.06)	1425 (12.81)	57 (9.2)
30-39	5221 (15.70)	3058 (25.63)	3128 (28.13)	123 (19.8)
40-49	13,397 (40.29)	3886 (32.41)	3594 (32.32)	189 (30.4)
50-59	8607 (25.89)	2880 (24.15)	2388 (21.47)	186 (29.9)
60-69	1839 (5.53)	613 (5.14)	520 (4.68)	61 (9.8)
70+	234 (0.70)	74 (0.62)	66 (0.59)	6 (1.0)
Half-marathon	**N=23,957**	**N=8299**	**N=4553**	**N=4689**
16-29	4259 (17.78)	2241 (27.00)	795 (17.46)	1120 (23.89)
30-39	6960 (29.05)	2726 (32.85)	1466 (32.20)	1299 (27.70)
40-49	6561 (27.39)	1934 (23.30)	1231 (27.04)	1095 (23.35)
50-59	4878 (20.36)	1138 (13.71)	873 (19.17)	909 (19.39)
60-69	1124 (4.69)	225 (2.71)	167 (3.68)	236 (5.03)
70+	175 (0.73)	35 (0.42)	20 (0.44)	30 (0.64)

Table 3. Answers given for selected questions (prerace questionnaire, Q_1) on (non) motivation (No. 7, No. 6), privacy (No. 10), and data sharing (No. 11) by age group. Values in round brackets represent the proportion of runners who answered this question in the respective age group.

Answer options	Age groups (years), n (%)						Total
	16-29	30-39	40-49	50-59	60-69	70+	
Users with device Q_1, No. 7[a], n=617[b]	**N=192**	**N=245**	**N=268**	**N=262**	**N=47**	**N=9**	**N=1023**
Gift	5 (4.2)	7 (4.8)	9 (5.4)	11 (7.4)	1 (3.1)	0 (0.0)	33
Incentive program by health insurance	0 (0.0)	0 (0.0)	2 (1.2)	2 (1.3)	1 (3.1)	0 (0.0)	5
Recommendation by physician/general practitioner	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.7)	0 (0.0)	0 (0.0)	1
Health aspects	12 (10.2)	13 (9.0)	20 (12.0)	33 (22.1)	5 (15.6)	1 (14.3)	84
Self-motivation	50 (42.4)	57 (39.3)	54 (32.5)	44 (29.5)	5 (15.6)	0 (0.0)	210
Curiosity	19 (16.1)	27 (18.6)	25 (15.1)	21 (14.1)	4 (12.5)	1 (14.3)	97
Exercise control	103 (87.3)	131 (90.3)	151 (91.0)	132 (88.6)	30 (93.8)	7 (100.0)	554
Trend setter	0 (0.0)	2 (1.4)	3 (1.8)	3 (2.0)	0 (0.0)	0 (0.0)	8
Other	3 (2.5)	8 (5.5)	4 (2.4)	15 (10.1)	1 (3.1)	0 (0.0)	31
Users without device Q_1, No. 6[a], n=228[b]	**N=57**	**N=27**	**N=56**	**N=77**	**N=39**	**N=10**	**N=266**
Costs	5 (10.2)	1 (4.0)	0 (0.0)	2 (2.9)	0 (0.0)	0 (0.0)	8
Lack of trust	1 (2.0)	1 (4.0)	4 (8.7)	4 (5.9)	1 (3.0)	1 (14.3)	12
Bad experiences	4 (8.2)	1 (4.0)	2 (4.3)	0 (0.0)	0 (0.0)	0 (0.0)	7
Technical barriers	4 (8.2)	1 (4.0)	9 (19.6)	7 (10.3)	5 (15.2)	3 (42.9)	29
I trust my body	24 (49.0)	17 (68.0)	30 (65.2)	55 (80.9)	26 (78.8)	3 (42.9)	155
Other	13 (26.5)	3 (12.0)	10 (21.7)	7 (10.3)	5 (15.2)	3 (42.9)	41
Don’t know	6 (12.2)	2 (8.0)	1 (2.2)	1 (1.5)	1 (3.0)	0 (0.0)	11
Not stated	0 (0.0)	1 (4.0)	0 (0.0)	1 (1.5)	1 (3.0)	0 (0.0)	3
Privacy concern Q_1, No. 10, n=617[b]	**N=118**	**N=145**	**N=166**	**N=149**	**N=32**	**N=7**	**N=617**
Yes	31 (26.3)	42 (29.0)	59 (35.5)	68 (45.6)	14 (43.8)	2 (28.6)	216
No	53 (44.9)	69 (47.6)	71 (42.8)	56 (37.6)	10 (31.3)	0 (0.0)	259
Doesn’t matter	25 (21.2)	27 (18.6)	25 (15.1)	18 (12.1)	5 (15.6)	2 (28.6)	102
Don’t know	9 (7.6)	7 (4.8)	11 (6.6)	7 (4.7)	3 (9.4)	3 (42.9)	40
Data sharing Q_1, No. 11[a], n=617[b]	**N=273**	**N=306**	**N=325**	**N=257**	**N=58**	**N=8**	**N=1227**
Employer	3 (2.5)	3 (2.1)	2 (1.2)	1 (0.7)	0 (0.0)	0 (0.0)	9
Physician	44 (37.3)	45 (31.0)	51 (30.7)	46 (30.9)	11 (34.4)	2 (28.6)	199
Family	60 (50.8)	67 (46.2)	75 (45.2)	54 (36.2)	12 (37.5)	0 (0.0)	268
Fitness platform	17 (14.4)	22 (15.2)	23 (13.9)	12 (8.1)	3 (9.4)	0 (0.0)	77
Research	17 (14.4)	25 (17.2)	17 (10.2)	19 (12.8)	9 (28.1)	1 (14.3)	88
Friends	78 (66.1)	83 (57.2)	88 (53.0)	58 (38.9)	9 (28.1)	3 (42.9)	319
Health insurance	23 (19.5)	17 (11.7)	17 (10.2)	14 (9.4)	3 (9.4)	0 (0.0)	74
Social media	5 (4.2)	10 (6.9)	7 (4.2)	5 (3.4)	2 (6.3)	0 (0.0)	29
Everybody	14 (11.9)	10 (6.9)	11 (6.6)	6 (4.0)	0 (0.0)	0 (0.0)	41
Nobody	12 (10.2)	24 (16.6)	34 (20.5)	42 (28.2)	9 (28.1)	2 (28.6)	123

[a]Denotes Questions in Q_1 which allowed multiple answers.

[b]Number of runners, rather than number of responses, that is, can have multiple responses per runner.

Figure 2. Answers given to Q1, No. 8 (n=2473) for each parameter (multiple answers possible). Numbers on the y-axis represent answers given per parameter; percentages at the top of each bar correspond to the relative proportion of runners (n=617) who selected one or more activity parameters.

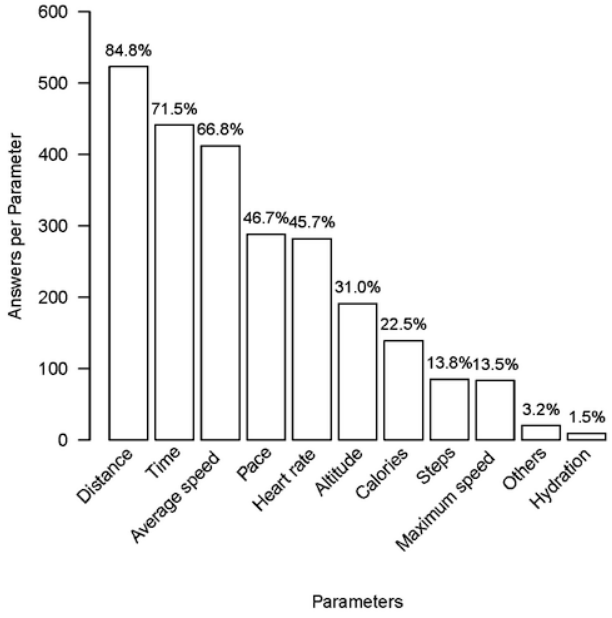


Table 4. Binary logistic regression analysis of the parameter age for the factors: (1) Trust in own body (n=221, excluded: 7 participants in the age group 70 to 79 years), (2) Data sharing (n=617), and (3) Privacy concerns (n=474, excluded: "Doesn't matter", "Don't know", and 1 participant in the age group 70-79 years).

Model	Odds ratio (95% CI)	*P* value
Trust in own body, age range (years)		
16-29 (Ref[a])	1.0	—[b]
30-39	2.214 (0.824-6.321)	.12
40-49	1.953 (0.862-4.523)	.11
50-59	4.407 (1.966-10.301)	<.001
60-69	3.869 (1.470-11.203)	.09
70-79	—	—
Data sharing, age range (years)		
16-29 (Ref)	1.0	—
30-39	0.570 (0.264-1.176)	.14
40-49	0.440 (0.209-0.869)	.02
50-59	0.288 (0.138-0.562)	<.001
60-69	0.289 (0.109-0.783)	.01
70-79	0.283 (0.054-2.123)	.16
Privacy concerns, age range (years)		
16-29 (Ref)	1.0	—
30-39	1.041 (0.371-0.905)	.89
40-49	1.421 (0.813-2.506)	.22
50-59	2.076 (1.183-3.686)	.01
60-69	2.394 (0.957-6.183)	.06
70-79	—	—

[a]Ref: Reference group in the respective regression model.

[b]Not applicable.

Privacy

Users of smart technology were asked about their views and opinions on nonvoluntary sharing of training or exercise data. In particular, this question included health insurance companies or device vendors, which could share such data for commercial or other purposes. As presented in Table 3, 2 out of 5 runners (42.0%, 259/617) stated that they would not be concerned if data were shared in such a manner. By contrast, 35.0% (216/617) said that they would not accept a vendor sharing data without their consent. Out of 617, 102 (16.5%) participants had a neutral perspective ("Doesn't matter") and only a small fraction of runners (6.5%, 40/617) were undecided ("Don't know"). A detailed analysis revealed that runners in older age groups considered *privacy* a more important aspect of activity monitoring technology than younger users, see Table 4. This is a statistically significant finding for runners in the age group of 50 to 59 years (P=.01).

Data Sharing

In addition, we asked runners with whom they would share their personal training data on a voluntary basis. According to the results in Table 3, most participants that used technology preferred to share data only with their friends or family members (51.7%, 319/617 and 43.4%, 268/617). Nearly a third of the participants were open to sharing data with a physician (32.3%, 199/617). The public sharing of training data on social media platforms—for example, Twitter or Facebook—was only selected by 4.7% of the participants (29/617). One in every 7 runners (88/617, 14.3%) was open to sharing data for research purposes. Table 4 shows that voluntary data sharing with any other parties—that is, family, physician, employer, etc—decreases with higher age groups: 40 to 49 years (P=.02), 50 to 59 years (P<.001), and 60-69 years (P=.01).

Device Categories

Devices used by runners were classified into 6 device categories: (D_1) smartphones with related app, (D_2) GPS-enabled sports watches, (D_3) sports watches without GPS support, only heart rate monitors, (D_4) smart watches, (D_5) wristband activity trackers, and (D_6) other devices. These are the same categories as in the 2016 study; hence, comparison with previous results can be made.

The results presented in Table 5 reveal that 228 out of 845 (27.0%) runners did not use a device. This represents a slightly larger proportion when compared with the results from 2016.

The most popular device was the Polar M400 (7.5%, 66/881). The GPS-enabled sports watch segment (D_2) was dominated by the vendors Garmin (49.0%, 192/392) and Polar (31.9%, 125/392). If runners used their smartphone with a companion app, most of them preferred a model sold by Apple (49.4%, 78/158). The most popular running app was Runtastic/Runtastic Pro (63.9%, 101/158) followed by Nike+ Run Club, Strava, Sports Tracker and several other apps. These findings were very similar to the results from 2016 (see Table 2 in [4]). Devices in the categories D_4 and D_5 were found more frequently than in the previous year. However, these categories accounted for a smaller share than other device categories.

Adoption of Wearable Devices

Regression analysis showed that use of wearable devices was associated with runners of younger age groups, see Table 6. This finding was statistically significant for 30 to 39 years age group (P=.002). Older age groups were less likely to use such devices. However, this finding was only significant for the age group of 60 to 69 years (P=.005). Being a participant of the walking/Nordic walking course was predictive of using no technology when compared with the reference group of half-marathon runners (P=.005). Marathon and relay runners were more likely to use wearable devices (odds ratio [OR] 1.368 and OR 1.458). Moreover, female runners seemed not to rely on technology (OR 0.745), although these findings were not statistically significant.

Table 5. Devices (D) used by category in 2017 compared with 2016 (n=653 devices used by 845 runners). Values in brackets denote the relative proportion of each category. Note: some runners (4.2%, 36/845) used more than one device.

Category	2017 (N=881), n (%)	2016 (N=978), n (%)
D_1–Smartphone and app	158 (24.2)	181 (24.4)
D_2–GPS[a]-equipped sports watch	392 (60.0)	437 (58.8)
D_3–Heart rate monitor	25 (3.8)	37 (5.0)
D_4–Smart watch	22 (3.4)	14 (1.9)
D_5–Wristband activity tracker	33 (5.1)	28 (3.6)
D_6–Other devices	23 (3.5)	47 (6.3)
No device	228 (27.0)	234 (26.1)

[a]GPS: Global Positioning System.

Table 6. Binary logistic regression of sex, age, and course type for the dependent variable "wearable device use" (n=845).

Feature	Odds ratio (95% CI)	*P* value
Sex		
Male (Ref[a])	1.0	—[b]
Female	0.745 (0.528-1.054)	.09
Age (years)		
16-29 (Ref)	1.0	—
30-39	2.357 (1.378-4.115)	.002
40-49	1.485 (0.920-2.403)	.11
50-59	0.904 (0.572-1.424)	.67
60-69	0.417 (0.226-0.765)	.01
70-79	0.637 (0.188-2.243)	.47
80+	<0.001	.98
Course type		
Half-marathon (Ref)	1.0	—
Marathon	1.368 (0.875-2.191)	.18
Marathon relay	1.458 (0.875-2.191)	.51
Walking or Nordic walking	0.391 (0.202-0.751)	.01
Unknown	0.725 (0.280-2.032)	.52

[a]Ref: Reference group in the regression model.

[b]Not applicable.

Discussion

Principal Findings

One aim of the study was to gain insights into the reasons for use and privacy concerns of healthy active citizens with regard to wearable devices. The literature [19,21,23] yields an unclear picture whether privacy is a concern for using tracking technology. The results in this paper confirmed the plurality of opinions on data privacy and voluntary sharing aspects present in a heterogeneous population. Approximately 35% of runners raised concerns regarding whether vendors would share or sell their individual tracking data to third parties without explicit consent. By contrast, approximately 42% were not concerned and almost 17% did not care at all. Runners in older age groups considered privacy to be more important (50-59 years: *P*=.01) than in younger age groups. This is in line with our findings on voluntary data sharing (see Table 3), that is, openness to sharing activity data with family members, friends, and physicians, decreased for older age groups.

Our findings revealed that the primary reason for technology use is to monitor exercise levels (approximately 89.8%), followed by self-motivation (34.0%), curiosity (15.7%), and personal health aspects (13.6%). The main reason for using no technology at all was that runners prefer to "listen to their own body" (68.0%). The analysis showed that there are significant differences between age groups: when compared with runners in the 16 to 29 years age group, runners in the 50 to 59 years age group (*P*<.001) and 60 to 69 years age group (*P*=.008) had higher trust in listening to their own body feedback.

With respect to the second aim of the study, the analysis of adoption rates of wearables showed that 3 out of 4 runners used tracking technology. This is in line with our findings from the previous edition of the Trollinger Marathon study (see [4]). Most runners preferred to use a GPS-enabled sports watch (D_2: 60.0%), followed by mobile phones with apps (D_1: 24.2%). Smart watches (D_4: 3.4%) and wristband activity trackers (D_5: 5.1%) were less frequently used even though their relative share increased slightly compared with 2016. Overall, 76.8% of these runners stated that they always trust the tracking data of their personal device.

Limitations

The data were collected through a cross-sectional survey, which may be subject to bias. For this reason, several limitations apply.

The cohorts for marathon and half-marathon runners were samples of randomly chosen registrants of the Trollinger Marathon. The age distribution of the event in Heilbronn was similar to those in Berlin, Hamburg and Frankfurt/Main. Age and sex distributions of the study sample were similar to the proportions published in the official starter list. However, Chi-square analysis revealed that only the marathon subcohort can be considered as representative for the respective group. Although statistical tests indicated no representativeness for the other subcohorts, we consider our data to be a valid sample, at least for the running community in (southwestern) Germany.

As in 2016, the response rate for the (Nordic) walking event was quite low (n=47, with a total of 1015 registered participants). The major part of the walkers were employees of

the main sponsor of the event, and the handing out of number bibs for those participants was only conducted on May 7 and at a different location. More interviewer staff would have been necessary to cover this separate location, which was not feasible.

Questionnaires used in this survey were developed by the authors. Items in Q_1 were either used in [4] or have been raised during interviews by study participants of the previous year (2016) or were derived from existing literature [24,27]. Therefore, we assume that Q_1 achieves at least a moderate level of content-related validity. However, no evaluations on construct- or criterion-related validity and/or reliability were conducted, which poses a limitation for this study.

Due to heavy rainfall in the Heilbronn region on May 7, no postrace survey (Q_2) on wearables' accuracy could be conducted by our interviewer staff. Therefore, no results on the tracked course distances can be reported in the 2017 edition of the Trollinger Marathon study.

Potential Pitfalls

In 2017, the deployment and use of our Trolli survey app at the event site helped to prevent or reduce (1) capturing data manually, (2) questionnaire transcription errors, (3) incomplete questionnaires, and (4) increase the postrace data analysis efficiency. Moreover, less time was needed to train the interviewing staff, and the handover between interviewer shifts was more streamlined, as the app could be preinstalled and tested individually. However, the mobile phones of some interviewers were outdated, which meant extra effort was required to set them up for the interviews on-site.

During crowded times in the registration area, the on-site cellular network was not able to handle all connection attempts initiated by numerous runners and our interviewer team. For such a scenario, unsent survey records were stored locally, and a built-in app feature allowed interviewers to resend those records to the study's Web service. Unfortunately, for a small number of survey records, some interviewers initiated the transmission of a record multiple times, thereby skipping the server-side (asynchronous) response receipt. The resulting duplicates had to be identified and cleared afterward with the help of (1) server log files, (2) screenshots of related survey smartphones, and/or (3) time stamps available in the survey database. Checks for duplicates and stricter confirmation mechanisms were missing during the interview phase and could have prevented these issues.

Comparison With Prior Work

The principal findings of the Trollinger Marathon of 2016 [4] on device adoption rates, usage of specific device categories, and the most popular devices and apps were replicated in 2017. In both editions, the results of binary logistic regression analysis support that younger age, male sex, and choice of long-distance running course are predictive of using technology in running activities.

Several studies on user acceptance of wearables exist [16,17,19,29,38-40]. However, these studies are based on surveys of between 16 and 260 participants that have a specific demographic background, for example, older adults or adolescents. By contrast, our study cohort included 845 users and nonusers of smart technology of both sexes and distributed over almost all age groups. The study revealed that runners without a device represent 27.0% (228/845) of all runners of which 12.7% (29/228) stated technical barriers as the primary reason for not using a wearable. When compared with the nonuser proportion caused by technological barriers (17.5%) reported in the study by Hermsen et al [8], the aforementioned fraction is slightly lower. The reason might be that the Trollinger study participants were given more answer options, in particular, "I trust my body", "Lack of trust," and the more general option "Bad experiences" (see Table 3, Q_1, No. 6). Our results suggest that runners who do not use a device instead *use* their body to gather feedback (68.0%, 155/228).

A survey conducted by Deloitte among 2000 Germans in 2016 found that more than half of the participants (55%) were willing to share health data with a general practitioner [27]. By contrast, only 32.3% of the technology equipped runners of the sampled cohort were open to sharing exercise data with a physician. According to the Deloitte survey, a small fraction was open to sharing data with either device manufacturers (7%) or other internet companies (7%). This is comparable with our findings: fitness platforms (12.5%) or social media (4.7%) are channels to which users would upload their data.

Puri et al reported for 20 elderly people in Canada that "privacy was less of concern of older adults" and linked this to a potential "lack of understanding" [19]. This is not in line with our findings as runners of older age groups expressed privacy concerns more frequently than younger participants. We assume that runners who use wearable technology have at least a basic understanding on data collected and potential risks, supported by Huckvale et al [20]. However, a substantial fraction of the device users in our sample did not have concerns or reported that it did not matter (combined 58.5%) if their activity data were shared or sold by vendors to third parties. This particular finding slightly disagrees with the results of the study by Lidynia et al on privacy concerns. The authors report (n=82 German citizens—36 device users, 46 nonusers) "that the participants would prefer to keep said data to themselves." [23]. According to their study, runners of older age groups are more hesitant to share data publicly, for example, via social media. This can be supported by our findings.

In a cross-sectional Australia-wide telephone survey (n=1257), Alley et al found that the "use of advanced trackers compared with pedometers was higher in males [...] and younger participants" [24]. The results of our logistic regression analysis on the same factors is in line with the findings from Australia.

The Online Eindhoven Running Survey 2014 (ERS14) comprised 2172 participants of the Half Marathon in Eindhoven [41]. In terms of adoption rates, the results were similar to our findings. In the ERS14 study, more than 86% of the participants used such a device, whereas 73% used at least one in the Trollinger study cohort. In the study by Janssen et al, the sports watch segment was also dominated by Garmin. However, a larger proportion of ERS14 participants used mobile phones in combination with apps (54.9%), which is not supported by our results (24.2%, see Table 5). This difference is likely to originate

from the fact that the ERS14 study was conducted as an online survey and not as a field study at the site of an actual running event. Janssen et al reported age as a predictor for app and sports watch usage. Our results are in line with this finding: younger and middle-aged runners (16-29 years, 30-39 years, and 40-49 years) are more likely to use monitoring devices than runners in older age groups (50-59 years, 60-69 years, and 70-79 years).

Becker et al chose a qualitative approach with semistructured interviews (n=16) on factors influencing "continuous use of fitness trackers" [38]. They reported on perceived benefit, perceived privacy, perceived deficiency, and related subthemes. The interviews lasted 25 min on average. By contrast, for field studies—such as in our setting—and with a limited amount of time per interview, merely predefined answer options and short questionnaires were applicable. From a methodological perspective, it is important to keep in mind that runners are not willing to participate in long interviews on-site.

Future Directions

Given that certain outcomes of the ERS14 study differ from those of our field study, it could be interesting for other researchers to replicate the design of this study in a different setting; that is, to analyze running communities in other countries. Moreover, a detailed study of female long-distance runners could provide new insights into their preferences toward tracking technology. Therefore, we encourage other international research groups to support the idea of interviewing runners at the site of an actual running event.

Alternatively, future surveys could investigate whether runners have changed their lifestyle, diet, or activity patterns as they started using a device for activity tracking. This could shed light on open questions, for example, "does wearable technology have long-term health effects besides being a stylish gadget?" or even "should health insurance companies promote wearable devices via incentive programs"?

Moreover, a question remains whether there is a link between *active* citizens and the openness toward collecting data and sharing it with others. Our study provides a first indication that both positions—openness and being concerned—exist within the German running community. These findings might, however, be different in other countries or communities.

In addition, it is still unclear whether citizens use wearable technology for other reasons than sport. This could, for instance, include monitoring of sleep, hydration, or blood glucose levels. In this context, other motivational patterns could be present and interesting for future studies.

Conclusions

Use of technology for training or during running events is prevalent in the running community (approximately 73% in Southwestern Germany). Male long-distance runners and runners of younger age groups are more likely to use wearable devices. In total, 136 distinct devices by 23 vendors and 17 running apps were identified.

As expected, runners use wearable technology primarily for monitoring personal exercise levels (90%). The second most prevalent reason is self-motivation (34%), which is more important for younger runners. External incentives or recommendations are of marginal importance (<1%).

Three out of 4 runners always trusted the data tracked by their personal device. Two out of 5 runners (42%) explained that they were not concerned whether data collected by their device might be shared without explicit consent. By contrast, 35% said that they would not accept a vendor sharing data with third parties for commercial purposes. In the case of voluntary data sharing, runners preferred to give it to friends (52%), family members (43%), or a physician (32%), whereas only a small fraction (<2%) would give these data to their employer.

A large proportion (68%) of runners not using technology stated that they preferred to trust in the feedback from their own body. Approximately 11% answered that they experienced technical barriers when using wearables.

Future research might focus on preferences of female runners as they seem to be less likely (OR 0.745) to use tracking technology for running. Whether women perceive wearables and potential benefits differently from men remains a question open for research.

Acknowledgments

The authors would like to thank the following persons for conducting interviews: K Blesch, L Brandl, T Eichner, J Fischer, R Kappler, S Koch, W Schramm, S Schweizer, A Speer, and I Steinmetz. The authors would like to thank the Heilbronn Marketing GmbH—the organizer of the Trollinger Marathon—for their support. The authors would also like to thank Dr Caroline S Clarke for her constructive criticism and proofreading of the manuscript. Moreover, the authors explicitly thank Nasse-Design KG, Daniel Nasse [42] for contributing the article's TOC image and granting publishing permissions.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Prerace questionnaire Q1.

[PDF File (Adobe PDF File), 49KB - mhealth_v6i12e201_app1.pdf]

Multimedia Appendix 2

Postrace Questionnaire Q2.

[PDF File (Adobe PDF File), 41KB - mhealth_v6i12e201_app2.pdf]

Multimedia Appendix 3

Prerace questionnaire Q1 in the original German version.

[PDF File (Adobe PDF File), 54KB - mhealth_v6i12e201_app3.pdf]

Multimedia Appendix 4

Postrace questionnaire Q2 in the original German version.

[PDF File (Adobe PDF File), 44KB - mhealth_v6i12e201_app4.pdf]

Multimedia Appendix 5

Relational database schema of the Trolli 2017 study database.

[TXT File, 9KB - mhealth_v6i12e201_app5.txt]

Multimedia Appendix 6

Device categories, vendors, models, and apps.

[PDF File (Adobe PDF File), 52KB - mhealth_v6i12e201_app6.pdf]

References

1. Scheerder J, Breedveld K, Borgers J. Who Is Doing a Run with the Running Boom? In: Scheerder J, Breedveld K, Borgers J, editors. Running Across Europe - The Rise and Size of One of the Largest Sport Markets. London: Palgrave Macmillan; 2015:1-27.
2. Swan M. Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0. JSAN 2012 Nov 08;1(3):217-253. [doi: 10.3390/jsan1030217]
3. McCallum C, Rooksby J, Gray CM. Evaluating the impact of physical activity apps and wearables: interdisciplinary review. JMIR Mhealth Uhealth 2018 Mar 23;6(3):e58 [FREE Full text] [doi: 10.2196/mhealth.9054] [Medline: 29572200]
4. Pobiruchin M, Suleder J, Zowalla R, Wiesner M. Accuracy and adoption of wearable technology used by active citizens: a marathon event field study. JMIR Mhealth Uhealth 2017 Feb 28;5(2):e24 [FREE Full text] [doi: 10.2196/mhealth.6395] [Medline: 28246070]
5. Soutiere SE, Cox BD, Laird MD, Markwald RR, Heaney JH, Chinoy ED, et al. Defense Technical Information Center. San Diego, CA, United States: Naval Health Research Center San Diego; 2017. Wearable Activity Tracker Literature Review (January 2009 - July 2016) URL: http://www.dtic.mil/docs/citations/AD1041181 [accessed 2017-12-07] [WebCite Cache ID 6vWfYnooA]
6. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015;12(1):159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
7. Asimakopoulos S, Asimakopoulos G, Spillers F. Motivation and user engagement in fitness tracking: heuristics for mobile healthcare wearables. Informatics 2017 Jan 22;4(1):5. [doi: 10.3390/informatics4010005]
8. Hermsen S, Moons J, Kerkhof P, Wiekens C, De Groot M. Determinants for sustained use of an activity tracker: observational study. JMIR Mhealth Uhealth 2017 Oct 30;5(10):e164 [FREE Full text] [doi: 10.2196/mhealth.7311] [Medline: 29084709]
9. Clawson J, Pater J, Miller A, Mynatt E, Mamykina L. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. New York, NY, USA: ACM; 2015 Presented at: UbiComp '15 Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka, Japan p. 647-658 URL: http://ubicomp.org/ubicomp2015/index.html
10. Thompson W. Worldwide Survey of fitness trends for 2017. ACSMs Health Fit J 2016 Nov;20(6):8-17. [doi: 10.1249/FIT.0000000000000252]
11. Thompson WR. Worldwide survey of fitness trends for 2018: the CREP edition. ACSMs Health Fit J 2017 Nov;21(6):10-19. [doi: 10.1249/FIT.0000000000000341]
12. Wen D, Zhang X, Liu X, Lei J. Evaluating the consistency of current mainstream wearable devices in health monitoring: a comparison under free-living conditions. J Med Internet Res 2017 Mar 07;19(3):e68 [FREE Full text] [doi: 10.2196/jmir.6874] [Medline: 28270382]

13. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
14. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015 Oct;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
15. Kaewkannate K, Kim S. A comparison of wearable fitness devices. BMC Public Health 2016 May 24;16:433 [FREE Full text] [doi: 10.1186/s12889-016-3059-0] [Medline: 27220855]
16. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
17. Schaefer SE, Ching CC, Breen H, German JB. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. Am J Health Educ 2016 Jan 08;47(1):8-16. [doi: 10.1080/19325037.2015.1111174]
18. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
19. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
20. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
21. Li H, Wu J, Gao Y, Shi Y. Examining individuals' adoption of healthcare wearable devices: an empirical study from privacy calculus perspective. Int J Med Inform 2016 Apr;88:8-17. [doi: 10.1016/j.ijmedinf.2015.12.010] [Medline: 26878757]
22. Segura Anaya L, Alsadoon A, Costadopoulos N, Prasad PW. Ethical implications of user perceptions of wearable devices. Sci Eng Ethics 2017 Feb 02;24(1):1-28. [doi: 10.1007/s11948-017-9872-8] [Medline: 28155094]
23. Lidynia C, Brauner P, Ziefle M. A Step in the Right Direction - Understanding Privacy Concerns and Perceived Sensitivity of Fitness Trackers. In: Advances in Human Factors in Wearable Technologies and Game Design. AHFE 2017: Advances in Human Factors in Wearable Technologies and Game Design; 2017 Jul Presented at: AHFE 2017; July 17-21, 2017; Los Angeles, California, USA p. 42-53. [doi: 10.1007/978-3-319-60639-2_5]
24. Alley S, Schoeppe S, Guertler D, Jennings C, Duncan MJ, Vandelanotte C. Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. BMJ Open 2016 Dec 07;6(7):e011243 [FREE Full text] [doi: 10.1136/bmjopen-2016-011243] [Medline: 27388359]
25. Württembergischer Leichtathletik-Verband. Heilbronn Trollinger Marathon. 2017. Heilbronner Trollinger Marathon und Halbmarathon URL: http://www.trollinger-marathon.de/en/hauptmenue-links/home.html [accessed 2017-12-08] [WebCite Cache ID 6j5ssBLmp]
26. Heilbronner Trollinger-Marathon. Bergisch Gladbach, Germany: mika:timing GmbH Starterliste URL: http://trollinger.r.mikatiming.de/2017/?pid=startlist [accessed 2018-11-09] [WebCite Cache ID 6vUA7B3h8]
27. Deloitte Touche Tohmatsu Limited. Statista. 2016. Mit wem würden Sie mobil erhobene Gesundheitsdaten (z.B. Herzfrequenz, Blutzucker) teilen? [With whom would you share collected health data?] URL: https://de.statista.com/statistik/daten/studie/712642/umfrage/umfrage-zur-teilungsbereitschaft-von-gesundheitsdaten-in-deutschland/ [accessed 2017-12-08] [WebCite Cache ID 6vYQYicZv]
28. Ionic. 2017. Build Amazing Native Apps and Progressive Web Apps with Ionic Framework and Angular URL: https://ionicframework.com/ [accessed 2017-12-07] [WebCite Cache ID 6vWhE05ga]
29. Marakhimov A, Joo J. Consumer adaptation and infusion of wearable devices for healthcare. Computers in Human Behavior 2017 Nov;76:135-148. [doi: 10.1016/j.chb.2017.07.016]
30. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
31. R-Project. 2017. The R Project for Statistical Computing URL: http://www.R-project.org/ [accessed 2017-10-03] [WebCite Cache ID 6tvdqQY7e]
32. SCC Events. 2017. 44. BMW Berlin-Marathon - Results URL: http://results.scc-events.com/2017/?pid=list [accessed 2017-12-04] [WebCite Cache ID 6vSUxJGDm]
33. SCC Events. 2016. 36. Berliner Halbmarathon - Results URL: http://results.berliner-halbmarathon.de/2016/?page=5&pid=list [accessed 2017-12-04] [WebCite Cache ID 6vSW4fIgC]
34. Haspa Marathon Hamburg. Bergisch Gladbach, Germany: mika:timing GmbH; 2017. Haspa Hamburg Marathon - Results URL: http://hamburg.r.mikatiming.de/2017/?pid=list [accessed 2018-03-26] [WebCite Cache ID 6yCtJnkhu]
35. Hella Hamburg Halbmarathon. 2017. 23. Hella Halbmarathon Hamburg - Results URL: https://www.hamburg-halbmarathon.de/ergebnisse/ [accessed 2018-03-26] [WebCite Cache ID 6yCtOgPhM]
36. Mainova Frankfurt Marathon. 2017. Frankfurter Mainova Marathon - Results URL: http://live.frankfurt-marathon.com/2017/?pid=list [accessed 2018-03-26] [WebCite Cache ID 6yCtSaULe]

37. 16th Frankfurt Mainova Half Marathon. Bergisch Gladbach, Germany: mika:timing GmbH; 2018. 16. Frankfurter Mainova Halbmarathon - Results URL: http://frankfurter-hm.r.mikatiming.de/2018/ [accessed 2018-03-26] [WebCite Cache ID 6yCtWC1jc]
38. Becker M, Kolbeck A, Matt D, Hess T. Understanding the Continuous Use of Fitness Trackers: A Thematic Analysis. In: Proceedings of the PACIS 2017.: AIS eLibrary; 2017 Presented at: 21st Pacific Asia Conference on Information Systems (PACIS); July 16-20, 2017; Langkawi Island, Malaysia p. 40 URL: http://aisel.aisnet.org/pacis2017/40
39. Yang H, Yu J, Zo H, Choi M. User acceptance of wearable devices: an extended perspective of perceived value. Telematics Informatics 2016 May;33(2):256-269. [doi: 10.1016/j.tele.2015.08.007]
40. Fritz T, Huang E, Murphy G, Zimmermann T. A Study of Long-Term Use of Activity Sensing Devices for Fitness. New York, NY, USA: ACM; 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 26-May 01, 2014; Toronto, Ontario, Canada p. 487-496. [doi: 10.1145/2556288.2557383]
41. Janssen M, Scheerder J, Thibaut E, Brombacher A, Vos S. Who uses running apps and sports watches? Determinants and consumer profiles of event runners' usage of running-related smartphone applications and sports watches. PLoS One 2017;12(7):e0181167 [FREE Full text] [doi: 10.1371/journal.pone.0181167] [Medline: 28732074]
42. Nasse Design. Werbeagentur Heilbronn URL: http://www.nasse-design.de

Abbreviations

AIMS: Association of International Marathons and Road Races
ERS14: Eindhoven Running Survey 2014
GPS: Global Positioning System
IAAF: International Association of Athletics Federation
mHealth: mobile health
OR: odds ratio
Q_1: prerace questionnaire
Q_2: postrace questionnaire

Edited by G Eysenbach; submitted 11.12.17; peer-reviewed by P Boisrond, S Delfino Barboza, K Ng; comments to author 18.03.18; revised version received 05.04.18; accepted 29.09.18; published 14.12.18.

Please cite as:
Wiesner M, Zowalla R, Suleder J, Westers M, Pobiruchin M
Technology Adoption, Motivational Aspects, and Privacy Concerns of Wearables in the German Running Community: Field Study
JMIR Mhealth Uhealth 2018;6(12):e201
URL: http://mhealth.jmir.org/2018/12/e201/
doi:10.2196/mhealth.9623
PMID:30552085

Original Paper

An Activity Tracker and Its Accompanying App as a Motivator for Increased Exercise and Better Sleeping Habits for Youths in Need of Social Care: Field Study

Kari Rönkkö[1], Docent

Department of Design, Faculty of Business, Kristianstad University, Kristianstad, Sweden

Corresponding Author:
Kari Rönkkö, Docent
Department of Design
Faculty of Business
Kristianstad University
Elmetorpsvägen 15
Kristianstad, SE-291 88
Sweden
Phone: 46 0442503192
Email: kari.ronkko@hkr.se

Abstract

Background: The number of mobile self-tracking devices connected to the Web has exploded in today's society. With these wearable activity trackers related to Web 2.0 apps and social media have come new ways of monitoring, measuring, representing, and sharing experiences of the human body. New opportunities related to health and new areas of implementation for professionals have appeared, and one identified area that can benefit from mobile health technologies is social work.

Objective: There are still only a small number of papers reporting the results from studying wearable activity trackers and accompanying apps in the context of agency-based social work. This study aimed to contribute to the identified shortage by presenting results from a research project framed by the following overarching question: *What effects will the studied youths in need of social care experience in relation to exercise and sleep as the result of using a wearable activity tracker and its accompanying app?*

Methods: A field study framed by action research was performed. The study concerned vulnerable youths living in a Swedish municipality's care and accommodation home that tried out an activity tracker and its accompanying app.

Results: The results from the study confirm previously published research results reporting that instant graphical feedback, sharing information, and being part of a social community can have a positive impact on lifestyle changes. In addition, this study's main results are that (1) the most important factor for positive health-related lifestyle changes was the establishment of personal long-term goals and (2) professional social workers found the studied technology to function as a valuable counseling tool, opening up avenues for lifestyle talks that otherwise were hard to undertake.

Conclusions: This study demonstrates how an activity tracker and its accompanying app can open up a topic for discussion regarding how vulnerable youths can achieve digital support for changing unhealthy lifestyle patterns, and it shows that the technology might be a valuable counseling tool for professionals in social work.

(JMIR Mhealth Uhealth 2018;6(12):e193) doi:10.2196/mhealth.9286

KEYWORDS

mHealth; social work; youths; activity trackers; mobile applications; motivation; self-care; sleep hygiene; goals

Introduction

Over the last 10 years, mobile devices have become widely adopted in society. Today, small, portable digital devices are widely spread and able to connect remotely to the internet from most locations. Within the practice of medicine and public health supported by mobile devices, so-called mobile health (mHealth), there follow many promising developments. The option to use mobile technology to collect data on one's bodily functions and everyday activities has received much attention [1-4]. Frequent statements have also been made in the popular media and in the medical and public health literature about a revolution in health

care, preventive medicine, and public health driven by the use of mobile technologies [5]. Mobile technologies have the capacity to extend the body by supplying data that can be used to display both limits and capabilities, and they allow users to employ the data to work upon and improve themselves. Accounts of self-tracking technologies for health tend to place emphasis on the potential for the empowerment of laypeople and on the importance of taking responsibility for one's own health [5].

Previous studies have demonstrated that various technologies for health-promoting behavior allowing the user to measure and monitor their own behavior are promising [6-8]. It is also claimed that the concepts of health and health care are moving toward the notion of personalized preventive health and that personalized solutions could be the answer to solving public health challenges at their causal root [1]. It is clear that regular doses of physical movement have a positive influence on people's health and also that mHealth technology can help motivate people to move more. In relation to exercise by prescription, researchers have asked for more empirical studies on how apps can provide motivation [9]:

> *I think we have reached the next generation of studies now. They should be about how we get this to work for prescribers and patients. Perhaps patients can get the motivation, support, and monitoring via apps or other technical systems, which gradually give them the results showing that the training is paying off because this is strongly motivating. This is what we should study now when there exists evidence that supports exercise by prescription.*

Obviously, it has been believed for some years now that mHealth technologies have the potential to contribute to improved health through increased motivation and changes in bad habits.

Another interesting area besides exercise for mHealth technologies within social work is to provide feedback on bad sleeping patterns. Good sleeping habits can influence the ability to both learn and perform in society. Insufficient sleep is associated with a large number of problems and adverse health behaviors, that is, physical distress, frequent mental distress, activity limitations, depressive symptoms, anxiety, and pain [10]. Having adequate tools, methods, and knowledge to provide increased awareness of sleeping patterns [11] would benefit both people in need of care and professionals within health care and social work. In a systematic literature review published in 2017, evidence was presented that mobile phone interventions have the ability to attenuate sleep disorders and to enhance sleep quality. The same study also claimed that mobile phone intervention methods could provide better sleep solutions in comparison with other recognized treatments such as cognitive-behavioral therapy for insomnia [12].

The present author published results from a literature study in 2017 covering the state-of-the-art research surrounding motivation, wearable activity trackers, and accompanying apps in social work [13]. The main results are summarized here. On the basis of a bibliometric and literature study, it was demonstrated that there are relatively few research connections between concepts that are related to health apps and self-care versus motivation and social work. It seems that different journals have different interests in activity trackers and accompanying apps, for example, they are interested in the clinical, medical, user-oriented, interaction-oriented, management, or technical aspects of the technology. The strongest identified link-based relations between research areas were found in the borderland of medicine and computer science, which largely represented technical discussions of digital solutions in the health area. Connections to social work appeared to be isolated between psychology, psychiatry, and social psychology and also between systems-oriented computer science and user-oriented informatics. In relation to shared concepts between the research areas, it was found that the main focuses were on technology and technical systems; communication technologies and distance treatment; technical design; technology and medical treatment; data security; and privacy issues from technical, political, and legal points of view. In summary, the bibliometric study demonstrated that there are relatively few research connections between concepts that are related to health apps and self-care, motivation, and social work.

The literature review presented in the same publication [13] showed that health apps as a phenomenon had attracted a wide range of researchers from several disciplines. In general, high hopes were expressed for mHealth technologies, and self-care seems to be a growing area and a target for both health and industrial actors. It was believed that eHealth services could empower users to better manage their own health, and the main identified themes were related to digital technology and weight loss, psychological treatment, rapid technical developments, missing professional health competencies during the development of health technologies, and how to handle user data and integrity. Many of the papers focused on the technology itself rather than motivational work, health aspects, and social work. A systematic literature review by Ridgers et al [14] found that there is a paucity of research concerning the effectiveness and feasibility of activity trackers and accompanying apps as tools for increasing children's and adolescents' physical activity levels. There is a lack of contributions regarding the combination of sleeping habits, motivation, increased exercise, and social work [13]. There is also a limited amount of information about how mHealth technologies can sustain changes in health behavior and how they can be integrated into health care [15]. There seems to be room for more research regarding how to increase nonactive people's physical movement and how to improve their sleeping patterns with the help of activity trackers and accompanying apps in the field of social work.

Social work is lagging behind when it comes to adopting these technologies, and researchers have requested more empirical studies of these technologies in the context of social work [16-18]. Identified reasons for the lagging situation are a lack of technological training in the majority of social work degree programs, a lack of professional standards that define technological competency, and concerns that technology might interfere with the relationships on which social work is based [16-18.]. It has also been shown that the vulnerability of the clients in social work requires special ethical consideration when new digital technology is introduced [19]. The lagging also includes the quantified self, for example, with the help of

wearable activity trackers and accompanying apps [16]. The identified gap in knowledge is claimed to slow down the adoption of professional knowledge related to information technology [18,16,20]. Furthermore, it is suggested that harnessing technology in the context of social work will allow for more effective service development, planning, and delivery [21]. This paper contributes to the above by presenting results from a 4-month field study of how a wearable activity tracker and its accompanying app were perceived by 8 young people living in care and accommodation homes.

Despite the above-identified lack of rigorous research and conclusive results in this study's subject of interest, it was still possible to identify useful lessons that were applicable to this study's area of focus. These included findings that iterative feedback cycles between users (clients), health professionals, and researchers (as inventors) are necessary to successfully address real-world needs [22,23]; that privacy and users' attitudes are challenges that need to be considered early in the development of the technology [24-27]; that setting goals involving wearable activity trackers and their accompanying apps is motivating [6]; that instant gratification and graphical feedback are especially rewarding [7,28]; and that being part of a social community with friendly competition is motivating and might lead to increased physical activity [28].

This paper presents results from a field study framed by action research. The study aimed to better understand how a specific wearable activity tracker and its accompanying app might influence the motivation of vulnerable youths to exercise more and to improve their sleeping habits in the context of agency-based social work. The study's results suggest that social workers' counseling of youths can benefit from shared data from a wearable activity tracker and its accompanying app. The principal result of this paper is increased knowledge of how a wearable activity tracker and its accompanying app can open up for discussions of how activity and sleep patterns among vulnerable youths can be addressed in the context of agency-based social work.

Methods

Field Study Setting

The empirical investigation took place in cooperation with Helsingborg, the eighth largest city in Sweden with more than 100,000 inhabitants in its urban area. Helsingborg's interest in participation in the project was anchored in goal 7 of the city's strategic vision, that is, the city of Helsingborg will be a leader in utilizing digitization opportunities. They expressed 3 main reasons for their involvement in this study: (1) to learn more about new potential digital technology for preventive social work, (2) to identify new tools that could increase motivation for positive lifestyle changes, and (3) to better understand the potential of the *coolness factor* among children and youths. In comparison with previous research projects in the municipality, this project targeted vulnerable groups with a more complex composition of problems because of weakened or obstructed parenting. In relation to the target group, the municipality also expressed fear that the technology might create new types of problems among the youths. One fear was that the youths might feel obligated to participate and thus reveal habits and thoughts about themselves to their peers, which could be used against them. A second fear was that the technology might interfere with the relationships between clients and professionals on which social work is based. Another question they had was regarding how social workers would be able to identify and try out tools for measuring the health of youths with social problems within social care without being health professionals themselves.

In this project, young people living in care and accommodation homes (Home for Care or Living, HVB-home) in Helsingborg municipality were studied. From an organizational viewpoint, the studied HVB homes were under the social administration and targeted young people aged 13 to 20 years with social and psychosocial problems such as relationship problems, school problems, and incipient risk behavior that might lead to crime. Young people with active addiction are not allowed in the HVB homes. The HVB homes are open accommodations where young people are voluntarily placed, and the voluntary nature of the arrangement makes it easier for the staff to work with the youths. In the HVB homes studied for this project, there was room for 8 youths in a shared accommodation, 9 youths in training apartments, and 2 more youths at 2 ordinary homes. The training apartments were spread out in different neighborhoods in Helsingborg city. The joint youth accommodation consisted of a large house where the young people had their own room, including a large common social area and a shared kitchen and dining area.

The wearable activity tracker Jawbone UP24 (Jawbone Limited) was used in the project. The activity tracker synchronized with an app (the accompanying app) that had to be downloaded to the user's own smartphone or tablet. The app recorded the participant's daily physical activity and sleep patterns, which also could be followed over time. The physical activity was measured by counting the number of steps taken, and sleep was measured by the number of slept hours, that is, the number of hours of light and deep sleep. Jawbone UP24 allowed the users to set reminders, individual step and sleep goals, wake-up alarms, and other alarms. In the app, users could also form teams and follow each other's results and communicate in a chat forum. More information about Jawbone UP24 that is beyond the scope of this paper can be found in the study by Swider [29].

Framework for Cooperation and Study

As an overall framework for cooperation and research, the project applied action research [30,31]. This framework allowed us to not only study the social phenomena of common interest but also to jointly alter the object of study in close cooperation cycles. With this approach, we achieved a deeper joint understanding of those aspects that influence the situation of the observed target group at the same time as we could introduce improvements. We implemented a cyclical process of moving from observation to planning, to implementation, and back to observation (various examples of implementations that are beyond the scope of this paper, including some of the author's own previous experiences, can be found in the study by Dittrich et al [32]). The entire project, including planning, bibliometrics and literature study, implementation of the field study, and joint presentation of the results to the municipality, took place in the

period of September 2014 to June 2015. An important goal for the municipality was to achieve hands-on experiences based on academic knowledge. Thus, Helsingborg searched for approaches that could be spread in the municipality's organization and owned by the social workers themselves. A management and cross-municipal unit responsible for developing methods to improve the conditions for children and young people at risk continued the implementations based on this project's results, and when taking a retrospective look at the continued nonacademic implementations by Helsingborg, it was found that the academic study results were confirmed (more information about Helsingborg's continued implementations can be found in the study by Danielsson et al [33]).

Field Study

In the field study, we addressed the following research question: *What effects will the studied youths experience in relation to exercise and sleep as a result of using a wearable activity tracker and its accompanying app?* We chose to focus on interviews influenced by an ethnographic standpoint, meaning that we emphasized understanding the studied people's own point of view [34]. We wanted to know how recent digital technology in the form of an activity tracker and accompanying app could support users' needs and at the same time satisfy the municipality's ambitions regarding social work. The staff at the HVB homes did not want to be responsible for running the project because of previous negative experiences of project ownership. All the young people and staff in the HVB homes were offered the chance to participate. Because this was the first related project in Helsingborg, we could also offer the youths the chance to be involved in the choice of what type of wearable activity tracker and accompanying app should be used in the project. The field study for this group was conducted from November 2014 to February 2015. The period was planned deliberately to enable investigation of whether there might be a difference in young people's physical activity and sleep habits during the Christmas holidays. The youths were offered 2 follow-up meetings that were entirely voluntary, and 2 or 3 youths participated in each meeting. Besides these meetings, there was continuous communication with the youths via individual physical meetings, email, short message service text messages, and phone calls. The staff was followed up through group meetings and individually at the end of the project period.

We are aware that the choice of field methods could lead to an inherent bias in the kind and depth of information that can be obtained from the different communication channels, that is, the participant's personality and willingness to share input. Therefore, we worked hard to establish an atmosphere of mutual trust guided by ethical advice from ethnography [34]. Another potential bias was that the participants on their own initiative decided to introduce a step competition during the second month. They divided themselves into 2 teams and started to compete for a month's time. Hence, it was difficult to know what role the competition as such might have had on the youths' motivation during this month. Was it the competitive game or the technology that provided the motivation? One plausible interpretation is that the technology as such acted as an enabler that inspired them to come up with the idea, allowed them to make their results visible to all the participants in the competition, and provided an option to follow up on each other's results both continuously and historically. Our interpretation was that the technology became a tool for supporting the youths' ideas and desires, where the instant gratification and graphical feedback of achieved results played an important role for the level of motivation.

The field study included 8 youths aged 17 to 18 years and 12 staff, all of whom had access to the Jawbone UP24 wearable activity tracker linked to a smartphone or tablet with the accompanying app installed. The wearable activity tracker was synchronized with the accompanying app that was downloaded to the users' smartphones or tablets. The accompanying app recorded the participants' daily physical activity and sleep patterns, which could be followed over time. The physical activity was measured by counting movement or number of steps, and sleep was measured by counting the number of slept hours divided into hours of light and deep sleep. In the app, the user could set reminders, individual goals, and wake-up alarms. The users could also form teams, follow each other's results, and communicate with invited friends in a chat forum. The app also included a cost function, which allowed the users to record their daily food intake. This nutritional function was not included in the project because of a lack of expertise in the field of nutrition and also the risk of triggering eating disorders. The project had digital access to allow the researchers to follow all included participants' results regarding both daily steps and nightly sleep. This material was regularly compiled into an Excel file that was sent out to all members of the research group and that was taken up and discussed during the regular meetings of the research group.

The professional social caregivers played an important role in the project because they were the ones delivering therapy and advice about life to the youths at the shared accommodation. All caregivers voluntarily decided to use the wearable activity tracker and the accompanying app during the project, but their results from using the technology were not something that we registered. The choice to provide the same technology to the caregivers was important because it increased their commitment and understanding of what was being studied. During the implementation of the project, there were continuous informal meetings and conversation occasions with both the staff and the youths, so-called field observations. Data from many of these occasions confirmed the results from the interviews that we present in the Results section. These conversations and informal meetings were an inevitable and natural part of the project implementation related to the project's follow up of how the technology worked, and they provided understanding of the technology's usage as well as occasions for educating about the available features in the technology.

A 1-hour group interview was conducted with the staff at the youth HVB home, and 4 individual interviews were held with young people living at the same accommodation. We conducted open interviews focusing on the youths' own views and experiences, keeping the purpose of the study in mind during the interviews. Each individual interview lasted around 30 min. One of the youths was interviewed over Skype because of having moved abroad. All interviews were recorded and

transcribed. The selection of the interviewed respondents was based on the youths' availability and their willingness to participate because some of the youths were not comfortable with being formally interviewed. The interviews aimed at gaining a deeper level of knowledge regarding the perceived everyday experiences of using the wearable activity tracker and the accompanying app. A central subject in the interviews was whether and how the young people themselves felt that the activity tracker and accompanying app had influenced their situation regarding physical activity and sleeping patterns. Additional information that is beyond the scope of this paper can be found in the project report by Rönkkö et al [13] (in Swedish).

Results

Field Study Results

During the first 2 months, November and December, all 8 participants were physically active. Up to the Christmas holidays in late December, there were only minor breaks (a few days) during which some of the youths had not used the technology. The reasons for not using the device were because of illness or that they forgot to put it on again after taking a shower or after charging the activity tracker. After the Christmas holidays, only 4 of the participants were still using the activity tracker, and at the end of the project, only 3 were using the activity tracker and accompanying app. Although not all the participants were positive initially, all the interviewed participants expressed that they experienced increased motivation to exercise as a result of using the activity tracker and accompanying app. Table 1 shows the number of steps that the participants took on average per day during the entire study. Days when no data were recorded were not included and therefore did not affect the results.

Instant gratification and graphical feedback of specified goals were reported to be important, and the possibility to set up individual targets in the accompanying app together with instant gratification and graphical feedback of results motivated the participants to increase their physical activity. The average number of steps was consistently a bit over 10,000 steps per day, which can be considered good because this corresponds to about 8 km daily walking, a result one would expect from active people. Some comments in this regard were:

> *I'm proud of myself when I see that I have taken so many steps—it was motivating to see how many steps I had taken the last 7 days.*
>
> *Yes, that is what kept me going, it sort of motivated me to continue and it definitively also helped me in my physical development.*
>
> *For me it was motivating that we could see each other's results, to see each other's results triggered us all.*

We could also identify the importance of setting attainable goals that could be increased as the individual's physical and mental strength increased. Comments in this regard included:

> *Yes, it motivates me to do more, to move more each day because I always want to reach my daily goal* [...]
>
> *Sometimes when I see, oh shit, I've only walked 8,000 steps...but I can still go an extra round.*

All the youths reported trying to increase their physical activity during the project. Three of the youths started training at the gym during the project, 1 of them began taking walks with friends to a greater extent than before, and 1 began going by bicycle to school instead of taking the bus as well as taking walks during breaks:

> *I do not sit during breaks like before; instead, I choose to walk.*

Long-term goals and social attention from peers and the project influenced the youths' motivation. The study period was placed deliberately with a holiday break in the middle to explore what happens to participants' activity level after a 2-week break. The activity levels were plummeting for 4 of the participants who did not start using the activity trackers again after the Christmas break. We found that half of the participants stopped using the activity tracker and accompanying app during the Christmas break when the social attention from the project and their co-users was low. Only the most dedicated youths maintained their use of the device and app. An investigation revealed that the 4 participants who continued to use the technology had established their own long-term goals. The participants without long-term goals stopped using the technology during the holiday break. We concluded that social attention and having long-term goals were the 2 most important factors for motivation.

Social attention from their surroundings influenced the youths' attitudes toward technology and toward their health. We found that 1 participant often wore his activity tracker even when the battery was dead; however, at the same time, we learned that he spoke warmly about the benefits of the bracelet with many different people without really being a dedicated user. After closer investigation, we discovered that he found the social attention that came with the activity tracker stimulating. The remaining youths in the project confirmed the first nondedicated youth's experience by expressing how the project as such gave rise to positive social attention from their surroundings. They expressed that people were often curious to know more about the wearable activity tracker and the connection to health.

Friendly competition from peers was found to be motivating, although taking the competitive element too seriously could become stressful. The youths expressed how they were motivated by the project because it legitimized a social information-sharing culture among the youths. Four of the youths said that they gained increased strength by being part of a team where they could follow each other's developments and could encourage each other. Comments in this regard included:

> *For me, it was motivating that we could see each other's activities.*
>
> *Being in a team motivated me.*
>
> *We encourage each other through the sharing of results.*

Table 1. Average number of steps per day.

Months	November	December	January	February
Participants	8	8	4	3
Steps/day	12,204	10,799	11,523	12,854

It became clear to us that this type of friendly competition within a team played a key role in keeping up the participants' motivation and activity level. However, we also found that sharing goals within a team could lead to stress. Overall, 4 of the participants felt that it was stressful to see how physically active the others were in the app. Two of them were not pleased with their own results, although they expressed that they were more physically active than they had been before:

> *I'm not satisfied with my performance, and I'm also stressed by the others in the app.*

The stress became even more apparent when a step competition was held between teams, a proposal that came from the participants themselves. The youths divided themselves into teams and started to compete. Two of the youths said that it was hard to follow how many steps their teammates took and to contribute to the entire team's performance:

> *I have been sick, and this gives me anxiety because I'm not able to practice and contribute.*

Hence, they expressed an experience that leads to increased stress and diminished motivation.

The awareness of one's own sleeping patterns as a motivator for change was found to be more challenging than the awareness of one's own activity level as a motivator for change. It was obvious that the activity tracker and app provided useful feedback by visualizing patterns and habits in relation to sleep. Positive comments from the youths were:

> *It's a "coach" that tells me when I should go to bed—I have tried it, and it works.*
>
> *I try to sleep well and it helps me—I wake up easier and have the enough sleep.*
>
> *I feel clearer in my head now—I sleep better and feel more alert during the day.*

Overall, 5 youths had used the alarm function and were satisfied with it. It woke them up in the light sleep phase, whereby they experienced that it was easier to wake up feeling clearer in the head. Three youths had used the reminder function, notifying when it is time to go to bed, and said that it helped them. In addition, 1 youth used the reminder function to remember to take her vitamins, which also worked well. One negative comment was:

> *I want to the alarm to wake me up, but I do not notice the alarm.*

A more problematic comment was:

> *Now I know how I'm sleeping, but I cannot do anything about it.*

The last problem shows that the technology in itself did not generate solutions or provide useful enough information that could change the poor sleeping situation. It was a too difficult a challenge for the youths themselves to figure out what was needed on a practical level to establish a positive change. Going to bed earlier did not necessarily lead to more sleep if there were more complicated reasons behind the sleeping problems. Changing sleeping habits was not as straightforward as going out and walking the 1000 or 2000 steps that were needed to reach a preset goal of 10,000 steps. Our conclusion was that supplementary practical support by qualified professionals was needed in these cases.

Being aware of one's own habits had a positive value, but it was not enough to lead to sustainable change for all the youths. All participants expressed increased awareness of their own physical activity levels and sleeping habits. In relation to the 4 participants who took off the activity tracker during the Christmas break, we found that this awareness was still valuable for them. Some of the participants said that they felt comfortable with the awareness that they had achieved and the knowledge that they could form decent habits of physical activity if they wanted to. Two of the participants who had taken off the activity tracker expressed that they were more aware of their habits now, but they could not change either their sleeping habits or their attitude toward long-term exercise:

> *I'm more aware of my actual habits now, but I'm still not able to change the habits; I know more about myself today, but I still cannot change my situation.*

Not being able to change their habits after 2 months influenced their decision to stop using the activity tracker during the Christmas break.

The wearable activity tracker and its accompanying app were identified as a potential tool by the caregivers. We experienced a genuine interest among the staff to constantly improve their work and relationships with their clients. The staff in the HVB home had also previously asked their management for tools that could provide access to and improve their possibilities to influence young people's exercise and sleeping habits. The reason was that the staff had seen so many negative effects of moving too little and of getting poor sleep. At the end of the project, the staff expressed that the activity tracker and accompanying app could be such a tool. The staff expressed a clear value in being able to use this form of technology in their treatment work regarding the youth's habits and health. They saw a potential to support the creation of structure in young people's everyday lives:

> *We usually give advice to young people about picking up their clothes in the morning, eating before bedtime, not watching television just before bedtime, etc. This is a completely different way, to now and then meet the youths and talk about the results from the app.*

It became clear that the graphic results from the accompanying app could be used in conversations with the youths to motivate

or to concretize problems as well as to raise a topic of conversation about sleep and exercise more generally:

> *It is good for getting direct feedback that can be used in counseling.*

The staff expressed that the activity tracker and the accompanying app had opened up for a new type of talks about habits and health with the youths. The staff also told us that the technology had opened up for similar discussions between the employees themselves because they did not always touch upon these difficult subjects in their own group. One identified troublesome view here was that not all of the staff could claim to be good role models based on their own physical activity levels or sleeping habits. In any case, the staff perceived the activity tracker and accompanying app to be positive instruments for making visible the youths' habits, which enabled good discussions. To conclude, the staff experienced that the activity tracker and accompanying app could be used with youths for a certain period to increase their awareness and ultimately also their motivation for improving their physical activity levels and improving their sleeping habits.

Discussion

This study demonstrates how the activity tracker and its accompanying app can open up a topic for discussion regarding how poor lifestyle patterns among vulnerable youths in the context of agency-based social work can be addressed with the help of digital technology. Previous research on health apps to motivate physical activity have demonstrated that the most central factor for success was to include various forms of specific goals that individuals could work toward [6]. In relation to goals, instant gratification and graphical feedback are identified as key elements for motivation [7,8]. These are success factors that correspond well with our study's results. The studied youths in this project went from being nonexercising to walking around 8 km a day, a result one would expect from active people. Instant gratification and graphical feedback provide one explanation for the success. Additionally identified explanations in this project relate to social attention, long-term goals, and the possibility to adjust goals over time. Success in this project means that the technology helped to motivate 3 of the youths to exercise more to the end of the project and that all youths became more aware of their own physical activity levels and sleeping habits.

The study period was arranged deliberately with a holiday break in the middle to explore what happens with the activity level after a 2-week break. During the break, the youths visited their families. It was found that half of the participants stopped using the activity trackers when the social attention from the project and their co-users was lost. In this study, the youths experienced increased strength because of being part of a team where they could follow each other's developments and could encourage each other. In previous research, it has been demonstrated that *friendly* competition can increase the motivation for being physically active [28]. The studied youths talked about the social group and the internal competitive games as triggers for activities. The loss of participation took place when the social situation changed. With the lack of motivating social attention, the activity tracker and accompanying app lost their value for half of the studied group. We thus conclude that the 4 youths who stopped using the device over the Christmas break had been motivated primarily by the positive social attention from using the device. After the break, these 4 youths did not find enough motivation to start the social activity again.

Another interesting finding in relation to social attention was that 1 of the participants often wore his activity tracker in discharged mode. Without being a dedicated user, he spoke warmly about the benefits of the bracelet with many different people. This can be explained theoretically by the phenomenon called *ticket to talk* [35]. Ticket to talk is a phenomenon that opens up and legitimizes conversations between people who are unknown to each other; for example, if 1 person passes by an unknown person with a dog, the first person might start a legitimate conversation by commenting that the dog is cute. The ticket to talk, in this case, the cute dog, thus legitimizes 1 stranger starting a conversation with another stranger. After the opening question via the ticket and a first positive response, it is legitimate to talk about dogs in general and thereafter slowly enter other areas that would not have been legitimate to address directly between 2 strangers (different types of contexts and examples can be found in the study by Svensson and Sokoler [36]). In our case, the activity tracker constituted a ticket to talk regarding both health aspects and cool technology (one of the municipality's expressed interests). The fact that activity trackers were still sort of unusual when starting the project, and were visible, gave both unknown people and the youths themselves legitimacy to start up conversations about, for example, the activity tracker and its connection to exercise and health. People in the surroundings were curious and keen on knowing what the participant had on their wrist and how it worked. Interestingly, the youth in question was not a serious practitioner of physical exercise but rather a social promoter of the digital solution and the idea of a healthy lifestyle.

Returning to long-term goals and the possibility to adjust one's goals, we found that only the most dedicated youths continued to use the technology after the Christmas break. Interviews revealed that these 4 participants had established their own long-term goals, and the participants without their own long-term goals had taken off the activity trackers and stopped using them. The 3 youths who used the activity tracker and accompanying app to the end of the project also expressed that the app not only motivated them to move regularly but it also motivated them to steadily increase their physical activity over the course of the project. For these youths, the digital technology in itself and their own long-term goals were sufficient motivators. It would be interesting for future research to study if, how, and to what extent long-term goals for youths in need of social care can be supported by digital technology.

In this study, the studied youths all expressed that being aware of one's sleeping patterns was helpful. Half of the studied group improved their sleeping habits, whereas the other half did not. It was identified that the technology in itself did not generate solutions or provide sufficient information that could change the poor sleep habits of some of the youths. The challenge to figure out what was needed on a practical level to establish a positive change was too great for the youths themselves. Going

to bed earlier did not necessarily lead to more sleep if there were more complicated reasons behind the sleeping problems. The conclusion was that supplementary practical support by adequate professionals is needed. There exist mobile phone apps that have been developed by researchers aimed to improve sleeping habits [37], and these are also related to intervention programs [38]. The latter pioneer study presents results from a 4-week field study with 12 participants. Their results demonstrate that a very low effort, recommendation-based peripheral display can be an effective method for improving awareness of healthy sleep habits. The accompanying app included an option to specify judged need of sleep, which then was compared against measured sleeping results. If the problem with sleep actually was because of just having bad habits, then the activity trackers and the accompanying app were of help to visualize, monitor, and motivate change.

In previous research, a criticism has been raised regarding the lack of involvement of health professionals during the development of technology, the monitoring of its use, and the feedback given to users [22,23], and thus, it makes sense to clarify our situation in this respect. In the project, the professional social workers were involved as active users. The staff did not take any level of responsibility for the project's proceeding nor did they systematically test the app in their treatment process. This was because the staff had had a bad experience with taking responsibility in a previous externally introduced project. In this study, the staff chose to use the activity tracker and the accompanying app for reasons of curiosity. The staff also reflected over the results and the technology's potential usefulness related to their clients at the end of the study. The caregivers found that the activity tracker and its accompanying app opened up for new type of talks about healthy exercise and sleep habits and health with the youths. Hence, the study opens up a new topic for discussion concerning whether the technology might be a valuable tool to help professionals within social work to help youths to achieve better structure and lifestyle habits.

Conclusions

Motivation apps and their implementation for preventing health-related problems within social work have been identified as a gap in research. To contribute to this area, this research project was based on the following question: *What effects will the studied youths experience in relation to exercise and sleep as a result of using a wearable activity tracker and its accompanying app?* The youths' daily movement was high as they on average walked more than 10,000 steps a day, which corresponds to walking about 8 km daily. The reminder function in the app was useful for some youths because it made it easier to remember to go to sleep on time, which affected the next day in a positive manner. Furthermore, the youths who did not succeed in establishing good sleep and movement habits still expressed that they had gained a better awareness of themselves regarding sleep and exercise.

In general, all youths expressed that they were motivated by the technology and the social attention that followed from its use and from participation in the project. The instant graphical feedback and sharing of information played a crucial role here. When taking a closer look, we could see that the motivation came from different sources. Social attention, being a member of a social group, and the friendly competition motivated all the studied youths. When the friendly competition changed to real competition, however, some of the youths felt negative stress. Half of the studied youths had the additional motivating factor of having established long-term goals. The youths who had not established long-term goals stopped using the device during the holiday break when the social context was lost, whereas those who had established long-term goals continued to use the activity trackers and app after the break and to the end of the project. The increased awareness of one's own sleeping patterns did not by default generate motivation for all youths because it was hard for the youths to know what was needed to change their bad sleeping habits. Here, supplementary practical support by trained professionals is needed. When it comes to the staff, they emphasized that the activity tracker and accompanying app opened up for new types of talks about habits and health issues with the youths. They expressed that the device and app was a useful instrument for making visible the youths' habits, structures, and patterns that might influence health, and this enabled the establishment of good discussions. The staff saw great potential for this technology to assist in their work to create better structure and patterns in everyday life for the youths under their care.

In summary, increased personal awareness, support from social workers, and friendly competition all supported the establishment of health goals for the youths. Although both instant graphical feedback and sharing information through friendly competition had a positive impact, these were not influential beyond the moment and social context. Although this was a short study, having long-term goals was found to be the most powerful factor for influencing the youths to keep on using the app during the study. It is still difficult to predict what the long-term value of the awareness of one's own activity and sleeping habits will be, and longitudinal studies are needed here. More research is also needed relating to how activity trackers can support personal long-term health goals for youths and what influence and role professional caregivers can have here.

Acknowledgments

I would like to express my gratitude to all project members and people that contributed to the project and field study. I would like to thank Helsingborg City who sponsored the implementation of the project. Special thanks to PArT, who staffed the practical implementation of the project, and Fou Helsingborg who managed the coordination with Helsingborg City, PArT, and the Universities involved. Many thanks to Lund University for sponsoring researchers from their organization; and finally Kristianstad University for sponsoring the present author's time in the project, and the time spent to produce this paper.

Conflicts of Interest

None declared.

References

1. Swan M. Health 2050: the realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. J Pers Med 2012 Sep 12;2(3):93-118 [FREE Full text] [doi: 10.3390/jpm2030093] [Medline: 25562203]
2. Cummiskey M. There's an app for that smartphone use in health and physical education. J Phys Health Educ Recreat Dance 2011 Oct;82(8):24-30. [doi: 10.1080/07303084.2011.10598672]
3. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. J Med Internet Res 2012 Apr 20;14(2):e55 [FREE Full text] [doi: 10.2196/jmir.1950] [Medline: 22522112]
4. Lupton D. M-health and health promotion: the digital cyborg and surveillance society. Soc Theory Health 2012 Jun 27;10(3):229-244. [doi: 10.1057/sth.2012.6]
5. Lupton D. Quantifying the body: monitoring and measuring health in the age of mHealth technologies. Crit Public Health 2013 Dec;23(4):393-403. [doi: 10.1080/09581596.2013.794931]
6. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. Telemed J E Health 2011 Dec;17(10):801-803. [doi: 10.1089/tmj.2011.0055] [Medline: 22010977]
7. Adibi S. Biomedical sensing analyzer (BSA) for mobile-health (mHealth)-LTE. IEEE J Biomed Health Inform 2014 Jan;18(1):345-351. [doi: 10.1109/JBHI.2013.2262076] [Medline: 24403433]
8. Shigaki CL, Koopman RJ, Kabel A, Canfield S. Successful weight loss: how information technology is used to lose. Telemed J E Health 2014 Feb;20(2):144-151. [doi: 10.1089/tmj.2013.0163] [Medline: 24303931]
9. Lund A. Ki.se. 2018 May 31. Motion mot depression [Exercise against depression] URL: https://ki.se/forskning/motion-mot-depression [accessed 2018-10-15] [WebCite Cache ID 73BTmb7fC]
10. Strine TW, Chapman DP. Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. Sleep Med 2005 Jan;6(1):23-27. [doi: 10.1016/j.sleep.2004.06.003] [Medline: 15680291]
11. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, et al. Sleep quality prediction from wearable data using deep learning. JMIR Mhealth Uhealth 2016 Nov 4;4(4):e125 [FREE Full text] [doi: 10.2196/mhealth.6562] [Medline: 27815231]
12. Shin CJ, Kim J, Grigsby-Toussaint D. Mobile phone interventions for sleep disorders and sleep quality: systematic review. JMIR Mhealth Uhealth 2017 Sep 7;5(9):e131 [FREE Full text] [doi: 10.2196/mhealth.7244] [Medline: 28882808]
13. Rönkkö K, Bergman AK, Svensson M, Danielsson P, Åström F, Lundin T. Digital teknik och socialt arbete: Att motivera socialt utsatta ungdomar med stöd av motivationsappar och hälsoarmband [Digital technology and social work - To motivate socially disadvantaged young people, supported by motivational programs and health braces]. Kristianstad: Kristianstad University Press; 2017:1-46.
14. Ridgers DN, McNarry AM, Mackintosh AK. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
15. 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]
16. Berzin SC, Singer J, Chan C. Practice Innovation through Technology in the Digital Age: A Grand Challenge for Social Work. Cleveland: American Academy of Social Work and Social Welfare; 2015:3-23.
17. Reamer FG. Evolving ethical standards in the digital age. Aust Soc Work 2016 Apr 4;70(2):148-159 [FREE Full text] [doi: 10.1080/0312407X.2016.1146314]
18. Goldkind L, Wolf L. A digital environment approach: four technologies that will disrupt social work practice. Soc Work 2014 Oct 13;60(1):85-87. [doi: 10.1093/sw/swu045]
19. Willoughby M. A review of the risks associated with children and young people's social media use and the implications for social work practice. J Soc Work Pract 2018 May 8:1-14 [FREE Full text] [doi: 10.1080/02650533.2018.1460587]
20. Coulton CJ, Goerge R, Putnam-hornstein E, de Haan B. Harnessing Big Data for Social Good: A Grand Challenge for Social Work. Cleveland: American Academy of Social Work and Social Welfare; Jul 2015:1-20.
21. Herbert WJ. Grand challenges for social work: research, practice, and education. Soc Work Res 2016 Jun;40(2):67-70 [FREE Full text] [doi: 10.1093/swr/svw007] [Medline: 27257360]
22. Cheng NM, Chakrabarti R, Kam JK. iPhone applications for eye care professionals: a review of current capabilities and concerns. Telemed J E Health 2014 Apr;20(4):385-387. [doi: 10.1089/tmj.2013.0173] [Medline: 24476190]
23. Lalloo C, Jibb LA, Rivera J, Agarwal A, Stinson JN. "There's a Pain App for that": review of patient-targeted smartphone applications for pain management. Clin J Pain 2015 Jun;31(6):557-563. [doi: 10.1097/AJP.0000000000000171] [Medline: 25370138]

24. Wicks P, Stamford J, Grootenhuis MA, Haverman L, Ahmed S. Innovations in e-health. Qual Life Res 2014 Feb;23(1):195-203 [FREE Full text] [doi: 10.1007/s11136-013-0458-x] [Medline: 23852096]
25. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. J Med Syst 2014 Jan;38(1):9995. [doi: 10.1007/s10916-013-9995-7] [Medline: 24346929]
26. Musiat P, Goldstone P, Tarrier N. Understanding the acceptability of e-mental health--attitudes and expectations towards computerised self-help treatments for mental health problems. BMC Psychiatry 2014 Apr 11;14:109 [FREE Full text] [doi: 10.1186/1471-244X-14-109] [Medline: 24725765]
27. Sparkes J, Valaitis R, McKibbon A. A usability study of patients setting up a cardiac event loop recorder and BlackBerry gateway for remote monitoring at home. Telemed J E Health 2012;18(6):484-490. [doi: 10.1089/tmj.2011.0230] [Medline: 22676379]
28. Naslund JA, Aschbrenner KA, Barre LA, Bartels SJ. Feasibility of popular m-health technologies for activity tracking among individuals with serious mental illness. Telemed J E Health 2015 Mar;21(3):213-216 [FREE Full text] [doi: 10.1089/tmj.2014.0105] [Medline: 25536190]
29. Swider M. Techradar. 2014 Mar 27. Jawbone Up24 review: The fitness tracker with the most flair URL: https://www.techradar.com/reviews/gadgets/jawbone-up24-review-1230596/review [accessed 2018-10-15] [WebCite Cache ID 73BUROtJ8]
30. Lewin K. Frontiers in group dynamics: concept, method and reality in social science; social equilibria and social change. Human Relat 1947;1(1):5-41 First Published June 1, 1947. [doi: 10.1177/001872674700100103]
31. Mumford E. Advice for an action researcher. Info Technol People 2001 Mar;14(1):12-27. [doi: 10.1108/09593840110384753]
32. Dittrich Y, Rönkkö K, Eriksson J, Hansson C, Lindeberg O. Cooperative method development. Empir Software Eng 2007 Dec 18;13(3):231-260. [doi: 10.1007/s10664-007-9057-1]
33. Danielsson P, Ebrahimsson S, Lundin T, Mattisson S. Partinfo.se. Helsingborg: PART; 2015. Motivationsappar i socialt arbete: Erfarenheter och resultat från en pilotstudie mellan november 2014 – juni 2015 URL: http://www.partinfo.se/wp-content/uploads/PDF-Rapport-motivationsappar-FINAL-20160208.pdf [accessed 2018-10-15] [WebCite Cache ID 73BbYVYUp]
34. Rönkkö K. Ethnography. In: Laplante PA, editor. Encyclopedia of Software Engineering Three-Volume Set (Print). Boca Raton, Florida United States: Auerbach Pub; Nov 2010:278-346.
35. Sacks H. In: Jefferson G, Schegloff EA, editors. Lectures on Conversation, Volumes I and II. Oxford UK & Cambridge USA: Wiley-Blackwell; Jun 1995:1-1520.
36. Svensson MS, Sokoler T. Ticket-to-talk-television: designing for the circumstantial nature of everyday social interaction. In: NordiCHI '08 Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges.: ACM; 2008 Presented at: The 5th Nordic Conference on Human-Computer Interaction: Building Bridges; October 20-22, 2008; Lund, Sweden p. 334-333. [doi: 10.1145/1463160.1463197]
37. Munafo D, Hevener W, Crocker M, Willes L, Sridasome S, Muhsin M. A telehealth program for CPAP adherence reduces labor and yields similar adherence and efficacy when compared to standard of care. Sleep Breath 2016 May;20(2):777-785 [FREE Full text] [doi: 10.1007/s11325-015-1298-4] [Medline: 26754933]
38. Bauer JS, Consolvo S, Greenstein B, Schooler J, Wu E, Watson NF, et al. ShutEye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. In: CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.: ACM; 2012 May Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 05-10, 2012; Austin, Texas, USA p. 1401-1410. [doi: 10.1145/2207676.2208600]

Abbreviations

HBV: Home for Care or Living (In Swedish, HVB_hem: Hem för vård eller boende)
mHealth: mobile health

Edited by G Eysenbach; submitted 27.10.17; peer-reviewed by F Akbar, S Aggarwal; comments to author 15.03.18; revised version received 03.08.18; accepted 16.08.18; published 21.12.18.

Please cite as:
Rönkkö K
An Activity Tracker and Its Accompanying App as a Motivator for Increased Exercise and Better Sleeping Habits for Youths in Need of Social Care: Field Study
JMIR Mhealth Uhealth 2018;6(12):e193
URL: https://mhealth.jmir.org/2018/12/e193/
doi:10.2196/mhealth.9286
PMID:30578186

Original Paper

Accuracy of Wrist-Worn Activity Monitors During Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study

Ravi Kondama Reddy[1], MSc, PhD; Rubin Pooni[2], BSc; Dessi P Zaharieva[2], BSc, MSc; Brian Senf[3], BSc; Joseph El Youssef[3], MD; Eyal Dassau[4], PhD; Francis J Doyle III[4], PhD; Mark A Clements[5], MD, PhD; Michael R Rickels[6], MD; Susana R Patton[7], PhD; Jessica R Castle[3], MD; Michael C Riddell[2], PhD; Peter G Jacobs[1], PhD, MSEE

[1]Department of Biomedical Engineering, Oregon Health & Science University, Portland, OR, United States

[2]School of Kinesiology and Health Science, York University, Toronto, ON, Canada

[3]Harold Schnitzer Diabetes Health Center, Oregon Health & Science University, Portland, OR, United States

[4]Harvard John A Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, United States

[5]Children's Mercy Kansas City, Kansas City, MO, United States

[6]Institute for Diabetes, Obesity & Metabolism, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, United States

[7]Department of Pediatrics, University of Kansas Medical Center, Kansas City, KS, United States

Corresponding Author:
Peter G Jacobs, PhD, MSEE
Department of Biomedical Engineering
Oregon Health & Science University
3303 SW Bond Avenue
Portland, OR, 97239
United States
Phone: 1 503 358 2291
Email: jacobsp@ohsu.edu

Abstract

Background: Wrist-worn activity monitors are often used to monitor heart rate (HR) and energy expenditure (EE) in a variety of settings including more recently in medical applications. The use of real-time physiological signals to inform medical systems including drug delivery systems and decision support systems will depend on the accuracy of the signals being measured, including accuracy of HR and EE. Prior studies assessed accuracy of wearables only during steady-state aerobic exercise.

Objective: The objective of this study was to validate the accuracy of both HR and EE for 2 common wrist-worn devices during a variety of dynamic activities that represent various physical activities associated with daily living including structured exercise.

Methods: We assessed the accuracy of both HR and EE for two common wrist-worn devices (Fitbit Charge 2 and Garmin vívosmart HR+) during dynamic activities. Over a 2-day period, 20 healthy adults (age: mean 27.5 [SD 6.0] years; body mass index: mean 22.5 [SD 2.3] kg/m^2; 11 females) performed a maximal oxygen uptake test, free-weight resistance circuit, interval training session, and activities of daily living. Validity was assessed using an HR chest strap (Polar) and portable indirect calorimetry (Cosmed). Accuracy of the commercial wearables versus research-grade standards was determined using Bland-Altman analysis, correlational analysis, and error bias.

Results: Fitbit and Garmin were reasonably accurate at measuring HR but with an overall negative bias. There was more error observed during high-intensity activities when there was a lack of repetitive wrist motion and when the exercise mode indicator was not used. The Garmin estimated HR with a mean relative error (RE, %) of −3.3% (SD 16.7), whereas Fitbit estimated HR with an RE of −4.7% (SD 19.6) across all activities. The highest error was observed during high-intensity intervals on bike (Fitbit: −11.4% [SD 35.7]; Garmin: −14.3% [SD 20.5]) and lowest error during high-intensity intervals on treadmill (Fitbit: −1.7% [SD 11.5]; Garmin: −0.5% [SD 9.4]). Fitbit and Garmin EE estimates differed significantly, with Garmin having less negative bias (Fitbit: −19.3% [SD 28.9], Garmin: −1.6% [SD 30.6], P<.001) across all activities, and with both correlating poorly with indirect calorimetry measures.

Conclusions: Two common wrist-worn devices (Fitbit Charge 2 and Garmin vívosmart HR+) show good HR accuracy, with a small negative bias, and reasonable EE estimates during low to moderate-intensity exercise and during a variety of common daily

activities and exercise. Accuracy was compromised markedly when the activity indicator was not used on the watch or when activities involving less wrist motion such as cycle ergometry were done.

(JMIR Mhealth Uhealth 2018;6(12):e10338) doi:10.2196/10338

KEYWORDS

heart rate; energy metabolism; fitness trackers; high-intensity interval training; artificial pancreas

Introduction

Background

Consumer-based wrist-worn multisensor activity monitors have emerged as an increasingly popular way to track various physiological metrics such as heart rate (HR) and physical activity levels, with the latter being typically expressed in the form of step counts or energy (caloric) expenditure (EE). Sales of activity monitors have doubled from approximately 30 million units in 2014 to approximately 70 million units in 2017 [1,2]. The growth in activity monitors has been largely driven by consumer interest in monitoring and sometimes sharing physical activity levels, workouts, and total daily EE within social networks. In the scientific community, there is increasing interest in whether activity monitors may also be used within a health care setting to collect these same data and help patients and health care providers better manage weight and/or chronic illnesses. For example, in people with type 1 diabetes, aerobic exercise is known to cause steep drops in blood glucose levels, whereas anaerobic exercise can cause glucose levels to rise [3]. Monitoring of patient physical activity levels may be helpful in implementing insulin and/or nutritional strategies to optimize glucose control in type 1 diabetes [4]. In theory, activity monitors can be used in conjunction with on-body continuous glucose monitors, an insulin pump and a control algorithm to adjust insulin delivery, and perhaps glucagon delivery in real time [5,6]. Activity monitors can also be used within algorithm-driven decision support systems to help avert exercise-induced hypoglycemia or late onset hypoglycemia. Automated insulin delivery systems can potentially modify insulin dosing in response to activity monitors to reduce the risk (or severity) of exercise-induced hypoglycemia in people living with type 1 diabetes [7-10]. For any medical system utilizing an activity monitor, the accuracy of the HR and EE estimates by the activity monitor is critical as it can influence medical dosing decisions and patient outcomes. There are 3 distinct challenges with using the activity monitors within medical systems, namely, detecting the onset of the activity, distinguishing the type of the detected activity, and estimating the intensity and duration of the activity, as each of these functions can determine how medical systems may behave. In this paper, we explore the accuracy of HR and EE estimates from 2 popular activity monitors to determine if the accuracy of these wearables is sufficient for use within medical applications such as automated insulin delivery systems for use within type 1 diabetes glucose management.

In the earlier models of activity monitors, only accelerometers were used to estimate EE [11], but in more recent multisensor models, photoplethysmography (PPG) is being used to estimate HR [12] and, potentially, to improve the accuracy in estimating EE [13]. With the inclusion of HR as measured by the PPG sensor and acceleration as measured by the accelerometer, the accuracy of the estimated EE is expected to be improved in newer models. For example, Zakeri et al [14] showed that EE can be estimated using both accelerometry and HR along with several additional patient-specific parameters such as age, weight, and height. The Zakeri et al algorithm utilizing accelerometry and HR to estimate EE and metabolic equivalents (METs) has been used in the past to inform an automated insulin delivery system during physical exercise [6]. In a post hoc analysis that combined both HR and accelerometer signals, researchers demonstrated that steady-state aerobic exercise could be detected early before rapid changes to glucose occurred [15]. In recent studies involving predominantly steady-state aerobic activities, wrist-worn activity monitors have been shown to have reasonable accuracy in HR estimation (approximately 5% error) but a poor estimate of EE, where the error was found to be closer to approximately 30% with a negative bias [16,17]. In free-living conditions, however, activity monitors are worn typically on the nondominant wrist during multiple forms of exercise in nonsteady states, not just aerobic exercise on a treadmill performed at a constant workload or intensity (steady-state). For example, in free-living conditions, many individuals often perform resistance exercise using free weights or their own body weight, followed by some form of high-intensity interval training (HIIT) within the same session. In fact, in the diabetes population, patients are encouraged to perform both resistance and aerobic training all in one session. HIIT has recently been recommended to rapidly improve fitness, body composition, and overall glycemic control [18-20].

Presently, there are at least 4 studies [21-24] that have investigated the accuracy of wearable devices during resistance exercises and none during HIIT training. Bai et al [21] reported that EE measured during an unstructured resistance exercise protocol in which participants selected exercises and loads was inaccurate across numerous devices. The devices included 5 wrist-worn devices (Fitbit Flex, Jawbone Up24, Misfit Shine, Nike+ Fuelband SE, and Polar Loop) and 2 research monitors (Actigraph GT3X+ on the waist and the BodyMedia Core on the arm). In this study, 52 participants tested these 7 different devices, and the wearable devices had lower accuracy for EE when compared with a metabolic analysis system. None of the devices in this study reported HR measures. Horton et al [22] assessed the validity of HR only using the Polar M600 when compared with a 3-lead electrocardiogram (ECG) during both aerobic and resistance exercises. The accuracy of the wearable device was reported to be better during aerobic exercise (92%) as compared with only 35% accurate during the resistance exercises. In this study, participants completed squats, shoulder shrugs, bicep curls, and lunges with dumbbells at a self-selected weight. Jo E et al [23] reported poor correlation and HR

accuracy in the Fitbit Charge HR device. In this study, subjects completed a short-resistance exercise bout involving resisted arm raises, resisted lunges, and isometric plank. In a large cohort study, Bourdreaux et al [24] standardized the selection of the weights utilized during the resistance exercises: 2 upper body exercises (chest press, latissimus dorsi pulldown) and 2 lower body exercises (leg extension and leg curl) among the subjects using a standardized 10-rep max protocol. Results from this study demonstrated that HR measured by nonwrist worn devices were relatively accurate, whereas wrist-worn devices showed poor correlations (R<.8) and higher error during resistance exercises (mean absolute percent error [MAPE] >9%). They also showed that the EE measured by the devices was poor, with MAPE values ranging between 43% and 57%.

Objectives

The primary aim of this study was to examine the accuracy of both HR and EE across a wide range of dynamic activities including resistance training, HIIT, and aerobic training. A secondary aim was to examine the accuracy when the optional *activity mode* is not selected on the wearable. There may be times when people exercise, but they do not indicate that they are exercising; we wanted to determine the accuracy both when they do and do not indicate that they are exercising.

Methods

Participants

The experimental protocol conformed to the standards set by the Declaration of Helsinki and was approved by the institutional review board at the Oregon Health and Science University (OHSU, Portland Oregon) and by the research ethics board at York University (Toronto, Canada). This study recruited 20 healthy adults (11 females; 10 subjects at OHSU; 10 at York University) who all provided informed consent before taking part in the study. Participants were screened for any cardiovascular complications using a Physical Activity Readiness Questionnaire [25].

Study Protocol

Participants attended the research laboratory on 2 separate occasions, separated by 24 hours. Each visit involved simultaneous recordings of HR (beats per minute) and EE (kcals and METs) from the respective criterion measures during a series of physical activities and structured exercises. On the first visit, a stadiometer (Seca, model220, Hamburg, Germany) was used to measure height to the nearest 0.25 cm (without shoes) and body mass was measured to the nearest 0.1 kg using a scale (Seca, model 707, Hamburg, Germany), with the participant dressed in workout clothes. As per the manufacturer's instructions, age, gender, height, and weight were used to initialize the wearable devices and associated applications. These same data were also inputted to a portable metabolic unit (Cosmed, Rome, Italy). Two wearable devices (one of each brand) were tested at the same time on all participants (one on each wrist as per manufacturer's instructions) using a randomized and counterbalanced method. On each visit, participants undertook 2 activity blocks (see below for further details) following setup of the devices and synchronization of all the devices to a single clock before the exercise protocol commenced.

Activities

At visit 1, participants performed 2 blocks of physical activity separated by a 30-min rest period. In the first block, participants performed a graded maximal aerobic exercise test (treadmill or cycle ergometer, 10 subjects per mode) to volitional exhaustion (ie, progressive to peak oxygen consumption, VO_2 peak). These will be referred to as MAX-T (MAX-treadmill) and MAX-C (MAX-cycle ergometer) tests. During MAX-T, each participant began with a 5-min standing rest, followed by 4 min of walking as a warm-up (3.0 mph, 0% grade for 2 min then at 5% grade for 2 min). After the warm-up, participants self-selected a comfortable running speed between 4 to 6 mph, and subsequently, the treadmill incline was increased by 2% every 2 min until the participant reached volitional exhaustion. At each workload stage, participants were asked to assess their level of physical exertion using the Borg Rating of Perceived Exertion (RPE) 10-point scale [26]. For the participants performing the MAX-C test, each participant began with a 5-min seated rest followed by 4 min of warm-up cycling at a moderate cadence (approximately 50-60 revolutions per minute [rpm]) at zero load. After this, cycling cadence was maintained at 60 rpm, and the power output was increased every 2 min by 30 watts until the participant reached volitional exhaustion. Borg RPE was assessed at the end of each 2-min stage. For both MAX-T and MAX-C protocols, the wearables were placed in the appropriate exercise setting (ie, running or cycling) and worn on the wrist as per manufacturer's specifications. Following the exercise test, the participants rested for 30 min. In the second block of activity on the same day, a resistance circuit workout was performed (2 sets of 8 repetition max of all the major muscle groups). Subjects selected a suitable dumbbell weight that they could maintain a proper form for 8 repetitions before muscular fatigue. The following 6 exercises were performed: dumbbell bicep curls, Romanian deadlifts, Bulgarian split squat, dumbbell bench press, dumbbell shoulder press, and dumbbell step ups. After a 20-min cool-down, participants then left the laboratory.

At visit 2, performed the next day, participants undertook 2 new activity blocks. The first activity block consisted of 28 min of routine activities of daily living (ADLs), while the second block included high-intensity interval training (HIIT) for 27 min (including warm-up and cool-down). Six ADLs were performed to simulate daily chores. Each activity was 3 min in duration. Activities included sitting on a chair or lying on a bed, washing of dishes and simulated loading and unloading of a dishwasher, sweeping or vacuuming of a small room, organizing a room or adjusting furniture in the room, scrubbing of walls and carpet/floor, and self-paced ascending and descending of a flight of stairs. These activities were preceded and followed by two 5-min segments of seated rest. In the second activity block, participants executed the same exercise mode (ie, treadmill and cycle ergometer) as was done in the peak exercise test. The high-intensity activities are referred to as HIIT-T (HIIT-treadmill) and HIIT-C (HIIT-cycle ergometer). For HIIT-C, participants were asked to cycle at approximately 60 rpm for 2 min at a low intensity with low resistance,

corresponding to approximately 30% of their peak power output in watts (as measured during MAX-C), and then at a high intensity (60 rpm), at a power output corresponding to approximately 80% of their peak power output for 2 min, for a total of 5 cycles. For the treadmill intervals, participants were asked to walk for 2 min at a treadmill speed and slope corresponding to approximately 30% HR reserve (as measured during MAX-T), and then run/jog at a speed and slope corresponding to approximately 80% of their HR reserve for 2 min, for a total of 5 cycles. This session was completed following a cool-down period of 5 min.

Wearables Devices

Although multiple devices were available that could provide the relevant exercise metrics, we chose the following 2 devices mentioned below after considering their costs and their ability to integrate with a control system running on an Android platform. Henriksen et al provided a detailed review of the many devices that are available and have been tested over the last few years [27].

Garmin vívosmart HR+

The Garmin vívosmart HR+ (2016 version, Garmin International Inc, Kansas, US) is a multisensor activity monitor that has an accelerometer, global positioning system, and built-in PPG sensor that uses the "Elevate" wrist HR technology to measure HR at the wrist. According to the device specifications, the frequency at which HR is measured is normally once every 15 seconds, but triggering the *device key* button and setting the wearable to an activity mode (eg, run) increases the frequency at which HR is measured. EE values are reported in calories for a given activity session, also when the device key is pressed. Garmin provided a special interface to export data from the device when the *device key* button was not indicated. This provided a reliable method to download data. The firmware version of the device was 3.20. Data were exported via Bluetooth low energy (BTLE) to the Garmin-Connect App version 3.17.

Fitbit Charge 2

The Fitbit Charge 2 (2017 version, Fitbit Inc, California, US) is a multisensor activity monitor that has an accelerometer and built-in PPG sensor that uses the "PurePulse" wrist HR technology to measure HR at the wrist. The sample rate at which HR is measured varies and depends on the level of activity; the Charge 2 uses SmartTrackTM to automatically detect and record select exercises, but the manufacturer recommends using the exercise menu to improve the precision of HR and EE measurements. All data collected from the Fitbit were collected with the particular exercise selected as recommended by the manufacturer. Data could not be exported reliably without the type of exercise selected via the button press; this prevented the collection of data without the users pressing the button indicating the type of activity selected. According to the manufacturer, the frequency at which HR is measured during activity mode is once every second. EE values are reported in calories for a given exercise session. Data were exported via BTLE to the Fitbit App version 2.35. The firmware version of the device was 22.54.6. Data were downloaded at the highest sample rate possible through Fitabase (Small Steps Labs, California, US), a third party research platform designed to collect data from Fitbit using the developer application programming interface (API). The use of Fitbit with Fitabase also allows for estimates of METs for an additional assessment of the relative energy costs of a given activity, compared with rest, and for the determination of estimated oxygen consumption (VO_2) expressed in ml $O_2 \cdot kg^{-1} \cdot min^{-1}$.

Heart Rate Criterion Measure

Participants wore the Polar H7 (BTLE version, Polar Electro, Kempele, Finland) chest strap HR monitor, which was secured tightly to ensure skin contact. The data from the Polar H7 was transmitted to the Polar A300 (Polar Electro, Kempele, Finland), and the second level data from this device was downloaded using the Polar Flow App. Although some studies have shown the limitation of these devices as compared with the gold standard ECG measure of HR [24,28], these chest-based HR monitors have been used to inform glucose control systems of exercise [8,10,15].

Energy Expenditure Criterion Measure

Participants wore a portable indirect calorimeter, Cosmed K4b2 or Cosmed K5 (Rome, Italy), which collected breath-by-breath data on the ventilatory parameters (ie, oxygen consumption, VO_2). EE was estimated from the direct measurement of oxygen consumption and carbon dioxide production. The units were calibrated before each session according to the manufacturer's instructions. EE data were downloaded from the cardiopulmonary exercise testing suite.

Statistical Analysis

Statistical analysis was performed separately for HR and EE. Data from the indirect calorimetry (VO_2 and carbon dioxide consumption [VCO_2]) served as the reference standard measurement for calculations of EE (kcal/min). Data from the Polar HR monitor served as the as the reference standard for HR (beats per minute, bpm). In this analysis for both EE and HR, we analyzed all the data collected from each device, and error was calculated as device measurement-reference standard, and mean relative error (RE, %) was calculated as the mean of the device measurement-reference standard × 100/reference standard. We also report MAPE as the mean of the absolute value of device measurement-reference standard × 100/reference standard. Error in HR was calculated at each measurement using the closest data collected from the reference standard as the reference measurement. We observed in our data that the sample rate of the devices varied, with the reference standard Polar measuring the HR every second, the Fitbit measuring every 1 to 15 seconds, and Garmin measuring every 5 to 60 seconds. Pearson (*r*) correlation coefficient and Bland-Altman analysis were used to assess the mean bias and agreement between the devices and the reference standard. We adopted the widely accepted level of accuracy of 5% to be within the acceptable limits [16]. Student *t* test with the Satterthwaite approximation was performed to assess the difference in HR measured between Garmin devices when the activity mode was indicated and when it was not. We also performed the same statistical tests to assess the differences between the errors in the HR measurements for

activities with repetitive wrist motion (treadmill tests) as compared with activities with no repetitive wrist motion (ergometer tests). Error in EE was only calculated across an entire activity session as higher resolution data could not be obtained from the devices. Matched paired *t* tests were performed to assess the difference in RE and MAPE of EE between Fitbit and Garmin for each activity. One-way analysis of variance with a Tukey honest significant difference post hoc test was performed to assess the difference in RE and MAPE of EE between activities within each device. We used concordance class correlation to measure agreement between the devices tested. All statistical analyses were conducted in R (R Core Team, Vienna, Austria, version 3.4.2) and GraphPad Prism 7 (GraphPad Software, La Jolla, CA, version 7.0c) [29].

Results

Cohort

All 20 participants recruited for the study completed the procedures. Table 1 describes the participant characteristics.

Heart Rate Accuracy

We analyzed a total of 83,349 simultaneous HR pairs of data, whereby a pair is either a Garmin or a Fitbit measurement compared with the reference standard (Polar chest strap). There were a total of 61,499 pairs for the Fitbit HR data, 18,317 pairs of HR data from Garmin (with the activity mode indicated), and 3533 pairs of HR data from Garmin with no button press (activity mode not indicated). We analyzed data collapsed across all activities and also looked at accuracy during each individual activity. There was no difference in accuracy between the 2 devices when the activity mode was indicated. The overall performance was significantly worse if the activity mode was not indicated on the Garmin device compared with when activity mode was indicated (P<.001). Figure 1 shows results of the HR data across a test session for 1 subject. Both panels show that when the activity mode is not initiated on the wearable, there is less accuracy and also a distinct phase shift whereby the Garmin with no button trace appears to be shifted in time relative to the Polar. This shift in time is a minor contributor to the inaccuracy within the HIIT activities. The majority of error was from devices failing to track during dynamic activities.

For HR data collected with the activity mode indicated, a systematic negative bias was observed in both Fitbit and Garmin devices. The mean relative error, RE (SD) for the Fitbit device on the collapsed data was −4.71% (19.63), the mean RE (SD) for the Garmin (with activity mode indicated) was −3.33% (16.67), and the mean RE (SD) for the Garmin (with activity mode not indicated) was −5.47% (22.79; comparing the Garmin devices with activity mode indicated vs not indicated. P<.001). MAPE (SD) for the Garmin and Fitbit was 10.79 % (13.14) and 11.33% (16.71), respectively. Mean HR accuracy across each activity was analyzed and compared with the reference standard; these data are shown in Table 2.

The lowest mean error in measuring HR was observed during the HIIT-T (Fitbit: −1.7% [SD 11.5], Garmin: −0.5% [SD 9.4]), whereas the highest error was observed on both HIIT-C (Fitbit: −11.4% [SD 35.7], Garmin: −14.3% [SD 20.5]) and during MAX-C (Fitbit: −16.4% [SD 21.6], Garmin: −9.3% [SD 17.0]). Figure 2 shows the variability between and within activities. When the activity mode of the wearables are activated (panels A and B), median % relative errors are within the 5% error threshold for both devices. When the activity mode is not activated, as observed in panel C, the median % relative error significantly exceeds the 5% threshold across many of the activities.

The correlation between the HR values on the wearables and our gold standard chest band sensor was best during MAX-T (Fitbit: 0.94, Garmin: 0.94), whereas poor correlation between the HR values was observed during the HIIT-C (Fitbit: 0.46, Garmin: 0.71). The relative error across the collapsed data for the activities with repetitive motion of the upper torso (ie, treadmill tests) was observed to be significantly lower at −1.6% (SD 9.6) when compared with activities with no repetitive motion of the upper torso (ie, cycle ergometer tests) at −12.25% (SD 19.3; P<.001). Scatter plots between the simultaneous measures across all the activities are shown in Figure 3.

Bland-Altman plots indicated that all 3 devices underestimated the HR when compared with the reference standard as indicated in Figure 4. The variability between these devices was comparable. However, the wearable devices tended to have significantly higher error when the HR signal transitioned quickly and at higher intensity.

There was a generally small but significant impact of the wrist side worn (ie, left vs right) on the percent absolute relative error. Using a *t* test, the error was shown to be higher on the right hand versus the left hand for the MAX-T (6.6% vs 5.1%, P<.001), HIIT-T (6.72% vs 5.85%, P=.002), and ADLs (13.33% vs 11.17%, P<.001), whereas the error was higher on the left hand versus the right hand for resistance (15.0% vs 13.5%, P<.001) and MAX-C (9.53 vs 2.97%, P<.001).

Energy Expenditure Accuracy

Due to the limitation on the Garmin Connect application, EE data could only be compared at a low resolution, namely an average across each activity mode (eg, ADL, HIIT-C, or HIIT-T). Both Fitbit and Garmin performed reasonably well in estimating task-specific EE, when looking at the group as a whole, but considerable error was noted for some of the activities, particularly with cycling activities for Fitbit and resistance activities for Garmin. Fitbit and Garmin EE estimates differed significantly, with Garmin having less negative bias overall (Fitbit: −19.3% [SD 28.9], Garmin: −1.6% [SD 30.6]; P<.001). Table 3 shows the error in EE estimations for each of the activities for both devices.

Figure 5 shows the % relative error (RE) in EE for Fitbit and Garmin during each activity as scatter plots, when compared with the Cosmed indirect calorimeter.

Table 1. Participant characteristics (n=20). VO_2 max (maximal oxygen uptake) was measured at the incremental test to exhaustion.

Characteristic	Value
Age (years), mean (SD)	27.5 (6.0)
Height (cm), mean (SD)	173.2 (9.5)
Weight (kg), mean (SD)	67.9 (10.8)
Body mass index (kg/m^2), mean (SD)	22.5 (2.3)
VO_2 max (mL/min/kg), mean (SD)	48.0 (8.7)
Wrist (cm), mean (SD)	15.6 (2.0)
Race, n (%)	
White	17 (85)
Asian	2 (10)
Native American/Canadian	1 (5)

Figure 1. Two-day study protocol with "R" indicating the rest periods and "T" indicating the transition period between the different types of activities. Data are shown from 2 different participants wearing all devices in panels A and B. Note, Garmin devices were worn by the participants here in 2 different modes: one with the activity mode indicated (Garmin) and the other without (Garmin: No Button). Panel A shows the data during the cycle ergometer tests and panel B shows the data from the treadmill tests. Data in panel A highlight the error observed during higher intensity exercises where wrist movement was less pronounced during cycle ergometer testing. Panel B shows treadmill results when the Garmin, Fitbit, and Polar data are very closely matched across the exercise types. ADLs: activities of daily living; C: cycle ergometer; HIIT-C: high-intensity interval training-Cycle ergometer; HIIT-T: high-intensity interval training-Treadmill; T: treadmill.

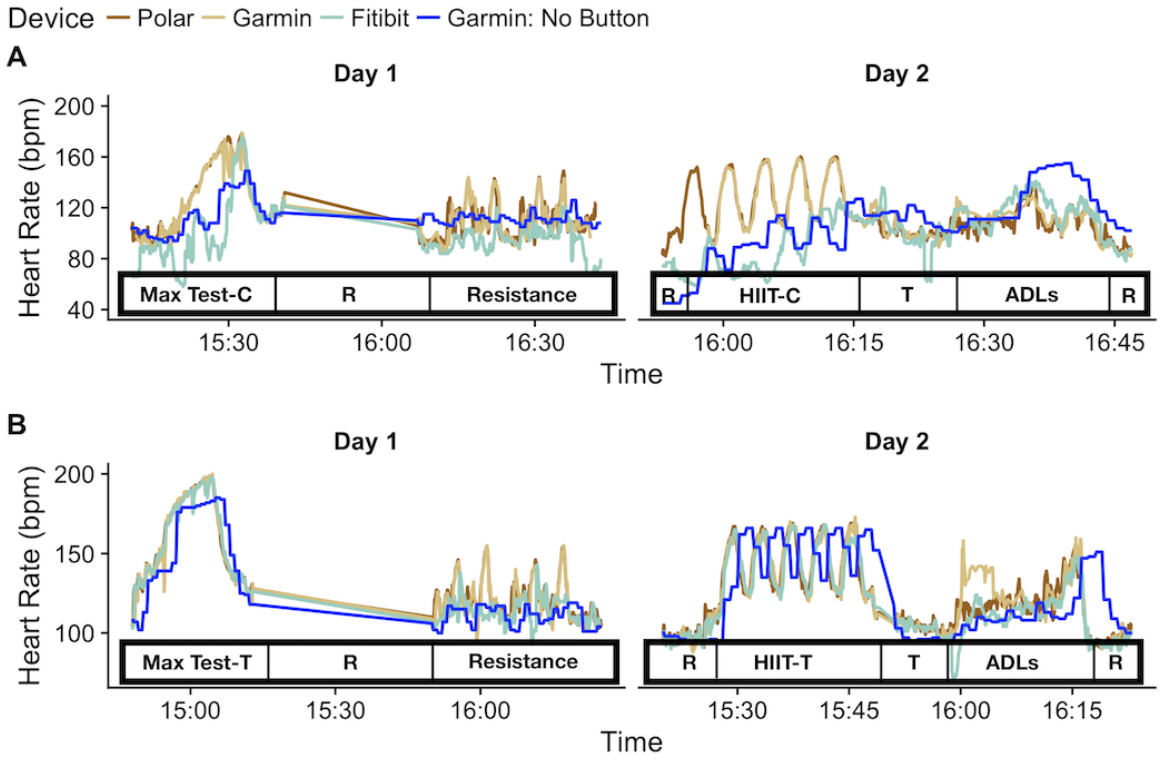


Table 2. Heart rate accuracy data across all subjects for the different activity types undertaken during the study: sample size, mean (SD) of each of the measured devices, mean (SD) of the difference between the device measurement and the reference standard, the mean relative difference (SD; %), the mean absolute difference (SD; %), and the correlation between the measures.

Heart rate (beats per minute) and measures	Fitbit	Garmin	Garmin + no button
Max test (treadmill): progressive exercise to volitional fatigue			
Pairs, N	7127	2037	476
Device, mean (SD)	129.6 (38.0)	139.6 (37.3)	112.2 (38.2)
Criterion, mean (SD)	137.2 (40.9)	144.7 (36.5)	122.3 (45.5)
Mean difference (SD)	−7.6 (13.6)	−5.1 (13.0)	−10.1 (21.5)
% mean relative error (SD)	−4.8 (10.3)	−3.3 (9.6)	−5.9 (16.6)
% mean absolute error (SD)	7.3 (11.8)	5.8 (8.4)	14.5 (10.1)
Concordance class correlation (95% CI)	0.92 (0.92-0.93)	0.93 (0.92-0.93)	0.84 (0.82-0.87)
Pearson correlation	.94	.94	.88
Max test (ergometer): progressive exercise to volitional fatigue			
Pairs, n	6375	1705	444
Device, mean (SD)	101.4 (31.2)	115.5 (34.0)	91.5 (21.3)
Criterion, mean (SD)	125.3 (32.7)	128.9 (33.3)	120.3 (34.1)
Mean difference (SD)	−23.8 (33.4)	−13.4 (25.6)	−28.8 (27.8)
% mean relative error (SD)	−16.4 (21.6)	−9.3 (17.0)	−20.6 (18.2)
% mean absolute error (SD)	17.9 (32.3)	11.8 (15.3)	22.9 (15.2)
Concordance class correlation (95% CI)	0.36 (0.34-0.37)	0.66 (0.62-0.68)	0.34 (0.29-0.39)
Pearson correlation	.46	.71	.58
Resistance exercise			
Pairs, n	17,420	5215	1200
Device, mean (SD)	105.9 (21.2)	112.9 (17.7)	91.8 (15.6)
Criterion, mean (SD)	114.4 (21.4)	119.5 (20.1)	104.6 (19.4)
Mean difference (SD)	−8.5 (14.4)	−6.5 (17.5)	−12.8 (17.4)
% mean relative error (SD)	−6.9 (12.0)	−4.2 (14.2)	−10.7 (14.9)
% mean absolute error (SD)	9.8 (12.1)	10.6 (10.4)	15.0 (10.7)
Concordance class correlation (95% CI)	0.72 (0.71-0.72)	0.54 (0.52-0.56)	0.4 (0.37-0.45)
Pearson correlation	.88	.9	.53
Daily chores and activities of daily living			
Pairs, n	14,883	3605	738
Device, mean (SD)	101.8 (20.5)	104.0 (22.0)	104.5 (20.8)
Criterion, mean (SD)	98.6 (20.8)	100.2 (21.8)	98.2 (17.0)
Mean difference (SD)	3.3 (15.2)	3.9 (17.4)	6.3 (18)
% mean relative error (SD)	3.3 (16.50)	5.6 (19.5)	7.4 (19.4)
% mean absolute error (SD)	11.4 (11.2)	13.0 (13.2)	14.0 (15.4)
Concordance class correlation (95% CI)	0.72 (0.71-0.73)	0.68 (0.66-0.69)	0.52 (0.47-0.57)
Pearson correlation	.73	.69	.56
Treadmill: intermittent high-intensity exercise			
Pairs, n	8105	3315	482
Device, mean (SD)	129.7 (28.0)	138.8 (26.9)	125.7 (38.1)
Criterion, mean (SD)	133.2 (30.6)	139.9 (26.3)	120 (35.4)
Mean difference (SD)	−3.5 (14.4)	−1.2 (11.9)	5.7 (33.5)

Heart rate (beats per minute) and measures	Fitbit	Garmin	Garmin + no button
% mean relative error (SD)	−1.7 (11.5)	−0.5 (9.4)	8.9 (33)
% mean absolute error (SD)	8.5 (10.0)	9.0 (6.0)	25.0 (23.3)
Concordance class correlation (95% CI)	0.87 (0.87-0.88)	0.90 (0.89-0.91)	0.58 (0.52-0.63)
Pearson correlation	.88	.9	.59
Ergometer: intermittent high-intensity exercise			
Pairs, n	7589	2440	193
Device, mean (SD)	110.6 (31.2)	110.9 (30.3)	100.4 (26.6)
Criterion, mean (SD)	127.0 (25.7)	131.2 (25.3)	131.2 (24.2)
Mean difference (SD)	−16.4 (27.2)	−20.3 (28.9)	−30.8 (27.4)
% mean relative error (SD)	−11.4 (35.7)	−14.3 (20.5)	−22.5 (19.8)
% mean absolute error (SD)	16.0 (24.4)	26.0 (17.6)	25.0 (13.4)
Concordance class correlation (95% CI)	0.47 (0.45-0.48)	0.37 (0.34-0.39)	0.24 (0.16-0.32)
Pearson correlation	.56	.47	.42

Figure 2. Percent relative error (RE) in heart rate (HR) across all the activities from all the devices tested. Percent error is calculated as (device measurement-reference standard) × 100/reference standard. The box-whisker plots indicate the error with the 25% quartile, median (50% quartile), and 75% quartile marked in each box plot. Gray horizontal dashed lines indicate the 5% error threshold, and the dotted lines indicate the 10% error threshold. ADLs: activities of daily living, HIIT-C: high-intensity interval training-Cycle ergometer, HIIT-T: high-intensity interval training-Treadmill, MAX-C: MAX-Cycle ergometer, MAX-T: MAX-Treadmill.

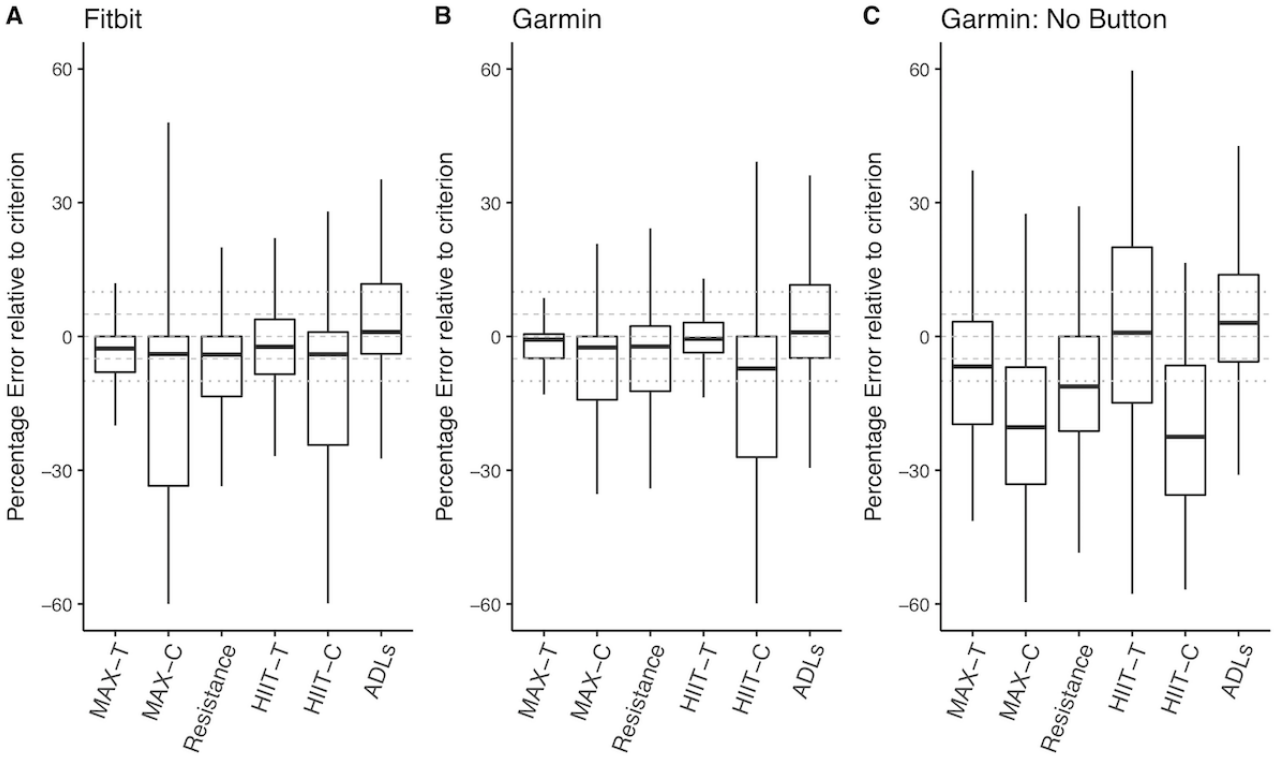


Figure 3. Scatter plots showing HR measurements from Fitbit and Garmin versus the reference standard Polar across all activities. Panel A shows the correlation plot comparing the Fitbit versus the Polar. Panel B shows the correlation plot comparing the Garmin (with activity mode indicated) versus the Polar. Panel C shows the correlation plot for a subset of the subjects comparing the Garmin (with activity mode not indicated) versus the Polar. ADLs: activities of daily living, HIIT-C: high-intensity interval training-Cycle ergometer, HIIT-T: high-intensity interval training-Treadmill, MAX-C: MAX-Cycle ergometer, MAX-T: MAX-Treadmill.

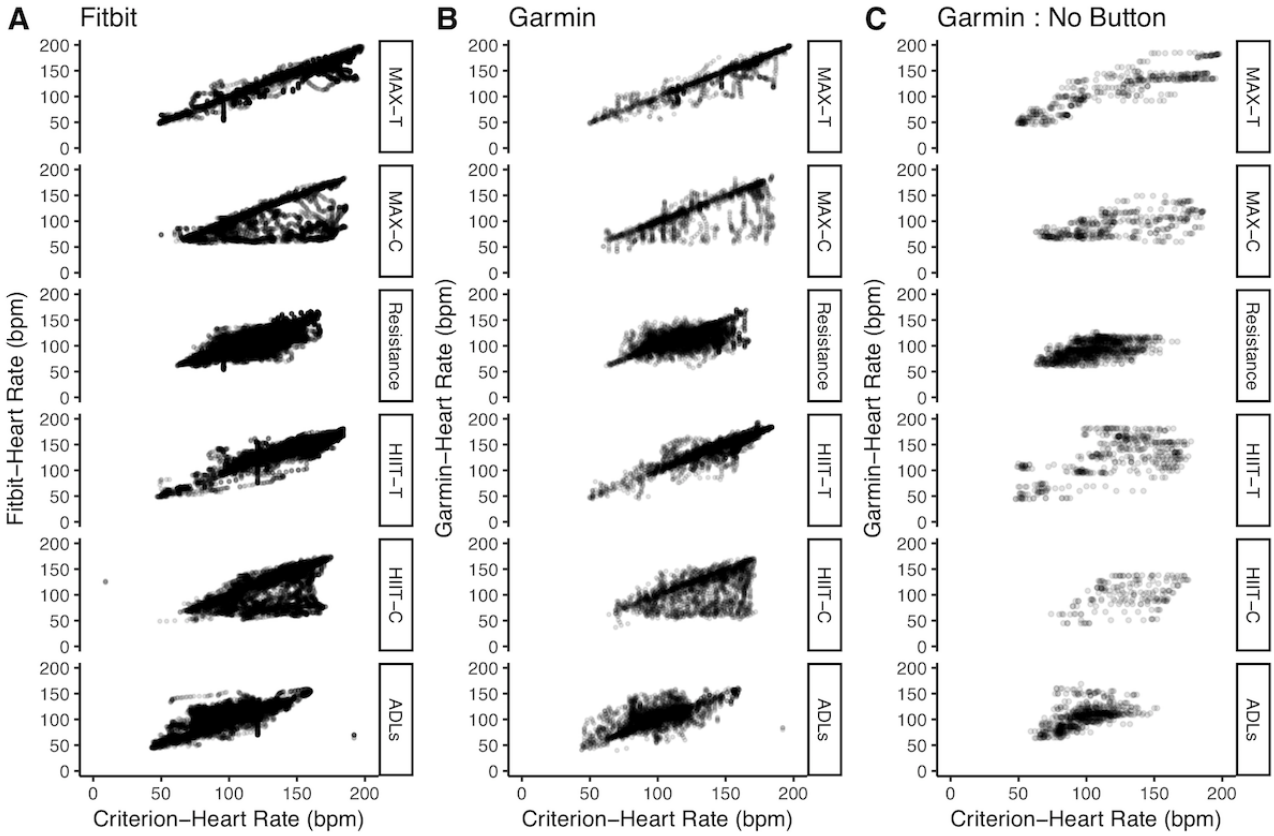


Figure 4. Bland-Altman plots showing heart rate measurements comparing Garmin and Fitbit relative to Polar for all data with activities indicated by color. Mean heart rate is shown on the x-axis, and the difference between the Garmin or Fitbit and the Polar heart rate is on the y-axis. The gray dotted line indicates the mean difference (bias) between the measurement, and the gray dashed lines indicate the limits of agreement. Panel A compares the Fitbit and the Polar. Panel B compares the Garmin (with activity indication) and the Polar. Panel C compares the Garmin (with no activity indication) and the Polar.

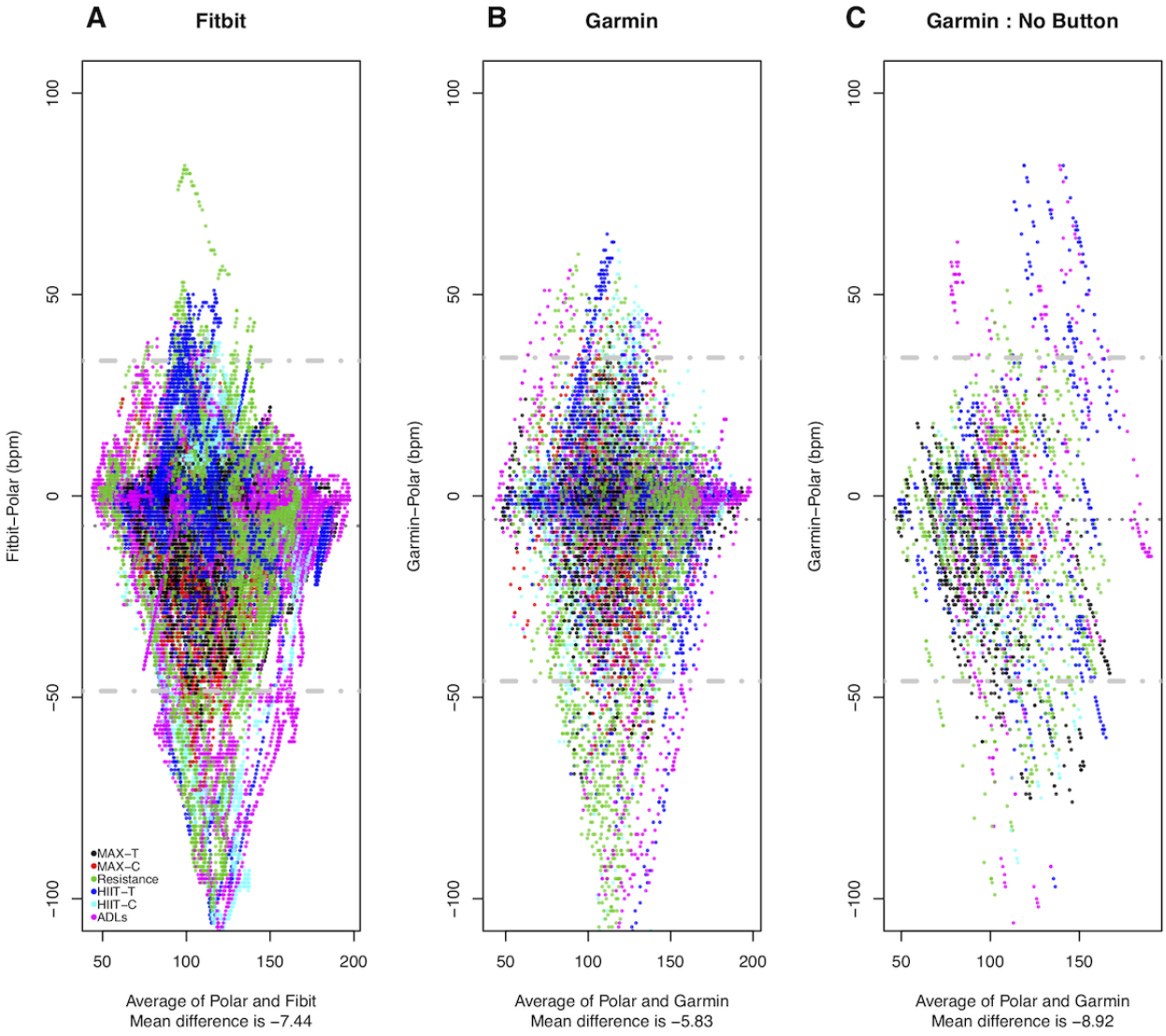
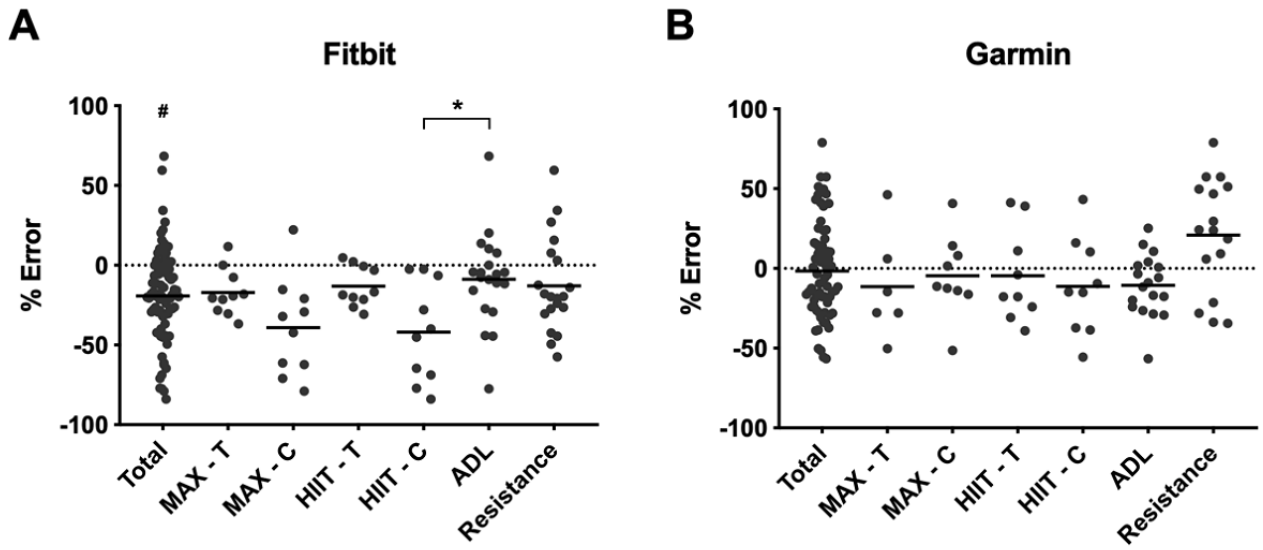


Table 3. Pooled energy expenditure data for the different types of activities undertaken during the study. Data are shown for each activity type. Sample size, mean (SD) of each of the measured device, mean (SD) of the difference between the device measurement and the reference standard, the mean relative difference (SD; %), the mean absolute difference (SD; %), and the correlation between the measures.

Energy expenditure (kcal) and measures	Fitbit	Garmin
Max test (treadmill): progressive exercise to volitional fatigue	**N=10**	**N=6**
Device, mean (SD)	192.1 (47.2)	216.5 (55.3)
Criterion, mean (SD)	237.3 (72.5)	260.5 (77.2)
Mean difference (SD)	−45.2 (44.4)	−44.0 (90.1)
% mean relative error (SD)	−17.0 (14.6)	−11.4 (33.7)
% mean absolute error (SD)	19.4 (11.0)	28.8 (17.2)
Pearson correlation	.81	.11
Max test (ergometer): progressive exercise to volitional fatigue	**N=10**	**N=9**
Device, mean (SD)	133.6 (77.6)	207.0 (48.7)
Criterion, mean (SD)	225.3 (74.7)	231.4 (76.5)
Mean difference (SD)	−91.7 (87.2)	−24.4 (63.9)
% mean relative error (SD)	−39.1 (30.6)	−4.5 (25.3)
% mean absolute error (SD)	43.5 (23.0)	18.9 (16.2)
Pearson correlation	.35	.56
Resistance exercise	**N=20**	**N=16**
Device, mean (SD)	130.2 (46.2)	179.8 (56.8)
Criterion, mean (SD)	153.1 (45.5)	155.2 (47.8)
Mean difference (SD)	−22.9 (44.0)	24.6 (56.6)
% mean relative error (SD)	−12.9 (29.7)	21.0 (35.7)
% mean absolute error (SD)	27.7 (15.9)	35.7 (19.7)
Pearson correlation	.54	.43
Daily chores and activities of daily living	**N=20**	**N=18**
Device, mean (SD)	103.5 (38.2)	100.6 (23.4)
Criterion, mean (SD)	114.4 (25.7)	114.8 (27.0)
Mean difference (SD)	−10.9 (39.4)	−14.3 (28.2)
% mean relative error (SD)	−8.8 (29.2)	−10.6 (19.3)
% mean absolute error (SD)	20.9 (21.8)	17.0 (13.7)
Pearson correlation	.29	.38
Treadmill: intermittent high-intensity exercise	**N=10**	**N=9**
Device, mean (SD)	211.1 (57.0)	226.9 (58.1)
Criterion, mean (SD)	246.6 (71.9)	249.7 (75.6)
Mean difference (SD)	−35.5 (34.6)	−22.8 (61.7)
% mean relative error (SD)	−13.1 (12.7)	−4.7 (29.3)
% mean absolute error (SD)	14.5 (10.9)	25.0 (3.4)
Pearson correlation	.88	.60
Ergometer: intermittent high-intensity exercise	**N=10**	**N=9**
Device, mean (SD)	128.2 (60.4)	205.8 (76.4)
Criterion, mean (SD)	232.8 (44.2)	234.9 (46.4)
Mean difference (SD)	−104.6 (83.8)	−29.1 (80.2)
% mean relative error (SD)	−41.9 (1.3)	−11.2 (30.8)
% mean absolute error (SD)	41.9 (31.3)	26.7 (17.0)

Energy expenditure (kcal) and measures	Fitbit	Garmin
Pearson correlation	–.26	.22

Figure 5. Percent relative error (RE) in energy expenditure (EE) across different exercise modalities for Fitbit (A) and Garmin (B). Negative bias in estimating EE is apparent across exercise modalities. The horizontal lines represent the mean. Asterisk indicates *P*=.03; # indicates *P*<.001 compared to Garmin. ADLs: activities of daily living, HIIT-C: high-intensity interval training-Cycle ergometer, HIIT-T: high-intensity interval training-Treadmill, MAX-C: MAX-Cycle ergometer, MAX-T: MAX-Treadmill.



MAPE (SD) for Garmin and Fitbit was 27.0% (SD 21.8) and 25.1% (SD 17.3), respectively. The lowest mean error in measuring EE was observed during ADL (−8.8% [SD 29.2]) for Fitbit and MAX-C (−4.5% [SD 25.3]) and HIIT-T (−4.7% [SD 29.3]) for Garmin. The highest error was observed during MAX-C (−39.1% [SD 30.6]) and HIIT-C (−41.9% [SD 31.3]) for Fitbit and resistance (21.0% [SD 35.7]) for Garmin. Figure 6 shows the relative error in EE for Fitbit and Garmin during all pooled treadmill and pooled cycle ergometer activities as scattered dot plots.

Both Fitbit and Garmin demonstrated negative bias when activities were performed on the treadmill (Fitbit: −15.1% [SD 13.5], Garmin: −7.4% [SD 30.1]; *P*=.18). For activities performed on the cycle ergometer, both devices displayed negative bias, but there was significantly higher mean error for Fitbit compared with Garmin (Fitbit: −40.5% [SD 30.2], Garmin: −7.9% [SD 27.6]; *P*<.001). Figure 7 shows the absolute percent error in EE during each activity as box-whisker plots for Fitbit and Garmin, compared with Cosmed-derived EE.

Garmin was significantly more accurate than Fitbit at estimating EE during MAX (Fitbit: 31.5% [SD 21.5], Garmin: 22.9% [SD 16.8]; *P*=.047) and all cycle ergometer activities (Fitbit: 42.7% [SD 26.8], Garmin: 22.8% [SD 16.6]; *P*=.03). Fitbit was significantly more accurate than Garmin at estimating EE during ADL (ADL: 20.9% [SD 21.8], ergometer: 42.7% [SD 26.8]; *P*=.02) and all treadmill activities (Treadmill: 16.9% [SD 10.9], ergometer: 42.7% [SD 26.8]; *P*=.003) compared with all activities performed on the cycle ergometer.

Spurious Heart Rate Measurements

During the early-phase testing of these devices, it was discovered that both devices would produce spurious HR measurements during periods of nonwrist use, such as when devices were stored in a backpack during commute. PPG sensors use a light source, commonly a group of light emitting diodes, to illuminate the tissue of the wrist, and the HR measurement is based on the differential reflection of the light as measured by the photodetector in response to the pulsatile nature of the blood perfusion in the superficial vessels. Under these working principles, if there is no light reflection from the surface, we suspected that the devices report HR measurements even if they are not *on body* (ie, spurious results). We performed a simple laboratory experiment to confirm this. Using a standard bench-top variable speed laboratory nutator (Fisher Sci # S06622), we simulated 3D wrist rotating motion at a fixed speed (22 rpm), and we recorded spurious HR results from both Garmin and Fitbit devices. The data and the experimental picture are shown in Figure 8.

Figure 6. Percent relative error (RE) in energy expenditure (EE) during the VO_2 peak test (MAX) and high-intensity interval training (HIIT) on the treadmill (A) and cycle ergometer (B) for Fitbit and Garmin. Negative bias in estimating EE is demonstrated by both devices during both modes of exercise, with the greatest mean error displayed by Fitbit during MAX and HIIT performed on the cycle ergometer. The horizontal lines represent the mean. # indicates *P*<.001.

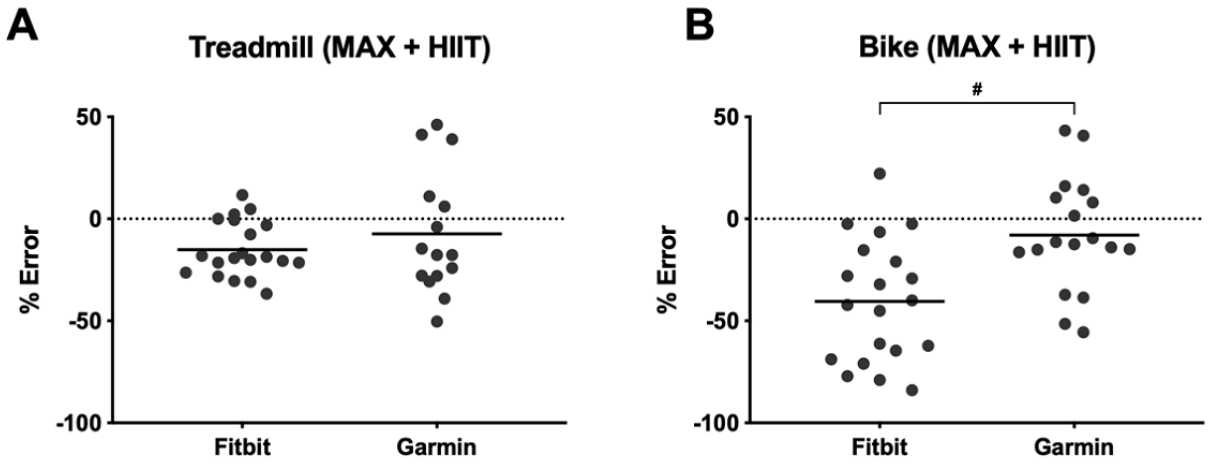


Figure 7. Absolute percent error in energy expenditure (EE) across different exercise modalities for Fitbit and Garmin. Each box-whisker plot consists of a box that extends from the 25% to the 75% quartile, with a line in the middle of the box representing the median (50% quartile). Each box has error bars that extend to the 5% and 95% quartiles, with outliers displayed with open circles. The *P* values listed on the right side display the difference in absolute percent error for EE between Fitbit and Garmin during each activity with italics indicating statistical significance. Asterisk and double asterisks indicate *P*=.02 and *P*=.003, respectively. ADL: activities of daily living, HIIT: high-intensity interval training.

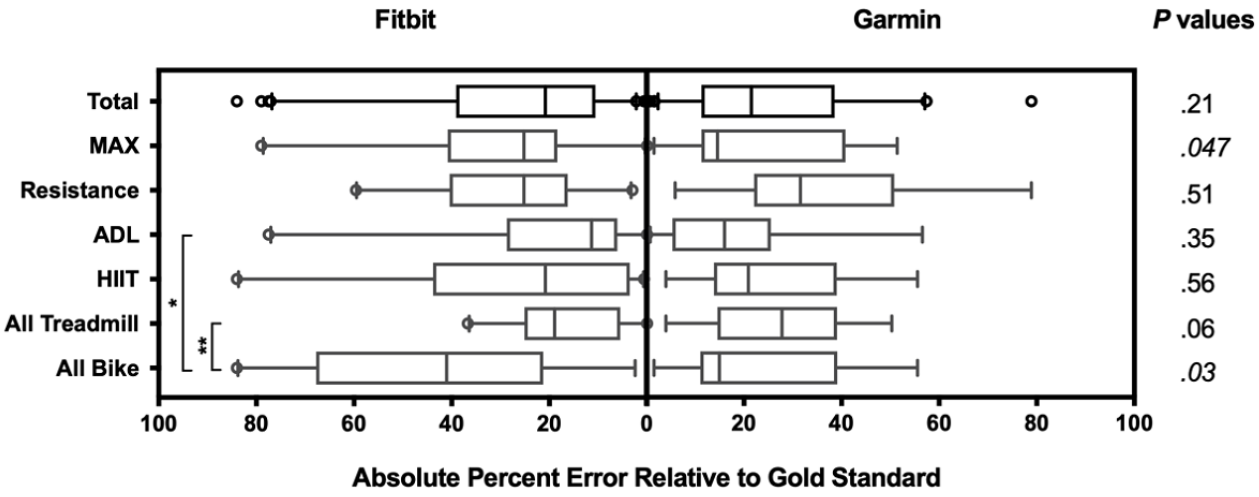
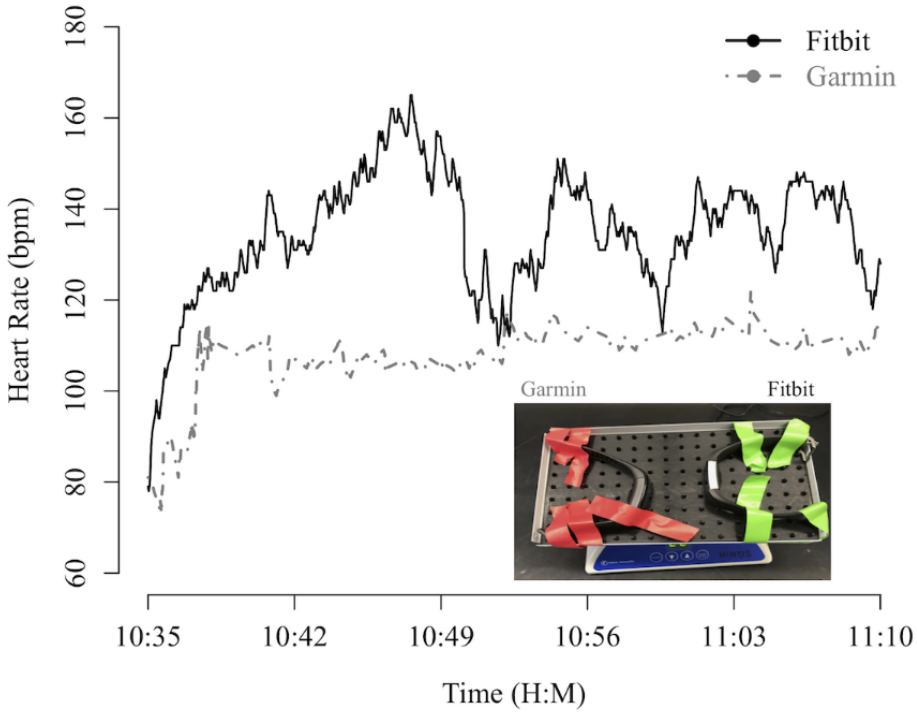


Figure 8. Spurious heart rate measured by the Garmin and Fitbit devices when placed on a shaker device (image of the experimental setup in the inset). H:M is hours:minutes.



Discussion

Principal Findings

This study examined the accuracy of 2 common wrist-worn, consumer-grade activity monitors for estimating HR and EE during a variety of nonsteady state activities. Similar to previous studies [13,28,30-32], we found reasonable accuracy in HR and EE estimations for these 2 devices under certain exercise conditions. Our findings are also in agreement with several prior studies that looked at HR and EE estimates across many different devices [16,17]; however, these 2 prior studies took measurements only at steady state conditions once HR had stabilized. A recent review by Bunn et al [33] showed that EE was generally underestimated by physical activity devices and that HR measurements were generally more accurate at rest or on a cycle ergometer as compared with treadmill. Dondzila et al [34] also looked at the Fitbit Charge HR and found that with aerobic exercise under laboratory conditions, the Fitbit Charge HR underestimated the HR compared with a Polar chest strap, with higher error at slower speeds. Jo et al [35] compared the Basis Peak and the Fitbit Charge HR with ECG and also found a negative bias of HR with respect to ECG measurements (−4.9 bpm for the Basis and −12.7 bpm for the Fitbit). In results presented in this paper, HR and EE measured by both the Garmin and Fitbit devices during the resistance exercise were similar to the measurements reported by Boudreaux et al [24]. Although, the resistance exercises were different, the intensity of the exercises was similar. There are 3 novel contributions from this study. First, we report HR accuracy in these activity monitors in modes not tested previously (eg, ADL and HIIT). Second, we show that HR accuracy as measured by these activity monitors is acceptable during low-intensity activities and high-intensity activities with repetitive wrist motion but that HR accuracy is poorer when there is no repetitive wrist motion and when any activity is at a high intensity (ie, ≥70% of maximal aerobic capacity). Prior research has suggested that PPG sensors used to measure the HR are liable to poor accuracy during activities with increased physical exertion or activities involving repetitive contractions of forearm skeletal muscles [36-38]. It has been suggested that during activities involving sustained muscle contractions or higher intensity exercises, the contact between the device's PPG sensor and skin is decreased, leading to a disruption in the signal quality and causing poor quality data [36,37]. Third, we show that HR, as measured by the Garmin, is significantly improved when the device is in the activity mode setting. As the HR measurement algorithm is proprietary to Garmin, we do not know why the accuracy is worse when activity mode is not indicated. It appears that the watch uses different HR measurement algorithms depending on the activity mode selected. It may be that the activity mode algorithms implement less smoothing than the nonactivity mode algorithm and are thereby designed to respond faster to rapid HR changes.

Although both activity monitors showed reasonable accuracy in HR, we did see differences between the 2 activity monitors in EE estimates across all activities, and both activity monitors correlated poorly with indirect calorimetry measures of EE. It is unclear why we found poor estimation of the EE. EE values are dependent on many anthropometric characteristics of the subject as well as the HR measurements [14]. We assume that the EE estimations provided by these devices are also utilizing this information, but these calculations are proprietary. According to the manufacturers, Fitbit's EE estimate includes both active calories and the basal metabolic rate (BMR), whereas

Garmin only reports active calories without BMR. Even with the inclusion of BMR in EE estimates, Fitbit still displayed a greater negative bias during most activities compared with Garmin. If EE estimates by Garmin included BMR, there would likely be greater accuracy in the EE values reported by these devices. At the time of testing, these activity monitors provide different ways to indicate the various types of activity such as running, stationary bike, strength training and "other," but there is not a clear indication for activities such as HIIT. Perhaps this is the reason for the high error rate recorded during these types of activities. As these consumer devices are constantly improved by the respective companies, the algorithms estimating EE should be improved or personalized to provide more accurate estimates. As these wearables transition from consumer reporting tools to clinical monitoring devices, a higher level of accuracy and precision is required. Clearly, the algorithms running on these wearables that estimate HR and EE are proprietary and can change without warning from the manufacturers, which poses further challenges for those wanting to integrate these devices into medical products. The onus of integrating these devices and assessing the level of accuracy and precision needed to make drug dosage decisions rests in the hands of those designing and evaluating medical algorithms.

Integrating these activity monitors into medical systems such as type 1 diabetes decision support systems or automated drug delivery systems in the future will require high fidelity data both from the HR signal and the EE estimates. The findings from this study point to shortcomings that could arise in both detecting activity and distinguishing the type of activity based on the HR signal. Although the mean error of the HR measurement was within the acceptable range for both devices, the range of the error was wider than anticipated. This issue and the inaccuracies associated with the EE data could lead to issues with estimating the intensity of the activity accurately. Additionally, short nonsteady state exercises such as a 10-second maximal sprint have been shown to influence the rapid change in glucose response to aerobic exercise [39], but findings from this study indicate that detecting these quick nonsteady exercises might be challenging for activity monitors to capture. We found spurious HR measurements when the activity monitor device is not worn on the wrist. Integration of these devices into a life-supporting drug delivery system must account for an on-wrist/off-wrist detection algorithms, which are currently not a part of the activity monitors evaluated. Another feature that could be integrated with further evaluation into a medical system is the exercise detection that is available on these devices. The Garmin device performed better when the exercise type was indicated through a button press on the watch. Future versions of these wearables are integrating automated exercise detection, and this is an area that should be further researched in terms of accuracy. Finally, if physical activity data are to be properly incorporated into medical systems including real-time drug delivery systems, access to the data in near real time (eg, every 5 min) would be important. In the automated insulin dosing scenario, for example, decisions would need to be made at the onset of exercise to prevent exercise-induced hypoglycemia. Currently, neither of these watches provide real-time access to their data streams. An approach to overcome some of the challenges associated with exercise detection and accuracy of detection would be to alert the individual before exercise dosing decisions are made. Effective integration of activity monitors is an active area of research in the medical community, and the findings from this study point to both the abilities and challenges associated with real-time monitoring and integrating into medical systems.

Limitations

Our study has a limitation in that we only tested 2 popular consumer-grade devices. The choice was based on the ubiquity of these sensors in the market, affordability, and potential to be easily integrated into existing medical system architectures through, for example, an API. Our current data and interpretations may be limited as we did not account for the skin color in our study. It has been reported that skin color could influence the accuracy of the HR measurement [16], and future studies should report the Fitzpatrick skin tone scale to account for this limitation. Another limitation of our study is that exercise was conducted in a laboratory setting as opposed to the real world. However, we attempted to capture several real-world ADLs to minimize this limitation, though these activities were also recorded within a lab. It would be important to do further investigations in real-world settings to corroborate our results. Another limitation was that HR measurements from the wearable devices were not compared against a true gold standard such as ECG.

Conclusions

We conducted a thorough assessment of 2 of the most popular low-cost consumer wrist-worn activity monitors during multiple exercise modalities and during daily activities. We found that during steady-state activities and during low-intensity activities, the HR measurements were within an acceptable error range (5%) but less accurate during higher intensity more dynamic activities that do not involve wrist motion. The EE estimates provided by these devices were inaccurate during all activities.

Acknowledgments

The authors would like to acknowledge the generous support from the Leona M and Harry B Helmsley Charitable Trust that funded the project through the grant "Exercise in Diabetes Initiative: The Effect of Exercise on Glycemic Control in Type 1 Diabetes (Pilot Study)."

Conflicts of Interest

None declared.

References

1. Manyika J, Chui M, Bisson P, Woetzel J, Dobbs R, Bughin J, et al. McKinsey. 2015. Unlocking the Potential of the Internet of Things URL: https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world [accessed 2018-10-29] [WebCite Cache ID 73XgkCKDc]
2. Haselton T. CNBC. 2017. Smart clothing and smartwatches will help double wearable market by 2021 URL: https://www.cnbc.com/2017/06/21/idc-wearables-market-to-double-by-2021.html [accessed 2018-10-29] [WebCite Cache ID 73XDKNzLk]
3. Riddell MC, Gallen IW, Smart CE, Taplin CE, Adolfsson P, Lumb AN, et al. Exercise management in type 1 diabetes: a consensus statement. Lancet Diabetes Endocrinol 2017 May;5(5):377-390. [doi: 10.1016/S2213-8587(17)30014-1] [Medline: 28126459]
4. Zaharieva D, Yavelberg L, Jamnik V, Cinar A, Turksoy K, Riddell MC. The Effects of Basal Insulin Suspension at the Start of Exercise on Blood Glucose Levels During Continuous Versus Circuit-Based Exercise in Individuals with Type 1 Diabetes on Continuous Subcutaneous Insulin Infusion. Diabetes Technol Ther 2017 Jun;19(6):370-378. [doi: 10.1089/dia.2017.0010] [Medline: 28613947]
5. Jacobs PG, El YJ, Reddy R, Resalat N, Branigan D, Condon J, et al. Randomized trial of a dual-hormone artificial pancreas with dosing adjustment during exercise compared with no adjustment and sensor-augmented pump therapy. Diabetes Obes Metab 2016 Dec;18(11):1110-1119 [FREE Full text] [doi: 10.1111/dom.12707] [Medline: 27333970]
6. Jacobs PG, Resalat N, El YJ, Reddy R, Branigan D, Preiser N, et al. Incorporating an Exercise Detection, Grading, and Hormone Dosing Algorithm Into the Artificial Pancreas Using Accelerometry and Heart Rate. J Diabetes Sci Technol 2015 Oct 05;9(6):1175-1184 [FREE Full text] [doi: 10.1177/1932296815609371] [Medline: 26438720]
7. Huyett LM, Ly TT, Forlenza GP, Reuschel-DiVirgilio S, Messer LH, Wadwa RP, et al. Outpatient Closed-Loop Control with Unannounced Moderate Exercise in Adolescents Using Zone Model Predictive Control. Diabetes Technol Ther 2017 Jun;19(6):331-339. [doi: 10.1089/dia.2016.0399] [Medline: 28459617]
8. Breton MD, Cherñavvsky DR, Forlenza GP, DeBoer MD, Robic J, Wadwa RP, et al. Closed-Loop Control During Intense Prolonged Outdoor Exercise in Adolescents With Type 1 Diabetes: The Artificial Pancreas Ski Study. Diabetes Care 2017 Dec;40(12):1644-1650. [doi: 10.2337/dc17-0883] [Medline: 28855239]
9. Turksoy K, Quinn LT, Littlejohn E, Cinar A. An integrated multivariable artificial pancreas control system. J Diabetes Sci Technol 2014 May;8(3):498-507 [FREE Full text] [doi: 10.1177/1932296814524862] [Medline: 24876613]
10. Breton MD, Brown SA, Karvetski CH, Kollar L, Topchyan KA, Anderson SM, et al. Adding heart rate signal to a control-to-range artificial pancreas system improves the protection against hypoglycemia during exercise in type 1 diabetes. Diabetes Technol Ther 2014 Aug;16(8):506-511 [FREE Full text] [doi: 10.1089/dia.2013.0333] [Medline: 24702135]
11. Lee J, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. Med Sci Sports Exerc 2014 Sep;46(9):1840-1848. [doi: 10.1249/MSS.0000000000000287] [Medline: 24777201]
12. Bai Y, Hibbing P, Mantis C, Welk GJ. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. J Sports Sci 2017 Dec 06:1-8. [doi: 10.1080/02640414.2017.1412235] [Medline: 29210326]
13. Chowdhury EA, Western MJ, Nightingale TE, Peacock OJ, Thompson D. Assessment of laboratory and daily energy expenditure estimates from consumer multi-sensor physical activity monitors. PLoS One 2017;12(2):e0171720 [FREE Full text] [doi: 10.1371/journal.pone.0171720] [Medline: 28234979]
14. Zakeri I, Adolph AL, Puyau MR, Vohra FA, Butte NF. Application of cross-sectional time series modeling for the prediction of energy expenditure from heart rate and accelerometry. J Appl Physiol (1985) 2008 Jun;104(6):1665-1673 [FREE Full text] [doi: 10.1152/japplphysiol.01163.2007] [Medline: 18403453]
15. Dasanayake IS, Bevier WC, Castorino K, Pinsker JE, Seborg DE, Doyle FJ, et al. Early Detection of Physical Activity for People With Type 1 Diabetes Mellitus. J Diabetes Sci Technol 2015 Jun 30;9(6):1236-1245 [FREE Full text] [doi: 10.1177/1932296815592409] [Medline: 26134831]
16. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. J Pers Med 2017 May 24;7(2) [FREE Full text] [doi: 10.3390/jpm7020003] [Medline: 28538708]
17. Abt G, Bray J, Benson A. Measuring Moderate-Intensity Exercise with the Apple Watch: Validation Study. JMIR Cardio 2018 Feb 28;2(1):e6. [doi: 10.2196/cardio.8574]
18. Helal L, Umpierre D, Moraes RS. High-intensity aerobic interval training improves aerobic fitness and HbA1c among persons diagnosed with type 2 diabetes: considerations regarding HbA1c starting levels and intervention design. Eur J Appl Physiol 2017 Nov;117(11):2365-2366. [doi: 10.1007/s00421-017-3682-1] [Medline: 28748370]
19. Wormgoor SG, Dalleck LC, Zinn C, Harris NK. Effects of high-intensity interval training on people living with type 2 diabetes: a narrative review. Can J Diabetes 2017 Oct;41(5):536-547. [doi: 10.1016/j.jcjd.2016.12.004] [Medline: 28366674]
20. Rooijackers HM, Wiegers EC, van der Graaf M, Thijssen DH, Kessels RP, Tack CJ, et al. A single bout of high-intensity interval training reduces awareness of subsequent hypoglycemia in patients with type 1 diabetes. Diabetes 2017 Dec;66(7):1990-1998 [FREE Full text] [doi: 10.2337/db16-1535] [Medline: 28420673]
21. Bai Y, Welk GJ, Nam YH, Lee JA, Lee J, Kim Y, et al. Comparison of consumer and research monitors under semistructured settings. Med Sci Sports Exerc 2015 Jul 6. [doi: 10.1249/MSS.0000000000000727] [Medline: 26154336]

22. Horton JF, Stergiou P, Fung TS, Katz L. Comparison of Polar M600 optical heart rate and ECG heart rate during exercise. Med Sci Sports Exerc 2017 Dec;49(12):2600-2607. [doi: 10.1249/MSS.0000000000001388] [Medline: 29135785]
23. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. J Sports Sci Med 2016 Sep;15(3):540-547 [FREE Full text] [Medline: 27803634]
24. Boudreaux BD, Hebert EP, Hollander DB, Williams BM, Cormier CL, Naquin MR, et al. Validity of wearable activity monitors during cycling and resistance exercise. Med Sci Sports Exerc 2018 Mar;50(3):624-633. [doi: 10.1249/MSS.0000000000001471] [Medline: 29189666]
25. Eparmedx. Physical Activity Readiness Questionnaire URL: http://eparmedx.com/wp-content/uploads/2013/01/January2018PARQPlus.pdf [accessed 2018-10-31] [WebCite Cache ID 73aq0EXsb]
26. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
27. Henriksen A, Haugen MM, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. J Med Internet Res 2018 Mar 22;20(3):e110 [FREE Full text] [doi: 10.2196/jmir.9157] [Medline: 29567635]
28. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, et al. Accuracy of wrist-worn heart rate monitors. JAMA Cardiol 2016 Oct 12. [doi: 10.1001/jamacardio.2016.3340] [Medline: 27732703]
29. R-project. Vienna, Austria: R Foundation for Statistical Computing; 2017. R: A Language and Environment for Statistical Computing URL: https://www.R-project.org/ [accessed 2018-10-29] [WebCite Cache ID 73XQR6GuD]
30. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. JMIR Mhealth Uhealth 2017 Mar 16;5(3):e34 [FREE Full text] [doi: 10.2196/mhealth.7043] [Medline: 28302596]
31. Stahl SE, An H, Dinkel DM, Noble JM, Lee J. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? BMJ Open Sport Exerc Med 2016;2(1):e000106 [FREE Full text] [doi: 10.1136/bmjsem-2015-000106] [Medline: 27900173]
32. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. PLoS One 2016;11(5):e0154420 [FREE Full text] [doi: 10.1371/journal.pone.0154420] [Medline: 27232714]
33. Bunn JA, Navalta JW, Fountaine CJ, Reece JD. Current state of commercial wearable technology in physical activity monitoring 2015-2017. Int J Exerc Sci 2018;11(7):503-515 [FREE Full text] [Medline: 29541338]
34. Dondzila C, Lewis C, Lopez J, Parker T. Congruent accuracy of wrist-worn activity trackers during controlled and free-living conditions. Int J Exerc Sci 2018;11(7):575-583 [FREE Full text]
35. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. J Sports Sci Med 2016 Sep;15(3):540-547 [FREE Full text] [Medline: 27803634]
36. Rafolt D, Gallasch E. Influence of contact forces on wrist photoplethysmography--prestudy for a wearable patient monitor. Biomed Tech (Berl) 2004;49(1-2):22-26. [doi: 10.1515/BMT.2004.005] [Medline: 15032494]
37. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007 Mar;28(3):R1-39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
38. Spierer DK, Rosen Z, Litman LL, Fujii K. Validation of photoplethysmography as a method to detect heart rate during rest and exercise. J Med Eng Technol 2015;39(5):264-271. [doi: 10.3109/03091902.2015.1047536] [Medline: 26112379]
39. Davey RJ, Bussau VA, Paramalingam N, Ferreira LD, Lim EM, Davis EA, et al. A 10-s sprint performed after moderate-intensity exercise neither increases nor decreases the glucose requirement to prevent late-onset hypoglycemia in individuals with type 1 diabetes. Diabetes Care 2013 Dec;36(12):4163-4165 [FREE Full text] [doi: 10.2337/dc12-2198] [Medline: 24130362]

Abbreviations

ADLs: activities of daily living
BMR: basal metabolic rate
BTLE: Bluetooth low energy
EE: energy expenditure
HIIT: high-intensity interval training
HIIT-C: high-intensity interval training-Cycle ergometer
HIIT-T: high-intensity interval training-Treadmill
HR: heart rate
MAPE: mean absolute percent error
MAX-C: MAX-Cycle ergometer
MAX-T: MAX-Treadmill
MET: metabolic equivalent
PPG: photoplethysmography
RPE: Rating of Perceived Exertion
VO_2: oxygen consumption

RE: mean relative error

Edited by S Kitsiou; submitted 19.03.18; peer-reviewed by G Dominick, I Mircheva, B Nelson; comments to author 23.04.18; revised version received 17.06.18; accepted 05.09.18; published 10.12.18.

Please cite as:
Reddy RK, Pooni R, Zaharieva DP, Senf B, El Youssef J, Dassau E, Doyle III FJ, Clements MA, Rickels MR, Patton SR, Castle JR, Riddell MC, Jacobs PG
Accuracy of Wrist-Worn Activity Monitors During Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study
JMIR Mhealth Uhealth 2018;6(12):e10338
URL: https://mhealth.jmir.org/2018/12/e10338/
doi:10.2196/10338
PMID:30530451

Original Paper

The Accuracy of Smart Devices for Measuring Physical Activity in Daily Life: Validation Study

Laurent Degroote[1,2], MSc; Ilse De Bourdeaudhuij[1], PhD; Maïté Verloigne[1], PhD; Louise Poppe[1,2], MSc; Geert Crombez[2], PhD

[1]Physical Activity & Health, Department of Movement and Sports Sciences, Ghent University, Ghent, Belgium

[2]Health Psychology, Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

Corresponding Author:
Laurent Degroote, MSc
Physical Activity & Health
Department of Movement and Sports Sciences
Ghent University
Watersportlaan 2
Ghent, 9000
Belgium
Phone: 32 9 264 62 99
Email: laurent.degroote@ugent.be

Abstract

Background: Wearables for monitoring physical activity (PA) are increasingly popular. These devices are not only used by consumers to monitor their own levels of PA but also by researchers to track the behavior of large samples. Consequently, it is important to explore how accurately PA can be tracked via these devices.

Objectives: The aim of this study was, therefore, to investigate convergent validity of 3 Android Wear smartwatches—Polar M600 (Polar Electro Oy, Kempele, Finland), Huawei Watch (Huawei Technologies Co, Ltd, Shenzhen, Guangdong, China), Asus Zenwatch3 (AsusTek Computer Inc, Taipei, Taiwan)—and Fitbit Charge with an ActiGraph accelerometer for measuring steps and moderate to vigorous physical activity (MVPA) on both a day level and 15-min level.

Methods: A free-living protocol was used in which 36 adults engaged in usual daily activities over 2 days while wearing 2 different wearables on the nondominant wrist and an ActiGraph GT3X+ accelerometer on the hip. Validity was evaluated on both levels by comparing each wearable with the ActiGraph GT3X+ accelerometer using correlations and Bland-Altman plots in IBM SPSS 24.0.

Results: On a day level, all devices showed strong correlations (Spearman *r*=.757-.892) and good agreement (interclass correlation coefficient, ICC=.695-.885) for measuring steps, whereas moderate correlations (Spearman *r*=.557-.577) and low agreement (ICC=.377-.660) for measuring MVPA. Bland-Altman revealed a systematic overestimation of the wearables for measuring steps but a variation between over- and undercounting of MVPA. On a 15-min level, all devices showed strong correlations (Spearman *r*=.752-.917) and good agreement (ICC=.792-.887) for measuring steps, whereas weak correlations (Spearman *r*=.116-.208) and low agreement (ICC=.461-.577) for measuring MVPA. Bland-Altman revealed a systematic overestimation of the wearables for steps but under- or overestimation for MVPA depending on the device.

Conclusions: In sum, all 4 consumer-level devices can be considered accurate step counters in free-living conditions. This study, however, provides evidence of systematic bias for all devices in measurement of MVPA. The results on a 15-min level also indicate that these devices are not sufficiently accurate to provide correct real-time feedback.

(JMIR Mhealth Uhealth 2018;6(12):e10972) doi:10.2196/10972

KEYWORDS

physical activity; fitness trackes; accelerometry

Introduction

Background

Physical inactivity is one of the major risk factors for mortality worldwide, causing an estimated 3.2 million deaths (6%) [1]. It accounts for approximately 21% to 25% of breast and colon cancers, 27% of type 2 diabetes, and 30% of burden because of ischemic heart disease [2,3]. It is hence recommended to perform a sufficient level of physical activity (PA). Physical activity is defined as "any bodily movement produced by skeletal muscles that require energy expenditure" [4]. PA can be classified according to the intensity of the activity using metabolic equivalents (METs). MET is the ratio of a person's working metabolic rate relative to their resting metabolic rate. One MET is defined as the energy cost in rest and is equivalent to a caloric consumption of 1 kcal/kg/hour. It is estimated that compared with sitting, a person's caloric consumption is more than 1.6 times higher and less than 3 times higher when being lightly active (1.6-3 METs), 3 to 6 times higher when being moderately active (3-6 METs), and more than 6 times higher when being vigorously active (>6 METs) [5]. Adults aged 18 to 64 years should accumulate at least 150 min of moderate-intensity aerobic PA throughout the week or do at least 75 min of vigorous-intensity aerobic PA throughout the week or an equivalent combination of moderate and vigorous intensity activity [5,6]. Another recommendation is to take at least 10,000 steps a day [7,8]. Nevertheless, 58% of the global population does not meet either of these recommendations [9].

Increasing the level of PA in the general population has proven notoriously difficult [10]. Scientists and practitioners have turned to behavior change theories to better understand the process of change and to better design interventions. Among various behavior change techniques, self-monitoring of the PA [11,12], has proven effective in changing PA levels. Consumer-level devices, also referred to as wearables, are increasingly used for the monitoring of PA [13]. They have built-in sensors to track and quantify daily movement [14].

Various wearables exist, and we can distinguish between activity trackers and smartwatches. Activity trackers (eg, Fitbit Flex, Misfit Shine, Garmin Vivosmart, and Xiaomi MiBand) are specifically built to track activity levels. Smartwatches (eg, Apple Watch, Samsung Gear, and Huawei Watch) also track activity levels but include other functions as well (eg, surfing the Web, receiving and answering mails or calls, playing music, and using the global positioning system). Furthermore, smartwatches allow downloading of apps and can be readily synchronized with a mobile phone. Smartwatches, therefore, have the potential to serve as a platform for app developers. They also have the potential to transform health care by supporting or evaluating health in everyday living because they (1) are familiar to most people; (2) are increasingly available as a consumer device; (3) enable near real-time continuous monitoring of PA and physiological measures; (4) support tailored messaging and reminders; (5) enable communication between patients, family members, and health care providers; and (6) allow for in situ mini-surveys and behavior verification based on sensor-based measure [15]. As wearables, both activity trackers and smartwatches, are increasingly popular not only with consumers but also with researchers [16], it is important to determine their accuracy for measuring PA variables such as step counts and minutes of MVPA.

Until now, only activity trackers have been scrutinized for their validity [17-23]. These studies found that most activity trackers (Fitbit Flex, Fitbit Zip, Fitbit One, Fitbit Charge HR, Jawbone Up, Nike+ Fuelband SE, Misfit Shine, and Withings Pulse) are valid for measuring steps but to a lesser extent, for measuring MVPA. For smartwatches, the validity for measuring PA variables (the number of steps and time spent in MVPA) has not been investigated. This is partly because of the recent rise in these devices: Up until 2014, about half of devices on the market were smartwatches. In 2015 and 2016, smartwatches represented 59.3% (143/241) of new devices on the market, whereas fitness trackers represented 40.7% (98/241) [24]. Furthermore, there is also a need for validation of wearables (both activity trackers and smartwatches) at a small time-scale. To our knowledge, all validation studies using activity trackers investigate validity on a daily level; however, validation using a smaller time-scale (eg, 15 min) is warranted. Increasingly, individual-focused interventions are developed that are based on real-time feedback. Examples are Just-In-Time adaptive interventions (JITAIs), which are the interventions that provide the right type and amount of support at the right time by adapting to an individual's changing internal and contextual state. By providing this personally tailored support, interventions can be more effective in guiding users toward a physically active lifestyle [25]. Due to the internal sensors, the larger screen, and the fact that the device can be consulted constantly as they are worn on the wrist, smartwatches have the potential to serve as a platform for a JITAI. Notwithstanding the potential of smartwatches for JITAIs, smartwatches should be accurate in measuring physical active or inactive behavior during a short time duration [25,26].

For example, when users engage in a 15-min jog, the device has to be able to correctly categorize this behavior as 15 min of MVPA. On the basis of this measurement, the appropriate intervention component is to give real-time feedback to the user that he or she is doing well without giving other suggestions for more PA. However, when the user is not physically active for 15 min, the device has to be able to correctly categorize this as 15 min of physical inactivity. On the basis of this measurement, the appropriate intervention component is to provide real-time feedback in the form of a tailored suggestion to the user to engage in more PA.

Objectives

The aim of this study was, therefore, to validate wearables in an adult population on both a day level as well as a 15-min level in free-living situations. We opted for a 15-min level because this is the smallest time level measured by the tested smartwatches. We opted for a validation in free living because this increases the external validity of our findings for use of wearables in daily life. We investigated convergent validity of 3 Android Wear smartwatches (Polar M600, Huawei Watch, and Asus Zenwatch3) and 1 activity tracker (Fitbit Charge). The number of steps and the time spent in MVPA measured by

consumer-level devices was compared directly with the measurements of an ActiGraph GT3X+ accelerometer.

Methods

Participants

In this study, 36 healthy participants (50% male; mean age 39.43 years, SD 17.77) aged between 20 and 65 years and living in the area of Ghent (Belgium) were recruited using purposeful sampling. The inclusion criteria were having no current physical limitations, medical conditions, or psychiatric conditions. Before participants were selected, they completed the International Physical Activity Questionnaire (IPAQ, long 7d version) to assess their current level of PA. This procedure allowed us to have variation in the participants' activity levels. The IPAQ was chosen for 2 reasons. First, a self-report measure was used for practical reasons. The self-report measurement allowed us to assess the current PA of people by letting them fill out a 10-min questionnaire, which makes it a very time-efficient measurement as opposed to objective measurement. Second, earlier research indicated that IPAQ is a reasonably reliable valid measurement tool for measuring habitual PA [27,28]. The International Physical Activity Questionnaire–Long Form (IPAQ-LF, last 7 days) asks participants to report the frequency and duration of activities in the last 7 days. Activities were classified into the domains of occupation, transportation, household, and leisure for each category of walking, moderate-intensity PA (MPA), and vigorous-intensity PA (VPA). Weekly and daily minutes of total PA, MPA, and VPA were computed.

On the basis of this assessment, we included 18 participants (50% male) who met the guideline of 30-min MVPA per day and 18 participants (50% male) who did not meet this guideline. All participants read and signed an informed consent form. The study protocol was approved by the ethics committee of the University hospital of Ghent (B670201731732).

Instruments

Convergent Measure

The ActiGraph GT3X+ (Actigraph, Pensicola, FL, USA), a triaxial accelerometer was used as reference or convergent measure. The ActiGraph GT3X+ has been found to be reliable and valid. The GT3X+ is valid for measuring step counts compared with direct observation by trained observers [29-31] and for MVPA compared with indirect calorimetry [32,33]. Accelerometer data were initialized, downloaded, and processed by using ActiLife version 5.5.5-software (ActiGraph, Fort Walton Beach, FL, USA). The Freedson Adult (1998) cut-points were used to categorize PA measured by the ActiGraph accelerometer (sedentary activity=0-99 counts/min, light activity=100-1951 counts/min, moderate activity=1952-5723 counts/min, and vigorous activity ≥5724 counts/min) [32]. A 15-s epoch was used when downloading the data.

Wearables

We tested 4 wearables: Fitbit Charge, Polar M600, Huawei Watch, and Asus Zenwatch 3. Fitbits are one of the most popular activity trackers on the market. Smartwatches from Polar, Huawei, and Asus were selected because they use the Android Wear platform that has a significant market share (18% during Quarter 1 2017) and provides easy opportunities to program smartwatches and develop apps [34]. Polar M600, Huawei Watch, and Asus Zenwatch were selected because of their potential for electronic health interventions at the time of data collection (beginning of 2017). All 4 devices measure steps and a specific variable that quantifies the degree of PA. For the Fitbit, we used the variable *active minutes*, which is divided into light active, fairly active, and very active minutes. To approach the MVPA variable, fairly and very active minutes were summed. For the Android Wear smartwatches, we used the variable *active time*, which is calculated by summing the time spent on various activities (walking, running, and biking) that are all covered by the definition of MVPA (>3.0 MET) [1]. As all the devices set a goal of 30-min PA per day (similar to the MVPA recommendations for adults), we assumed that the measured variable corresponded to MVPA as measured by the ActiGraph. However, specific information regarding intensity cut-points is not publicly available. All Fitbit data were exported in an XLS (Microsoft Excel) format using the Fitbit Dashboard Web app. Every minute was categorized as sedentary, lightly active, fairly active, or very active. Afterward, the data per minute were converted to data per 15 min. Data from the Android Wear smartwatches were exported in a CSV (comma-separated values) format from Google Fit using Google Take Out. Every 15 min, it was shown how many seconds were spent on various activities (walking, running, biking, and tilting)

Free-Living Protocol

As it was neither feasible nor comfortable to wear 4 wearables at the same time; participants were instructed to simultaneously wear 2 of the devices and the ActiGraph accelerometer for 2 consecutive days and then the other 2 wearables and the accelerometer for another 2 consecutive days. Between these 2 periods of 2 days, there was always a gap of 1 day on which devices were transferred from one participant to another. The devices were worn during all waking hours, except during water-based activities. All participants wore all 4 different wearables. All possible combinations of 2 wearables (a total of 6) were randomly assigned to the participants. Each combination was tested for 24 days in total, and each device was tested for 72 days. The ActiGraph GT3X+ was fitted to the right side of the participants' waist, and the wearables were placed on the nondominant wrist. Furthermore, participants were instructed to keep a short diary in which they wrote down when they put on the devices and when and why they took them off.

Statistical Analysis

Only days with valid data of the ActiGraph were included in the analysis. A valid day was defined as a 24-hour period in which at least 10 hours of data wear time was recorded. Nonwear time was analyzed as a run of zero counts lasting more than 60 min [35,36]. Analyses were performed using IBM SPSS Statistics version 24.0. All analyses were performed on a day level as well as a 15-min level. First, the correlation between the wearables and the ActiGraph accelerometer for measuring steps and MVPA was examined by calculating the Spearman *r* and ICC (absolute agreement, 2-way random, single measures,

and 95% CI). Both analyses were conducted to take into account the possible systematic difference between the measurements, which is taken into account by the ICC, but not by the Spearman correlation. The following cut-off values were used to interpret the Spearman correlation: r<.20=very weak; .20 to .39=weak; .40 to .59=moderate; .60 to .79=strong; and .80 to 1.0=very strong [37]. The cut-off values to interpret the ICC were <.60=low; .60 to .75=moderate; .75 to .90=good; and >.90=excellent [38]. Second, to examine the level of agreement between the wearables and the convergent measure, Bland-Altman plots were constructed with their associated limits of agreement.

Results

Participants' Characteristics

Participants' characteristics are presented in Table 1. All 36 participants wore the devices as planned. Some data were lost because of device malfunction (2 days MVPA or steps for Asus) and participant error such as not charging the device (4 days MVPA or steps for Asus, Polar, Fitbit, and Huawei). No data were lost from the ActiGraph GT3X+ accelerometers.

Validation at a Day Level

In Table 2, the mean steps and mean minutes of MVPA (SD) per day are presented for all wearables and ActiGraph accelerometer. Moreover the statistical significance (P value) of the difference between the ActiGraph accelerometer and the wearables is presented. This table shows that every wearable overestimated the number of steps per day (not significant for Asus). For MVPA, Huawei, Asus, and Fitbit underestimated, whereas Polar overestimated the number of minutes of MVPA (not significant for Fitbit).

Table 1. Participant characteristics (N=36).

Characteristic	Minimum-maximum[a]	Mean (SD)[a]
Age (years)	20-65	39.43 (17.77)
Height (cm)	150-186	172.28 (8.22)
Weight (kg)	42-98	68.43 (12.09)
BMI[b] (kg/m²)	17.51-32.00	23.00 (3.50)
MVPA[c] (min/day)[a]	0-178.29	43.70 (42.02)

[a]Based on the International Physical Activity Questionnaire data.

[b]BMI: body mass index.

[c]MVPA: moderate to vigorous physical activity.

Table 2. Mean steps and minutes of moderate to vigorous physical activity per day measured by Huawei, Asus, Polar, and Fitbit and the corresponding ActiGraph measurements and statistical significance (P value) of the difference between the ActiGraph accelerometer and the wearables.

Variable	Wearable, mean (SD)/day	ActiGraph accelerometer, mean (SD)/day	P value
Huawei			
Steps	8625 (4514)	7148 (3761)	.02
MVPA[a] (min)	27.24 (31.59)	36.97 (27.63)	.07
Asus			
Steps	7662 (4380)	7082 (4148)	.42
MVPA (min)	27.14 (33.18)	39.53 (36.33)	<.001
Polar			
Steps	10,864 (7517)	7234 (4076)	<.001
MVPA (min)	59.77 (62.94)	36.51 (28.31)	.03
Fitbit			
Steps	9127 (5381)	7459 (3661)	.004
MVPA (min)	35.47 (49.18)	41.98 (34.40)	.39

[a]MVPA: moderate to vigorous physical activity.

Table 3. Correlation coefficients, intraclass correlation coefficients, and 95% CI of the measurements at a day level.

Variable	Spearman *r* (95% CI)	ICC[a] (95% CI)
Huawei		
Steps	.892[b] (0.779-0.930)	.885[b] (0.822-0.926)
MVPA[c]	.577[b] (0.346-0.752)	.606[b] (0.433-0.736)
Asus		
Steps	.757[b] (0.605-0.881)	.723[b] (0.590-0.817)
MVPA	.557[b] (0.349-0.724)	.517[b] (0.324-0.669)
Polar		
Steps	.847[b] (0.659-0.937)	.695[b] (0.553-0.798)
MVPA	.529[b] (0.292-0.724)	.377[b] (0.159-0.560)
Fitbit		
Steps	.885[b] (0.798-0.939)	.792[b] (0.686-0.866)
MVPA	.564[b] (0.358-0.738)	.660[b] (0.504-0.774)

[a]ICC: interclass correlation coefficient

[b]*P*<.001.

[c]MVPA: moderate to vigorous physical activity.

Correlations

For measuring steps on a day level, all wearables showed strong to very strong correlations based on the Spearman *r* and moderate to good agreement based on the ICC. Correlations between the MVPA levels from the wearables and the MVPA levels from the ActiGraph accelerometer were moderate based on the Spearman *r*. Agreements for MVPA between the wearables and the ActiGraph accelerometer were low. The correlation coefficients, ICC values, and associated 95% CI are shown in Table 3. The correlations are also illustrated in Figure 1. This figure shows that the scatter of the points around the line, reflecting the perfect agreement between measurements is larger for measuring MVPA than for measuring steps.

Level of Agreement

Bland-Altman plots indicated the differences between the ActiGraph accelerometer and the wearables (y-axis) against the average number of steps or number of minutes of MVPA of the 2 devices (x-axis). Mean differences with the ActiGraph accelerometer and the limits of agreement for each wearable are presented in Figures 2 and 3. A positive value of the mean difference indicates an underestimation of the wearable compared with the golden standard, and a negative value indicates an overestimation. The systematic differences (mean differences) and the range between the upper and lower limits of agreement are important to make a statement about the validity of these wearables. The broader the range between the lower and the upper limit, the less accurate the measurements are. All wearables showed broad limits of agreement. For measuring steps, the plots (presented in Figure 2) showed the narrowest limits for Huawei (7759 steps) and the broadest limits for Polar (18,379 steps). The Bland-Altman plots for measuring MVPA are presented in Figure 3. For measuring MVPA, the narrowest limits were found for Fitbit (94 min), and the broadest limits were found for Polar (212 min).

Figure 1. Correlations between the activity estimates per day from the wearables and the ActiGraph. Spearman *r* values and intraclass correlation coefficient values denote the correlation for measuring moderate-to-vigorous physical activity or steps between the wearable and the ActiGraph. a) P<.001. MVPA: moderate to vigorous physical activity; ICC: intraclass correlation coefficient.

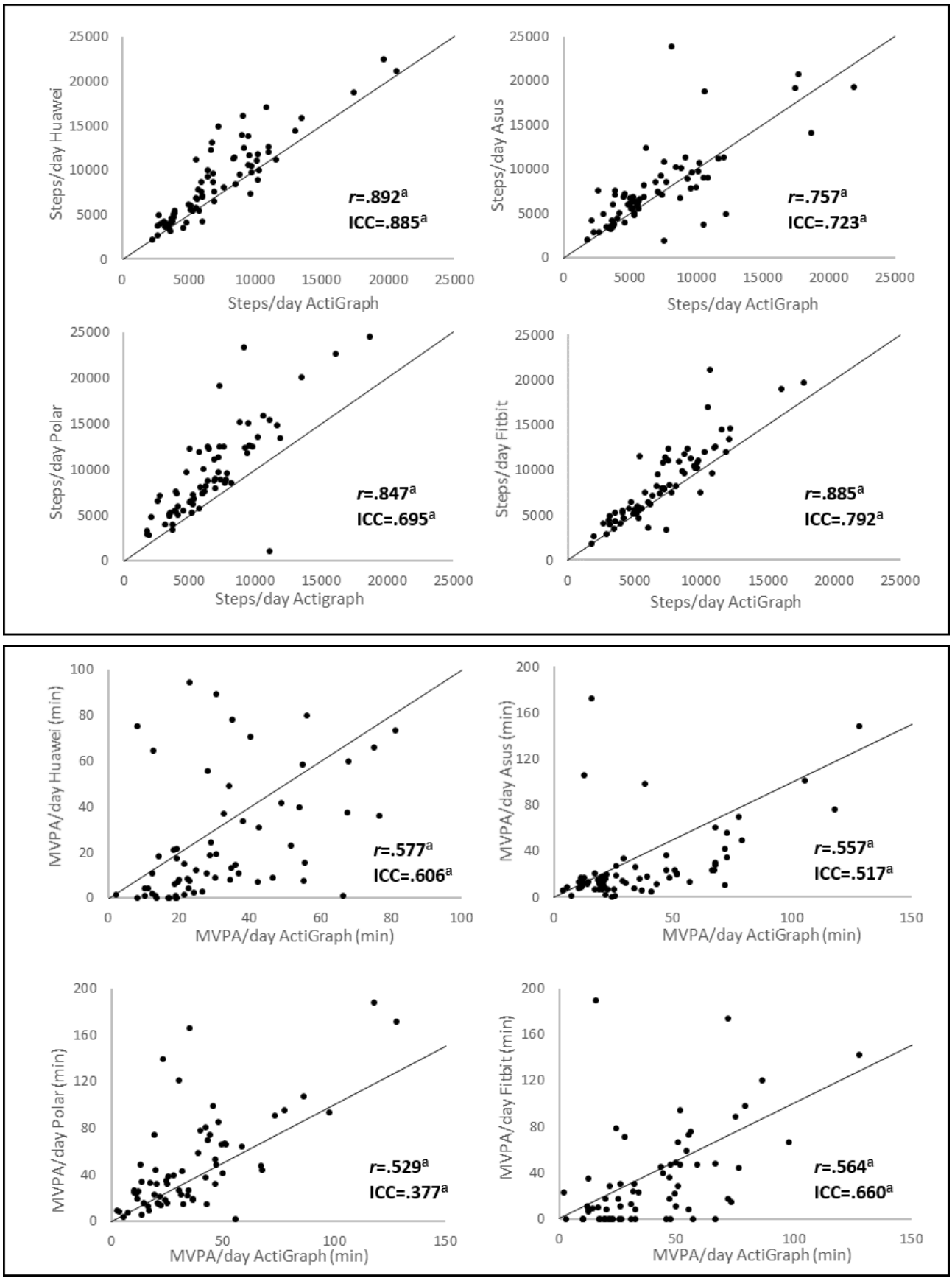
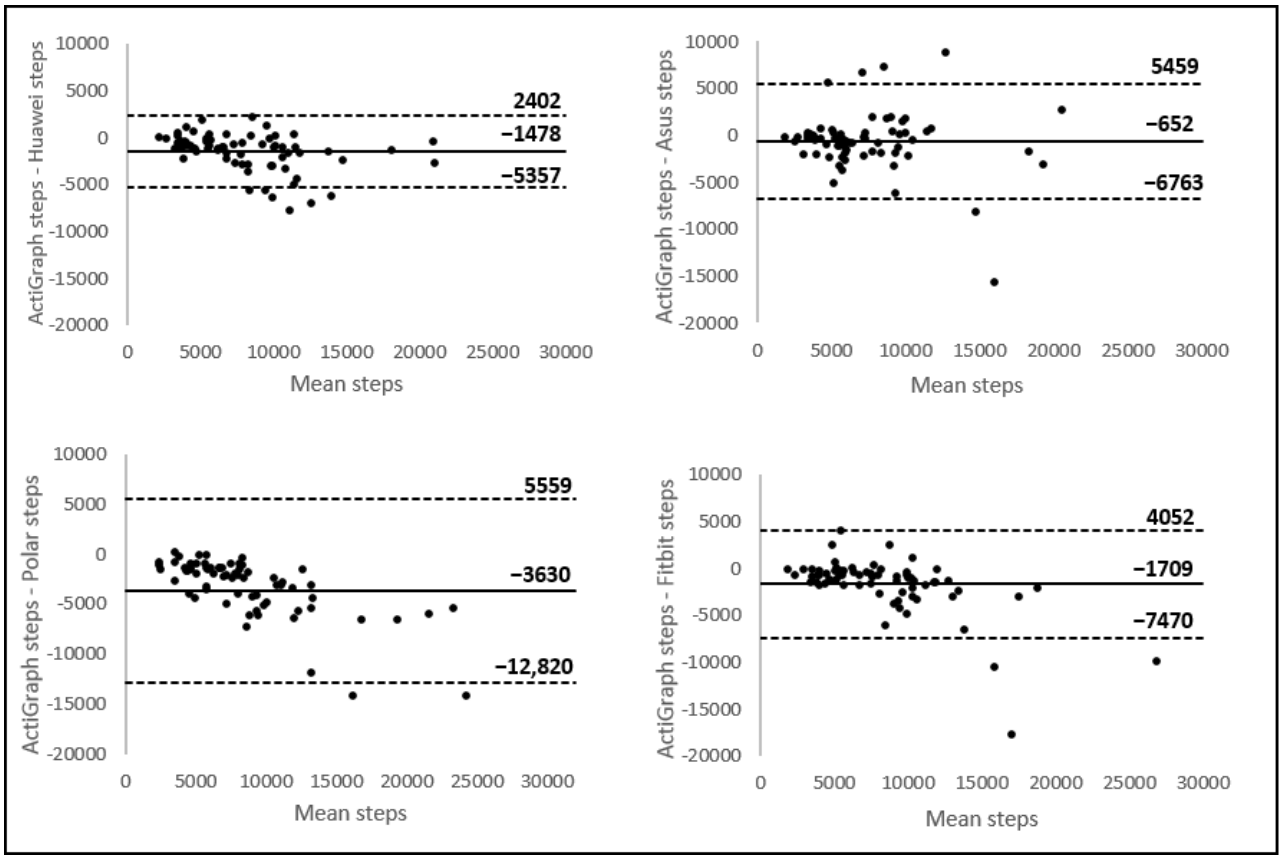


Figure 2. Bland-Altman plots of the wearables. The middle line shows the mean difference (Positive values indicate an underestimation of the wearable and negative values indicate an overestimation) between the measurements of steps of the wearables and the ActiGraph, and the dashed lines indicate the limits of agreement (1.96 × SD of the difference scores).



Validation at a 15-Minute Level

In Table 4, the mean steps per 15 min and mean minutes of MVPA per 15 min are presented for all devices. Moreover, the statistical significance (*P* value) of the difference between the measurements of the ActiGraph accelerometer and the wearables is presented. The results are displayed for (1) all 15-min time periods (including those with no MVPA) and (2) only the 15-min time periods in which MVPA was displayed with and without data revealing no MVPA. We opted to also present the latter to avoid distortion of the results. As users did not perform any PA during most periods of the day, a good agreement would be easy to obtain because of the many zero measurements by both measuring devices (wearable and ActiGraph accelerometer). In addition, this would reflect the validity of measuring physical inactivity rather than validity of measuring PA. Table 4 shows that every wearable device overestimated the number of steps per 15 min (all significant). For MVPA, Asus underestimated, whereas Huawei, Polar, and Fitbit overestimated the number of minutes of MVPA (not significant for Asus).

Correlation

All devices showed strong to very strong correlation based on the Spearman *r* and good agreement based on the ICC for measuring steps. For measuring MVPA (only including the data without zeros), correlations between readings from the wearables and the ActiGraph accelerometer were very weak to weak based on the Spearman *r*. Agreement between all the wearables and the ActiGraph accelerometer was low. The correlation coefficients, ICC values, and associated 95% CIs are shown in Table 5. The correlations are also illustrated in Figure 4. This figure revealed a systematic difference between the measurements of the wearables and the ActiGraph. The systematic difference increased as the number of steps or number of minutes MVPA increased. For example, an overestimation of 20% results in a difference of 200 steps on a day with 1000 steps. On a day, however, with 8000 steps, the difference between the measurements is 1600 steps. This is also evident from the Bland-Altman plot (Figure 5).

Level of Agreement

Mean differences with the ActiGraph accelerometer and the limits of agreement for each wearable device for measuring steps and MVPA are presented in Figure 5. For measuring steps, Huawei (503 steps) had the narrowest limits and Polar (770 steps) had the broadest limits. For MVPA, Asus (13.14 min) had the narrowest limits, and Fitbit (17.26 min) had the broadest limits.

Figure 3. Bland-Altman plots of the consumer-level devices. The middle line shows the mean difference (Positive values indicate an underestimation of the consumer-level device and negative values indicate an overestimation) between the measurements of moderate-to-vigorous physical activity of the device and the ActiGraph, and the dashed lines indicate the limits of agreement (1.96 × SD of the difference scores). MVPA: moderate to vigorous physical activity.

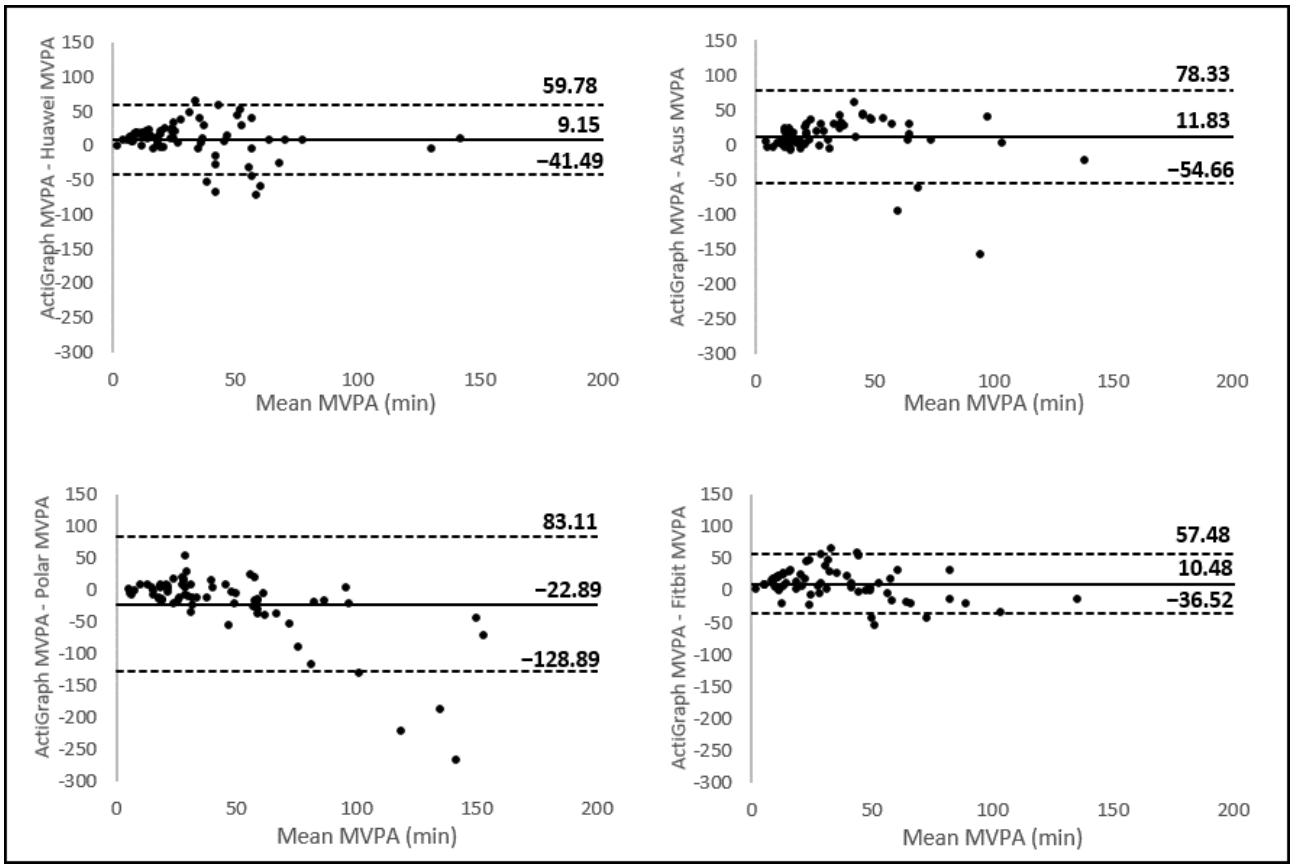


Table 4. Mean steps and minutes of moderate to vigorous physical activity per 15 min measured by Huawei, Asus, Polar, and Fitbit and the corresponding ActiGraph measurements and statistical significance (*P* value) of the difference between the ActiGraph accelerometer and the wearables.

Variable	Wearable, mean (SD)/15 min	ActiGraph accelerometer, mean (SD)/15 min	*P* value
Huawei			
Steps	184 (263)	148 (236)	<.001
MVPA[a] with zeros deleted (min)	2.91 (4.68)	2.53 (3.19)	.11
MVPA (min)	0.86 (3.06)	0.54 (1.79)	<.001
Asus			
Steps	166 (228)	147 (241)	.04
MVPA with zeros deleted (min)	2.44 (3.62)	2.50 (3.67)	.76
MVPA (min)	0.60 (2.07)	0.61 (2.12)	.76
Polar			
Steps	231 (358)	145 (241)	<.001
MVPA with zeros deleted (min)	3.75 (4.40)	1.62 (2.86)	<.001
MVPA (min)	1.20 (3.03)	0.52 (1.78)	<.001
Fitbit			
Steps	192 (304)	151 (247)	<.001
MVPA with zeros deleted (min)	3.59 (5.28)	2.86 (3.82)	.003
MVPA (min)	0.78 (2.86)	0.62 (2.13)	.01

[a]MVPA: moderate to vigorous physical activity.

Table 5. Correlation coefficients, intraclass correlation coefficients, and 95% CIs of measurements at a 15-min level.

Variable	Spearman *r* 95% CI	ICC[a] 95% CI
Huawei		
Steps	.752[b] (0.728-0.772)	.868[b] (0.859-0.877)
MVPA[c]	.177[b] (0.078-0.269)	.488[b] (0.424-0.547)
Asus		
Steps	.870[b] (0.851-0.880)	.837[b] (0.825-0.847)
MVPA	.208[b] (0.118-0.304)	.577[b] (0.524-0.625)
Polar		
Steps	.885[b] (0.875-0.898)	.792[b] (0.778-0.806)
MVPA	.153[b] (0.080-0.223)	.461[b] (0.408-0.512)
Fitbit		
Steps	.917[b] (0.906-0.928)	.887[b] (0.879-0.895)
MVPA	.116[b] (0.007-0.223)	.543[b] (0.481-0.599)

[a]ICC: interclass correlation coefficient

[b]*P*<.001.

[c]MVPA: moderate to vigorous physical activity.

Figure 4. Correlations between the activity estimates per 15 min from the wearables and the ActiGraph GT3X+, Spearman *r* values, and intraclass correlation coefficient values that denote the correlation for measuring moderate to vigorous physical activity or steps between the wearables and the ActiGraph. a) P<.001. ICC: intraclass correlation coefficient; MVPA: moderate to vigorous physical activity.

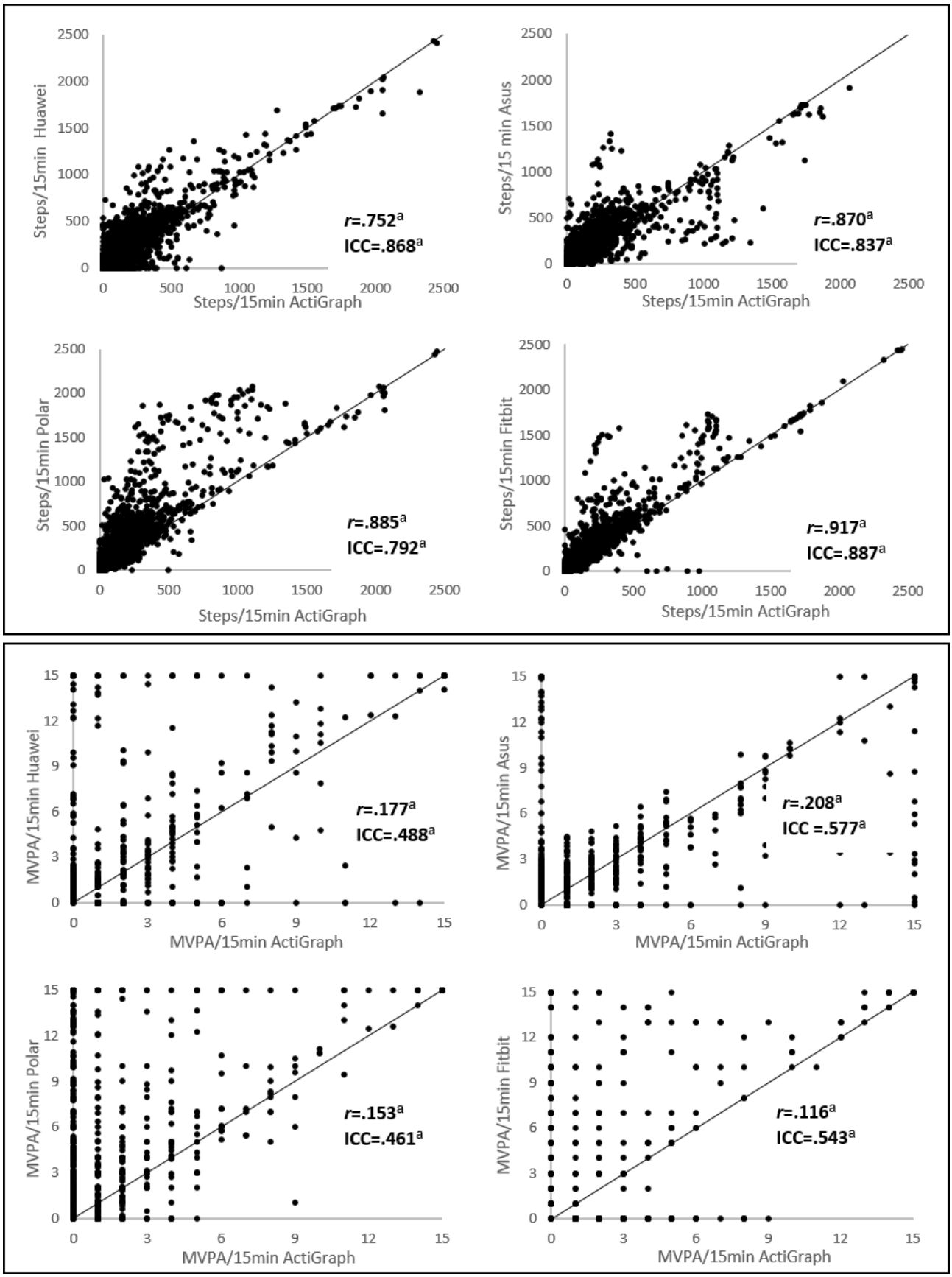
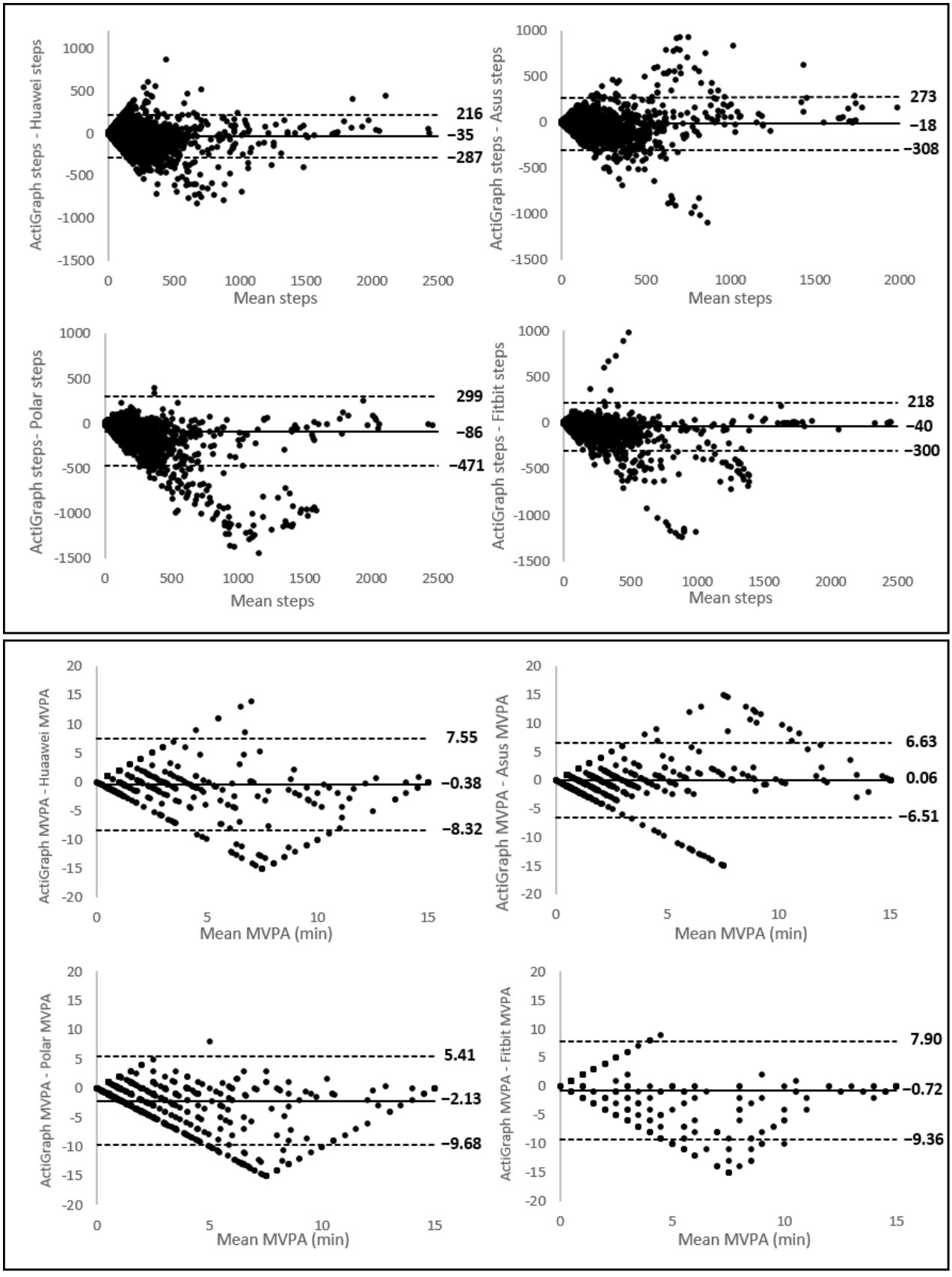


Figure 5. Bland-Altman plots of the wearables. The middle line shows the mean difference (positive values indicate an underestimation of the wearable and negative values indicate an overestimation) between the wearables and the ActiGraph, and the dashed lines indicate the limits of agreement (1.96 × SD of the difference scores). MVPA: moderate to vigorous physical activity.



Discussion

Principal Findings

This study investigated the validity of 4 wearables (3 smartwatches and 1 activity tracker) for measuring steps and MVPA in naturalistic situations. Validity was investigated separately for a day level and a 15-min level. The ActiGraph GT3X+ accelerometer was used as a convergent measure. The results can be readily summarized.

First, all 4 wearables showed good validity for measuring steps on a day level and a 15-min level. Nevertheless, all devices overestimated the number of steps. Second, for estimating MVPA, our study results demonstrated systematic bias for all wearables, both on a day level and a 15-min level, suggesting the validity is moderate to low for MVPA.

Although we cannot compare the overestimations of the steps per day for the smartwatches with previous studies, an overestimation for Fitbit has been reported before [18,35,39]. These studies showed that Fitbit overestimated steps on average by about 4% to 13% per day (step difference between wearable and Actigraph/steps measurement of the ActiGraph), which is a smaller overestimation than what we found. The overestimation for Fitbit (on average 1709/9126 steps, 18.72%), Huawei (on average 1477/8626 steps, 17.12%), and Polar (on average 3630/10,854 steps, 33.44%) was substantially larger. The overestimation on a day level was the smallest for Asus (on average 652 on 7662 steps; 8.50%). Moreover, on a 15-min level, all 4 devices overestimated the amount of steps: Huawei with on average 19.0% (35/184 steps), Asus with on average 10.8% (18/166 steps), Polar with 37.2% (86/231 steps), and Fitbit with 21.2% (41/193 steps). When looking at the limits of agreement on both levels, Polar shows the broadest limits, whereas Huawei shows the smallest limits. From this, it can be concluded that Polar is the least accurate device for measuring steps and that, despite the smallest mean difference being that of Asus, Huawei is the most accurate device for measuring steps. There are several reasons that may account for the systematic overestimation. First, the overestimation may also be explained by the different wear location of the devices. The ActiGraph GT3X+ is worn on the hip, whereas the wearables are worn on the wrist. This by itself could result in different measurements. Previous research concluded that wrist attachment devices detected consistently fewer counted steps than the waist attachment devices at most treadmill speeds during laboratory testing. In contrast, wrist attachment devices detected a higher average step count than the waist attachment devices under free-living conditions [40]. Second, the overestimation may also be explained by the algorithms used to convert raw activity data from the different sensors in the watches into steps. Companies may use a lower threshold for steps than the threshold for the ActiGraph accelerometer algorithm. In line with this hypothesis, the systematic error increased as the number of steps increased.

All devices displayed information on how much time per day was spent in PA of at least moderate intensity. In contrast to measuring steps, wearables showed only moderate validity for measuring MVPA relative to the ActiGraph GT3X+ accelerometer on a day level and even low validity on a 15-min level. Whether MVPA was overestimated or underestimated varied depending on the device type and the time level. On a day level, Fitbit, Huawei, and Asus underestimated MVPA with an average of 30% (10/35 min per day), 16% (9/57 min per day), and 36% (12/33 min per day), respectively, whereas Polar overestimated MVPA with 33% (23/70 min per day). When looking at the limits of agreement on a day level, Fitbit shows the narrowest limits, whereas Polar shows the broadest limits. Moreover, Huawei shows rather narrow limits, making it, in combination with the small mean difference, the most accurate for measuring MVPA on a day level. Polar, however, is the least accurate. On a 15-min level, Fitbit, Huawei, and Polar overestimated MVPA with 20% (0.72/3.60 min), 13% (0.38/2.91 min), and 57% (2.13/3.75 min), respectively, whereas Asus underestimated MVPA with 2% (0.06/2.44 min). Asus also showed the narrowest limits of agreement, meaning it is the most accurate wearable device for measuring MVPA on a 15-min level. The results of Fitbit Charge on a day level are in line with the findings of a validation study of Fitbit Flex in naturalistic settings in which an underestimation of 36% time spent on MVPA per day was found [21]. Other studies in naturalistic settings found an overestimation of the MVPA measurements by Fitbit on a day level with 77% to 153% per day [19,41]; however, in these studies, Fitbit was worn on the hip. The difference between the findings of these previous studies and this study can, therefore, be explained by the placement of the wearable. Ferguson et al and Reid et al investigated the validity of Fitbit One, Fitbit Zip, and Fitbit Flex. All these wearables are worn on the hip.

A possible explanation for the moderate to low validity found in our study could be that the PA variables measured by the devices were not explicitly identified as MVPA. However, because all devices had set a goal of 30 min PA per day (similar to the MVPA recommendations for adults), we assumed that the measured variable corresponded to MVPA as measured by the ActiGraph accelerometer. Nevertheless, specific information regarding intensity cut-points was not provided and publicly available from these 4 wearables. An earlier study showed that using different intensity cut-points in accelerometers resulted in different MVPA levels [42], suggesting that it is difficult to compare accelerometer MVPA measurements when intensity cut-points vary. This could be the case in this study, which makes it difficult to compare the Actigraph accelerometer MVPA measurements with the wearable MVPA measurements [43]. However, our results showed large inconsistent underestimations and overestimations between and within participants, which cannot only be attributed to the lack of definitional similarity of the measured variable. Therefore, the discrepancies here may be a result of both definitional and measurement problems (eg, sensitivity algorithm). These findings are in line with previous studies that have expressed concerns that such devices might not be able to provide adequate information to guide exercise intensity or detect MVPA [17].

The inclusion of 4 popular devices enables to draw conclusions on the validity of these 4 smartwatches and not only on a singular device. Moreover, to the best of our knowledge, this was the first study to explore validity of smartwatches to

measure steps and MVPA. The key strength of this study is the validation of the wearables on a 15-min level to investigate the potential of the devices to correctly situate physically active behavior over time to provide exact real-time feedback on PA behavior. Despite the clear results of this study, it is important to see them in the context of the purpose of the devices. The main purpose of these devices is to motivate the user to move more in everyday life, suggesting that 100% accurate measurements might not be needed. Modest accuracy can be good enough for this purpose [44]. Furthermore, this study has some other limitations. First, the choice of a 15-min level is arbitrary. It was the smallest data collection window in the Android Wear smartwatches. Ideally, validation on a smaller time-level, such as 1 or 5 min, should be performed to be able to better estimate the potential for providing real-time feedback. However, we can, based on the 15-min timescale, assume that these wearables will logically also not be accurately measuring MVPA on a smaller time-scale (eg, 10 min, 5 min, 1 min, and 30 s). Second, we used the ActiGraph accelerometer as convergent measure and not as a criterion measure, meaning it may not be considered the true golden standard. Although earlier studies showed good validity of the ActiGraph GT3X+ for measuring MVPA compared with indirect calorimetry, the main limitation for both uniaxial and triaxial accelerometers is the inability to accurately assess the movement associated with nonambulatory activity, such as cycling, especially with hip-worn devices [45]. For measuring steps, the golden standard is direct observation. For measuring MVPA, which is a complex and multifaceted construct, there is currently no consensus [46,47]. As by definition, PA leads to energy expenditure; the doubly labeled water (DLW) method, which assesses total energy expenditure over longer periods of time, is the golden standard to assess physical activities in laboratory settings [47,48]. However, because of feasibility, direct observation and DLW are impossible in free-living conditions. The ActiGraph was, therefore, by approximation, the best available golden standard. Third, the sample size was small but comparable with previous validation studies [19-21,38,41,49]. Fourth, the development of new wearables that appear on the market is going fast. Therefore, the need for further validation in naturalistic settings remains. Obviously, it is not possible to validate each single new device coming onto the market. However, we must always remain critical of measurements of PA by new devices, and research must continue to invest resources and time in this type of research, especially when new devices also have potential to be used within research. In this respect, it may be very useful in the future when manufacturers provide more insight into the cut-points and algorithms that were used to translate the raw data into useful information (such as steps and minutes of MVPA).

Conclusions

Generally, it can be concluded that all 4 consumer-level devices (Huawei Watch, Polar M600, Asus ZenWatch2, and Fitbit Charge) are valid devices to estimate the amount of steps in naturalistic situations on both a day level and 15-min level. Nevertheless, for estimating MVPA, our study reveals systematic bias for all devices, both on a day level and a 15-min level, suggesting the validity is moderate to low for MVPA. This suggests that these wearables cannot replace the current generation of research-based accelerometers such as the ActiGraph GT3X+ to assess MVPA. The MVPA results on a 15-min level also indicate that these devices are not accurate in giving feedback on how many minutes the user performed MVPA in the past 15 min. Although we were not able to investigate validity on a smaller time-scale, we can, based on the 15-min time-scale, assume that these wearables will not be accurate in measuring MVPA on a smaller time-scale as well (eg, 10 min, 5 min, 1 min, 30 s). Consequently, these wearables cannot be considered to have the potential to provide exact real-time feedback on minutes MVPA. Therefore, we conclude that these wearables cannot be used to inform the design of a JITAI or to serve as a platform for a JITAI to increase PA levels.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2016. Fiscal policies for diet and the prevention of noncommunicable diseases URL: http://www.who.int/dietphysicalactivity/publications/fiscal-policies-diet-prevention/en/ [accessed 2018-01-23] [WebCite Cache ID 6z8nx2Uwm]
2. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
3. Warburton DE, Charlesworth S, Ivey A, Nettlefold L, Bredin SS. A systematic review of the evidence for Canada's Physical Activity Guidelines for Adults. Int J Behav Nutr Phys Act 2010 May 11;7:39-220 [FREE Full text] [doi: 10.1186/1479-5868-7-39] [Medline: 20459783]
4. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985 Apr;100(2):126-131 [FREE Full text] [Medline: 3920711]
5. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1423-1434. [doi: 10.1249/mss.0b013e3180616b27] [Medline: 17762377]

6. World Health Organization. 2018. More Active People for a Healthier World URL: http://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf [accessed 2018-01-23] [WebCite Cache ID 71mlm7G6M]
7. Choi BC, Pak AW, Choi JC, Choi EC. Daily step goal of 10,000 steps: a literature review. Clin Invest Med 2007;30(3):E146-E151. [Medline: 17716553]
8. Tudor-Locke C, Bassett DR. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med 2004;34(1):1-8. [Medline: 14715035]
9. Dumith SC, Hallal PC, Reis RS, Kohl HW. Worldwide prevalence of physical inactivity and its association with human development index in 76 countries. Prev Med 2011 Jul;53(1-2):24-28. [doi: 10.1016/j.ypmed.2011.02.017] [Medline: 21371494]
10. Tuso P. Strategies to increase physical activity. Perm J 2015;19(4):84-88 [FREE Full text] [doi: 10.7812/TPP/14-242] [Medline: 26517440]
11. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322 [FREE Full text] [doi: 10.1093/her/cyr005] [Medline: 21321008]
12. Ashford S, Edmunds J, French DP. What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. Br J Health Psychol 2010 May;15(Pt 2):265-288. [doi: 10.1348/135910709X461752] [Medline: 19586583]
13. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR Mhealth Uhealth 2016;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
14. Rawassizadeh R, Price BA, Petre M. Wearables: has the age of smartwatches finally arrived? Commun. ACM 2014 Dec 23;58(1):45-47. [doi: 10.1145/2629633]
15. Reeder B, David A. Health at hand: a systematic review of smart watch uses for health and wellness. J Biomed Inform 2016 Dec;63:269-276 [FREE Full text] [doi: 10.1016/j.jbi.2016.09.001] [Medline: 27612974]
16. Piwek L, Ellis DA. Can programming frameworks bring smartphones into the mainstream of psychological science? Front Psychol 2016 Aug;7:1252 [FREE Full text] [doi: 10.3389/fpsyg.2016.01252] [Medline: 27602010]
17. Gorny AW, Liew SJ, Tan CS, Müller-Riemenschneider F. Fitbit Charge HR Wireless Heart Rate Monitor: validation study conducted under free-living conditions. JMIR Mhealth Uhealth 2017 Oct 20;5(10):e157 [FREE Full text] [doi: 10.2196/mhealth.8233] [Medline: 29055881]
18. Schneider M, Chau L. Validation of the Fitbit Zip for monitoring physical activity among free-living adolescents. BMC Res Notes 2016 Sep 21;9(1):448 [FREE Full text] [doi: 10.1186/s13104-016-2253-6] [Medline: 27655477]
19. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015 Mar;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
20. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot GM. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015 Oct;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
21. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of FitBit flex for step count, moderate to vigorous physical activity and activity energy expenditure. PLoS One 2016 Sep;11(9):e0161224 [FREE Full text] [doi: 10.1371/journal.pone.0161224] [Medline: 27589592]
22. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2014 Sep;17(5):496-500. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
23. Chu AH, Ng SH, Paknezhad M, Gauterin A, Koh D, Brown MS, et al. Comparison of wrist-worn Fitbit Flex and waist-worn ActiGraph for measuring steps in free-living adults. PLoS One 2017 Feb;12(2):e0172535 [FREE Full text] [doi: 10.1371/journal.pone.0172535] [Medline: 28234953]
24. Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. J Med Internet Res 2018 Mar 22;20(3):e110 [FREE Full text] [doi: 10.2196/jmir.9157] [Medline: 29567635]
25. Dunton GF, Dzubur E, Intille S. Feasibility and performance test of a real-time sensor-informed context-sensitive ecological momentary assessment to capture physical activity. J Med Internet Res 2016 Dec 1;18(6):e106 [FREE Full text] [doi: 10.2196/jmir.5398] [Medline: 27251313]
26. Sharmin M, Raij A, Epstien D, Nahum-Shani I, Beck JG, Vhaduri S, et al. Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions. Proc ACM Int Conf Ubiquitous Comput 2015 Sep;2015:505-516 [FREE Full text] [Medline: 26539566]
27. Kim Y, Park I, Kang M. Convergent validity of the international physical activity questionnaire (IPAQ): meta-analysis. Public Health Nutr 2013 Mar;16(3):440-452. [doi: 10.1017/S1368980012002996] [Medline: 22874087]
28. 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]

29. Rowlands AV, Stone MR, Eston RG. Influence of speed and step frequency during walking and running on motion sensor output. Med Sci Sports Exerc 2007 Apr;39(4):716-727. [doi: 10.1249/mss.0b013e318031126c] [Medline: 17414811]
30. Le Masurier GC, Lee SM, Tudor-Locke C. Motion sensor accuracy under controlled and free-living conditions. Med Sci Sports Exerc 2004 May;36(5):905-910. [Medline: 15126728]
31. Le Masurier GC, Tudor-Locke C. Comparison of pedometer and accelerometer accuracy under controlled conditions. Med Sci Sports Exerc 2003 May;35(5):867-871. [doi: 10.1249/01.MSS.0000064996.63632.10] [Medline: 12750599]
32. Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, Inc. accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [Medline: 9588623]
33. Santos-Lozano A, Santín-Medeiros F, Cardon G, Torres-Luque G, Bailón R, Bergmeir C, et al. Actigraph GT3X: validation and determination of physical activity intensity cut points. Int J Sports Med 2013 Nov;34(11):975-982. [doi: 10.1055/s-0033-1337945] [Medline: 23700330]
34. IDC Media Center. 2018 May 1. Global Wearables Market Grows 7.7% in 4Q17 and 10.3% in 2017 as Apple Seizes the Leader Position, Says IDC URL: https://www.idc.com/getdoc.jsp?containerId=prUS43598218 [accessed 2018-05-03] [WebCite Cache ID 6z8o9Zcue]
35. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fibit Zip™ physical activity monitor as a measure of free-living physical activity. BMC Res Notes 2014 Dec 23;7:952-955 [FREE Full text] [doi: 10.1186/1756-0500-7-952] [Medline: 25539733]
36. Bassett DR, Mahar MT, Rowe DA, Morrow JR. Walking and measurement. Med Sci Sports Exerc 2008 Jul;40(7 Suppl):S529-S536. [doi: 10.1249/MSS.0b013e31817c699c] [Medline: 18562970]
37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
38. Wahl Y, Düking P, Droszez A, Wahl P, Mester J. Criterion-validity of commercially available physical activity tracker to estimate step count, covered distance and energy expenditure during sports conditions. Front Physiol 2017 Sep;8:725 [FREE Full text] [doi: 10.3389/fphys.2017.00725] [Medline: 29018355]
39. Dominick GM, Winfree KN, Pohlig RT, Papas MA. Physical activity assessment between consumer- and research-grade accelerometers: a comparative study in free-living conditions. JMIR Mhealth Uhealth 2016 Sep 19;4(3):e110 [FREE Full text] [doi: 10.2196/mhealth.6281] [Medline: 27644334]
40. Tudor-Locke C, Barreira TV, Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. Med Sci Sports Exerc 2015 Apr;47(4):839-842. [doi: 10.1249/MSS.0000000000000476] [Medline: 25121517]
41. Reid RE, Insogna JA, Carver TE, Comptour AM, Bewski NA, Sciortino C, et al. Validity and reliability of Fitbit activity monitors compared to ActiGraph GT3X+ with female adults in a free-living environment. J Sci Med Sport 2017 Jun;20(6):578-582. [doi: 10.1016/j.jsams.2016.10.015] [Medline: 27887786]
42. Loprinzi PD, Lee H, Cardinal BJ, Crespo CJ, Andersen RE, Smit E. The relationship of actigraph accelerometer cut-points for estimating physical activity with selected health outcomes: results from NHANES 2003-06. Res Q Exerc Sport 2012 Sep;83(3):422-430. [doi: 10.1080/02701367.2012.10599877] [Medline: 22978192]
43. Bornstein DB, Beets MW, Byun W, Welk G, Bottai M, Dowda M, et al. Equating accelerometer estimates of moderate-to-vigorous physical activity: in search of the Rosetta Stone. J Sci Med Sport 2011 Sep;14(5):404-410 [FREE Full text] [doi: 10.1016/j.jsams.2011.03.013] [Medline: 21524938]
44. Gomersall SR, Ng N, Burton NW, Pavey TG, Gilson ND, Brown WJ. Estimating physical activity and sedentary behavior in a free-living context: a pragmatic comparison of consumer-based activity trackers and ActiGraph accelerometry. J Med Internet Res 2016 Sep 7;18(9):e239 [FREE Full text] [doi: 10.2196/jmir.5531] [Medline: 27604226]
45. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S523-S530. [Medline: 16294115]
46. Dencker M, Andersen LB. Accelerometer-measured daily physical activity related to aerobic fitness in children and adolescents. J Sports Sci 2011 Jun;29(9):887-895. [doi: 10.1080/02640414.2011.578148] [Medline: 21604226]
47. Hills AP, Mokhtar N, Byrne NM. Assessment of physical activity and energy expenditure: an overview of objective measures. Front Nutr 2014 Jun;1:5 [FREE Full text] [doi: 10.3389/fnut.2014.00005] [Medline: 25988109]
48. Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. Obes Rev 2013 Jun;14(6):451-462. [doi: 10.1111/obr.12021] [Medline: 23398786]
49. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015;12(1):159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]

Abbreviations

DLW: doubly labeled water
ICC: interclass correlation coefficient
IPAQ: International Physical Activity Questionnaire
JITAI: Just-In-Time adaptive intervention
METs: Metabolic Equivalents

MPA: moderate-intensity physical activity
MVPA: moderate to vigorous physical activity
PA: physical activity
VPA: vigorous-intensity physical activity

Edited by A Tsanas; submitted 04.05.18; peer-reviewed by A Maier, J Meng, JJ Murphy, A Rodríguez-Molinero; comments to author 27.06.18; revised version received 19.08.18; accepted 21.09.18; published 13.12.18.

Please cite as:
Degroote L, De Bourdeaudhuij I, Verloigne M, Poppe L, Crombez G
The Accuracy of Smart Devices for Measuring Physical Activity in Daily Life: Validation Study
JMIR Mhealth Uhealth 2018;6(12):e10972
URL: https://mhealth.jmir.org/2018/12/e10972/
doi:10.2196/10972
PMID:30545810

Corrigenda and Addenda

Correction: A Cardiopulmonary Monitoring System for Patient Transport Within Hospitals Using Mobile Internet of Things Technology: Observational Validation Study

Jang Ho Lee[1], MD; Yu Rang Park[2], PhD; Solbi Kweon[2], MS; Seulgi Kim[1], MS; Wonjun Ji[1], MD; Chang-Min Choi[1,3], MD, PhD

[1]Department of Pulmonology and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[2]Department of Biomedical System Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea
[3]Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

Corresponding Author:
Chang-Min Choi, MD, PhD
Department of Pulmonology and Critical Care Medicine
Asan Medical Center
University of Ulsan College of Medicine
88 Olympic-ro, 43-gil
Songpa-gu
Seoul, 05505
Republic of Korea
Phone: 82 2 3010 5902
Email: ccm@amc.seoul.kr

Related Article:

Correction of: http://mhealth.jmir.org/2018/11/e12048/

(JMIR Mhealth Uhealth 2018;6(12):e12864) doi:10.2196/12864

The authors of "A Cardiopulmonary Monitoring System for Patient Transport Within Hospitals Using Mobile Internet of Things Technology: Observational Validation Study" (JMIR Mhealth Uhealth 2018;6(11):e12048) wish to add their grant number to the Acknowledgments section.

Thus, the sentence "This study was supported by a grant from Korea Health Industry Development Institute" has been changed to "This study was supported by a grant from Korea Health Industry Development Institute (grant number HI17C0178)."

The correction will appear in the online version of the paper on the JMIR website on December 20, 2018, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Edited by G Eysenbach; submitted 19.11.18; this is a non–peer-reviewed article;accepted 28.11.18; published 21.12.18.

Please cite as:
Lee JH, Park YR, Kweon S, Kim S, Ji W, Choi CM
Correction: A Cardiopulmonary Monitoring System for Patient Transport Within Hospitals Using Mobile Internet of Things Technology: Observational Validation Study
JMIR Mhealth Uhealth 2018;6(12):e12864
URL: http://mhealth.jmir.org/2018/12/e12864/
doi:10.2196/12864
PMID:30578236

Original Paper

Combining Continuous Smartphone Native Sensors Data Capture and Unsupervised Data Mining Techniques for Behavioral Changes Detection: A Case Series of the Evidence-Based Behavior (eB2) Study

Sofian Berrouiguet[1,2,3,4], MD; David Ramírez[5,6], PhD; María Luisa Barrigón[7,8], MD, PhD; Pablo Moreno-Muñoz[5,6], MSc; Rodrigo Carmona Camacho[7], MD; Enrique Baca-García[7,8,9], MD, PhD; Antonio Artés-Rodríguez[5,6,9], PhD

[1]Department of Psychiatry and Emergency, Brest Medical University Hospital, Brest, France
[2]IMT Atlantique, Lab-STICC, F-29238, Brest, France
[3]SPURBO EA 7479, Université de Bretagne Occidentale (UBO), Brest, France
[4]CHRU Cavale Blanche University Hospital of Brest, Brest, France
[5]Universidad Carlos III de Madrid, Leganés, Spain
[6]Gregorio Marañón Health Research Institute, Madrid, Spain
[7]Department of Psychiatry, Fundación Jiménez Díaz Hospital, Madrid, Spain
[8]Department of Psychiatry, Autónoma University, Madrid, Spain
[9]Centro de Investigación en Salud Mental, Carlos III Institute of Health, Madrid, Spain

Corresponding Author:
Sofian Berrouiguet, MD
CHRU Cavale Blanche University Hospital of Brest
Boulevard Tanguy Prigent
Brest, 29609
France
Phone: 33 668294178
Email: sofian.berrouiguet@chu-brest.fr

Abstract

Background: The emergence of smartphones, wearable sensor technologies, and smart homes allows the nonintrusive collection of activity data. Thus, health-related events, such as activities of daily living (ADLs; eg, mobility patterns, feeding, sleeping, ...) can be captured without patients' active participation. We designed a system to detect changes in the mobility patterns based on the smartphone's native sensors and advanced machine learning and signal processing techniques.

Objective: The principal objective of this work is to assess the feasibility of detecting mobility pattern changes in a sample of outpatients with depression using the smartphone's sensors. The proposed method processed the data acquired by the smartphone using an unsupervised detection technique.

Methods: In this study, 38 outpatients from the Hospital Fundación Jiménez Díaz Psychiatry Department (Madrid, Spain) participated. The Evidence-Based Behavior (eB^2) app was downloaded by patients on the day of recruitment and configured with the assistance of a physician. The app captured the following data: inertial sensors, physical activity, phone calls and message logs, app usage, nearby Bluetooth and Wi-Fi connections, and location. We applied a change-point detection technique to location data on a sample of 9 outpatients recruited between April 6, 2017 and December 14, 2017. The change-point detection was based only on location information, but the eB^2 platform allowed for an easy integration of additional data. The app remained running in the background on patients' smartphone during the study participation.

Results: The principal outcome measure was the identification of mobility pattern changes based on an unsupervised detection technique applied to the smartphone's native sensors data. Here, results from 5 patients' records are presented as a case series. The eB^2 system detected specific mobility pattern changes according to the patients' activity, which may be used as indicators of behavioral and clinical state changes.

Conclusions: The proposed technique could automatically detect changes in the mobility patterns of outpatients who took part in this study. Assuming these mobility pattern changes correlated with behavioral changes, we have developed a technique that

may identify possible relapses or clinical changes. Nevertheless, it is important to point out that the detected changes are not always related to relapses and that some clinical changes cannot be detected by the proposed method.

(JMIR Mhealth Uhealth 2018;6(12):e197) doi:10.2196/mhealth.9472

KEYWORDS

behavioral changes; data mining; mental disorders; sensors; wearables

Introduction

Data Capture in Patient Environment

Web-based and smartphone apps offer new opportunities for patient monitoring. The integration of these tools into medical practice has heralded the electronic health (eHealth) era. eHealth involves the integration of new technologies into routine clinical practice by increasing networking possibilities between patients and clinicians. Recent trials using mobile electronic devices have proven successful in real-world and real-time monitoring and have improved the assessment possibilities in a large panel of clinical settings [1]. The assessment of patients' dynamic relationships between events and disease course is enhanced by the development of momentary data collection strategies such as experience sampling methods and ecological momentary assessment (EMA). These approaches, which rely on delivering informative contents and self-administered questionnaires, reduce the recall bias, as they are done in quasi-real time, but these face many limitations, including poor data reliability, burden and intrusiveness for patients, and data security issues [2].

In addition, electronic devices can perform passive (or autonomous) data gathering, that is, to extract information about users without any effort on their part. Actigraphy, geolocation, and communication activity are usual features of current smartphones and may be indicators of patients' behavior if they are properly processed. Advances in sensors technology and novel textile-electronic integration techniques also draw new perspectives for behavior ecological assessment. Moreover, it is currently possible to find commercially available wearable sensing technologies for several wellness and clinical purposes: simple heart rate monitors [3], rehabilitation after surgical intervention [4], and monitors of physical activity or sleep quality assessment [5]. Overall, an extensive panel of physical and mental conditions (eg, insomnia, diabetes, problems associated with older age, cardiac problems, or respiratory problems) can be remotely monitored by appropriate health care professionals—physicians, doctors, or nurses. These devices are often connected to a smartphone, which increases the networking capabilities and the user experience. Furthermore, the collected data can be processed and transferred over the internet to a remote clinical backend server for further analysis, assessment, and decision making and intervention if needed.

Monitoring Activities of Daily Living and Mobility Patterns

The emergence of smart homes and wearable sensor technologies allows nonintrusive collection of activity data [6]. Thus, health-related events, such as activities of daily living (ADLs; eg, feeding and sleeping) and patients' mobility patterns, can be captured without their active participation [7]. Monitoring behavioral changes of psychiatric patients and their ability to carry out their ADLs will likely improve the knowledge about the disease course. For example, the detection of changes in behavioral patterns may help in detecting emerging disorders [8]. In addition, smart home and ambient assisted living systems use sensors and other devices that are either wearable or integrated in the patients' home and have been used to assess the effect of undesirable symptoms and cognitive impairment on ADL functions [9] or to detect emerging disorders based on changes in patients' behavior [10]. The ease of access to smartphone technology for the general population and recent technological advances in smartphone-integrated sensors are paving the way for behavioral changes detection, based only on activity assessment. Physical activity assessment is usually based on findings from brief, regularly scheduled, in-person appointments or self-reported questionnaires [11]. Although widely used, this approach reduces the assessment in cross-sectional observations that miss essential information and are subject to recall bias. In this study, the data obtained from smartphones and integrated devices will be processed to identify mobility pattern changes, as they may be correlated with behavioral changes and clinical changes. For example, an increase in depressive symptoms is associated with a reduction of the patients' physical activity [12]. Thus, patient's mobility patterns may be used as proxies for behavioral changes. In a clinical setting, the detection of mobility pattern changes could be used by clinicians or caregivers as signals of (possible) behavioral changes in their patients.

Along the lines proposed in this work, recent studies have shown that smartphone data can be used to identify behavioral changes in patients. Abdullah et al. [13] reported that combining self-reported data with data from several smartphone sensors and communication patterns resulted in the reliable prediction of the Social Rhythm Metric, a clinically validated marker of stability and rhythmicity for individuals with bipolar disorder. Another system, Monsenso, collects and extracts voice features from phone calls that were made during everyday life in naturalistic settings [14]. Concretely, the MONARCA II Research Project, which uses Monsenso, obtained 6552 numerical features related to the pitch and voice variance that were extracted from patients' phone calls during their everyday life. Another platform is Beiwe, which is a research-oriented platform for digital phenotyping. Using Beiwe, Barnett et al. [15] developed a method to predict schizophrenia based on anomaly detection.

Considering the strengths and pitfalls of smartphone monitoring strategies, we have designed a system capable of performing continuous monitoring of patients using the smartphone and wearable sensors and data entry (data from phone calls,

messages, and so on). This Evidence-Based Behavior (eB^2) platform comprises a smartphone app, which collects these data, and a backend server, which stores and processes them. The eB^2 app collects data from inertial sensors, physical activity, phone calls and message logs, app usage, nearby Bluetooth and Wi-Fi connections, and location. In addition, using Google Play Services, the app can access detailed activity information and nearby location data. Moreover, wearable devices provide information like the body temperature, heart rate, or galvanic skin response. The app was developed to run in the background, and users only interact with the app for the initial configuration. Furthermore, it was designed with battery-safe considerations like noncontinuous recording schedule, automatic sleep and wake function, and it additionally notifies the operating system to relaunch itself when it is closed or stopped because of users' actions or failures and reboots.

Hypothesis and Principal Objective

We hypothesize that it is feasible to develop an analysis method capable of detecting mobility pattern changes based on the data acquired by the eB^2 system. Moreover, we believed that these changes might serve as proxies for behavioral changes. This study aims to assess the feasibility of detecting mobility pattern changes in a sample of outpatients using a smartphone app and an unsupervised detection method, which was run on a backend server.

Methods

Summary

We performed an unsupervised detection method and a qualitative analysis of a sample of 5 patients out of 38 outpatients enrolled in the eB^2 study between April 6 and December 14, 2017. The eB^2 study was (and still is) a 2-year, multicenter-controlled trial conducted by the Fundación Jiménez Díaz. Concretely, it was a prospective study that aimed to determine whether the behavioral changes detected by the eB^2 system correlated with any clinical change. Note, however, that in this preliminary work, we only focused on mobility pattern changes.

Participants

Patients who received psychiatric care in an outpatient mental health center of the Psychiatry Department at Fundación Jiménez Díaz, a University Hospital in Madrid, Spain, were approached to participate in this study. This department is part of the National Health Service and provides medical coverage financed by taxes to a catchment area of 420,000 people. The research followed the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Patient Inclusion and Exclusion Criteria

The inclusion criteria for this study were either male or female outpatients aged ≥18 years, diagnosed with mood disorders (ICD-10 codes F30-F39) or adjustment disorders (ICD-10 code F43.2), and coping with depression. Moreover, patients had to own a smartphone with an Android or iOS operating system, be connected to a Wi-Fi network, at least, once a week, and had to have given written informed consent for the eB^2 study. Participants were excluded if they were under the age of 18 years, illiterate, enrolled in other trials, or were in situations that did not allow obtaining written informed consent. Participants were not paid. Members of the study office (EBG, MLB, and RC) established an initial list of patients that met the inclusion criteria. The contents of the monitoring interviews were reviewed to identify patients who had attended, at least, 2 appointments. These criteria yielded the aforementioned 38 outpatients.

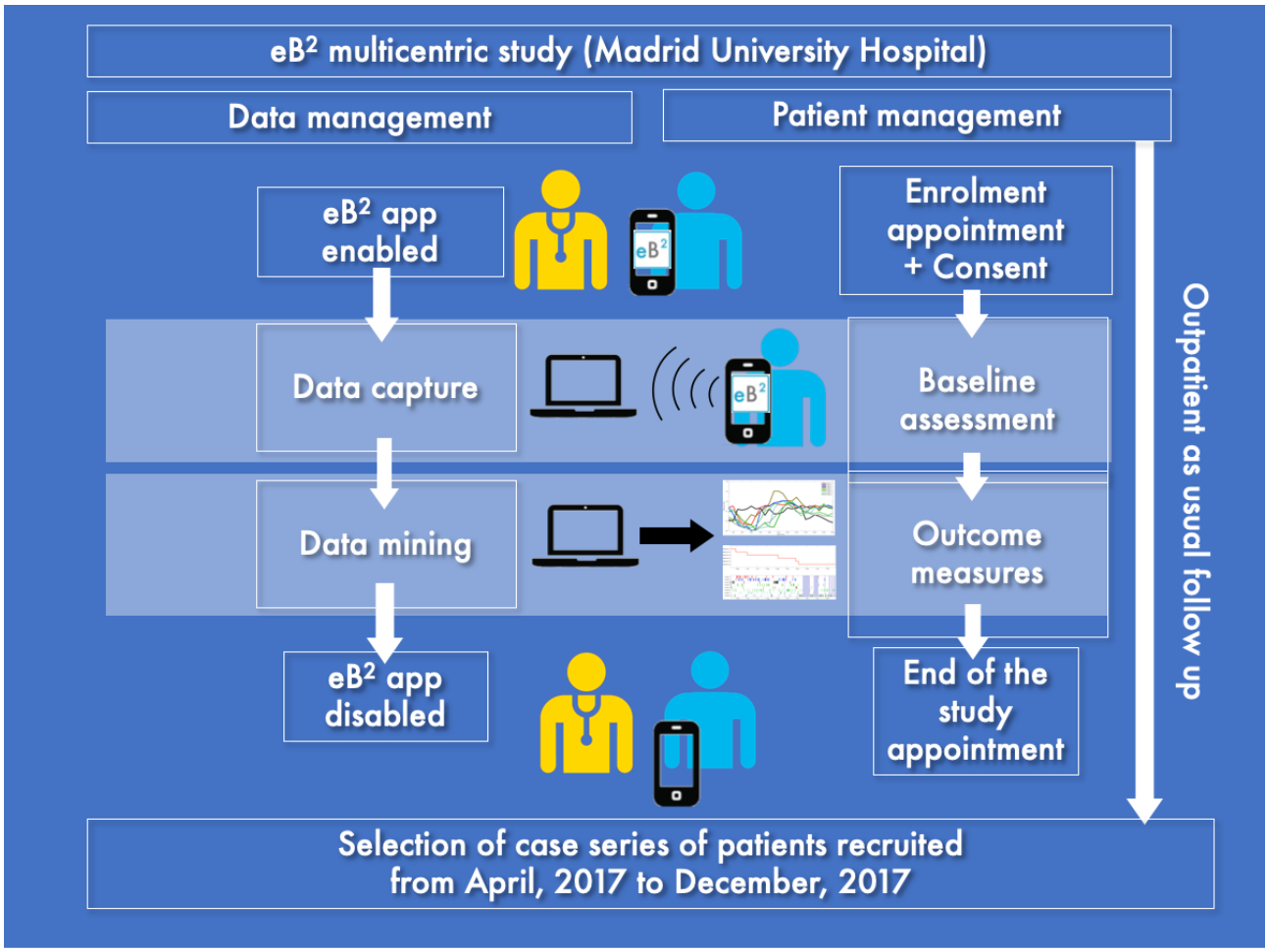
Research Protocol

The eB^2 app was downloaded on patients' smartphones on the day of recruitment and configured with the assistance of a physician (Figure 1). The app remained running in the background in patients' smartphones during the study. As previously pointed out, the app was designed with no patient interface, that is, no action of patients was required to capture data.

The eB^2 app collected the following data: actigraphy, global positioning system (GPS) location, Google location, app usage log, phone calls and message logs, nearby Wi-Fi and Bluetooth devices, and inertial measurement unit signals. The data gathered by the eB^2 app was anonymized if it was sensitive data (position and phone numbers), then it was translated to a unique data schema, and finally transmitted via Wi-Fi to the eB^2 backend server where it was stored. The transmission was done through a RESTful application programming interface (API), which had been developed using the JAVA Spring framework. This API is secure sockets layer protected and, to restrict access to the patients' information, a token-based access policy was implemented following the OAuth2 standard.

In addition to the data captured by the smartphone app, it was possible to collect data provided by third-party APIs, which were also translated into the common data schema by a service that ran at the server. The authorization to use these APIs was requested from the smartphone app and was also token-based. Moreover, data from wearable devices, like Fitbit or Microsoft Band 2, could also be uploaded. It is important to point out that, in this study, only the GPS location was used, which resulted in a simple technique and allowed for an easy clinical interpretation. Finally, signal processing and machine learning algorithms treated the acquired data to extract information, which was used afterwards by clinicians.

Figure 1. Visual representation of the study protocol. eB^2: Evidence-Based Behavior.



Baseline Characteristics

The baseline characteristics were recorded during an in-person interview for 38 patients enrolled in the eB^2 study. Variables collected for each patient profile were sex, age, Patient Health Questionnaire-9 (PHQ-9) score [16], diagnosis, and treatment. Clinical diagnoses were made by psychiatrists and coded according to the ICD-10 for mental disorders. Moreover, in each appointment, a psychiatrist administered the PHQ-9 questionnaire, which was designed to assess depression. These variables were entered manually into a secured electronic health record. Each patient was identified by a numeric code to ensure patient anonymity; this code was stored in the database and remained the same throughout all contact with patients. This study did not include a control group.

Outcome Measures

The principal outcome measure of this study was the identification of changes in patients' mobility patterns based on the smartphone's sensors data, which were processed by an unsupervised detection technique. We postulated that these changes could correlate with behavioral changes and relapses. That is, mobility patterns changes were proxies for (more general) behavioral changes. These data were interpreted for each selected patient in the light of the clinical data gathered in routine appointments during study participation.

Description of the Unsupervised Detection Technique

The proposed unsupervised detection technique comprised 2 algorithms. The first one was an unsupervised clustering technique that defined types of days. This classification was done according to the mobility profile, which was also learned in an unsupervised fashion. The mobility profiles could show, for instance, whether a patient was more active in the morning, afternoon, or evening, or even not active at all. The first step of the clustering technique was to summarize the measured distance acquired on an interval of a few minutes into larger 1-hour intervals and then the aggregated distances were stacked into 24-dimensional vectors; that is, each of these vectors corresponded to a given day, and each component was the cumulative distance traveled by the patient in the corresponding hour. Once we had these vectors, a clustering technique based on a mixture of Gaussians [17] was applied. The parameters of the model, that is, the means and covariance matrices (which we assumed diagonal with only 2 different values out of the 24 possible) were estimated using the Expectation-Maximization algorithm [17]. In particular, the estimated mean of each cluster defined what we called mobility profile, as it showed that in the corresponding cluster, the patient was more active (traveled more distance during the day, night, or at commuting hours). In addition, the Expectation-Maximization algorithm also allowed the handling of missing data, which corresponded to hours for which location data were not available. The final comment regarding the clustering step is the selection of the

number of clusters. That is, the allowed number of different profiles (or types of days). This selection obviously depended on the amount of available data, that is, more data allowed the technique to learn more profiles properly. However, an incorrect choice (too large or too small) would result in poor performance. Hence, we used an automatic method, which was based on the minimum description length (MDL) criterion [17].

Regardless of whether a patient was stable or not, these profiles were likely to change from day to day owing to weekends or public holidays. Hence, to detect mobility pattern changes, it did not suffice to detect profile changes (from one type of day or cluster to another). Concretely, we needed to detect changes in the distribution of these profiles. As an example, for a stable patient, the most likely profile was that of a workday, and a different profile could have appeared for the weekends. Nevertheless, the transition from one to another was not identified as a change. What we had to detect was, for instance, if these workday profiles started to appear less often because the patient stopped going to his or her work. Hence, we applied a change-point detection technique to identify when the probability (a portion of time) of each type of day suddenly changed. Moreover, this change-point detector could handle missing data. Then, the clustering technique handled missing hours, and the change-point detector handled missing days.

The technique described so far only exploited information given by the traveled distance, but both the technique and the eB^2 system may be generalized to incorporate other types of data. For instance, we may exploit how many phone calls were made every hour and, similar to the distance traveled profiles, we should detect changes in the distribution of these calls. Nevertheless, in this preliminary study, we wanted to study the feasibility of this detection based only on the traveled distance, as it resulted in a simple technique that was easier to interpret.

Results

Summary of the Results

Figure 2 presents the patient selection process of the case series. In the eB^2 study, 38 patients were recruited when we started the patient selection process for this case series. Nevertheless, of these 38 patients, only 18 had enabled the location (GPS), and of these 18, only 9 had the location enabled for >1 month, which was approximately the required time for the technique to work properly. That is, during the study, many patients disabled the location.

We addressed 9 patients for eligibility. Table 1 summarizes the results for those 9 patients, showing the number of monitored days, number of profiles or clusters, number of detected change-points, number of days between change-points, and the phone model and operating system version.

Figure 2. The patient selection process. eB^2: Evidence-Based Behavior; PHQ: Patient Health Questionnaire-9.

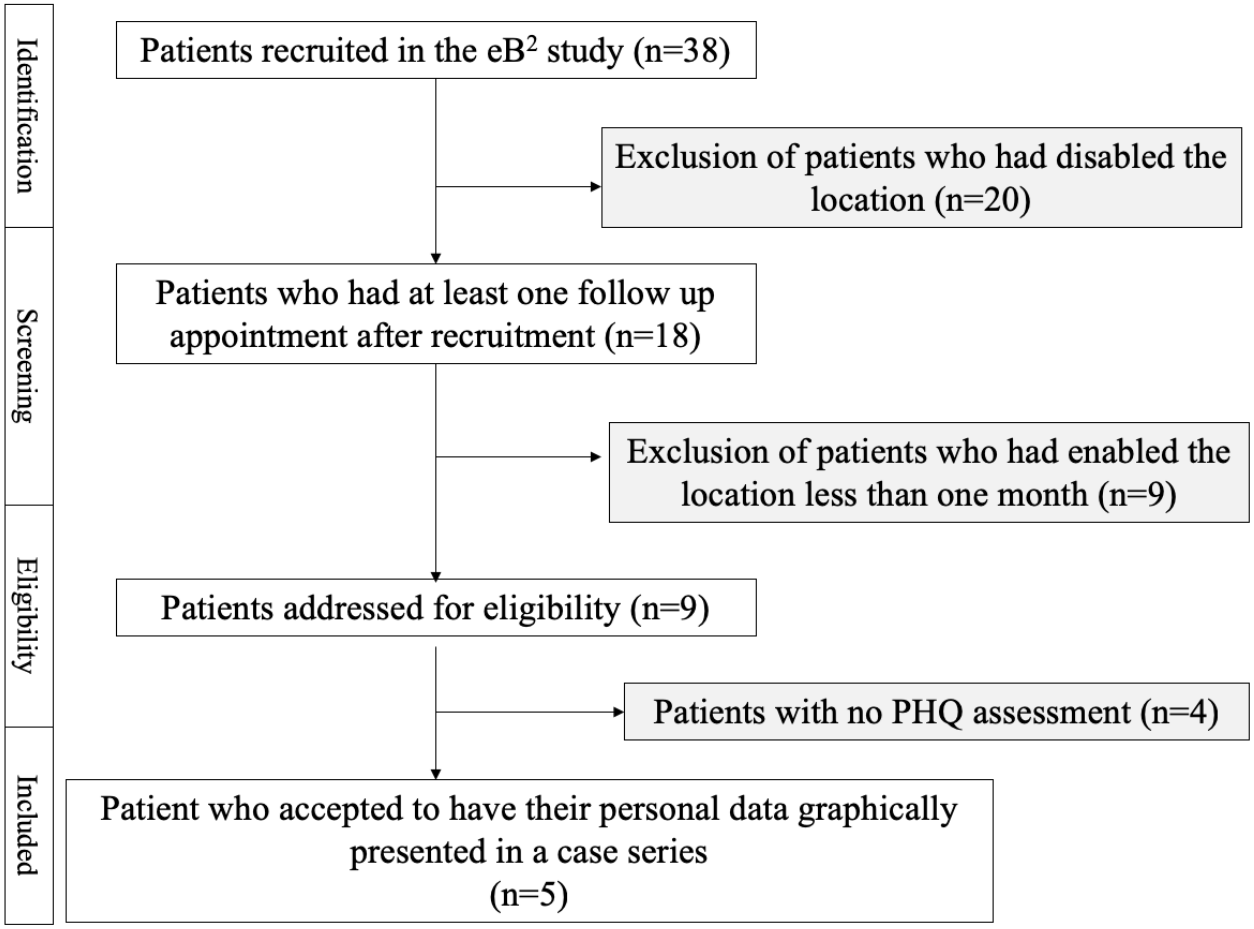


Table 1. Statistics of patients addressed for eligibility.

Patient	Monitored days (n)	Profiles (n)	Detected change-points (n)	Days between change-points (n)	Phone model-operating system
A	323	8	5	17, 35, 95, 54, 9, 113	Samsung Galaxy S7-6.0.1
B	240	5	0	N/A[a]	Samsung Galaxy A5-6.0.1
C	233	4	1	154, 79	Samsung Galaxy A5-7.0
D	162	4	0	N/A	Samsung Galaxy J7-6.0.1
E	75	3	1	49, 26	BQ Aquaris M5-6.0.1
F[b]	155	4	1	37, 118	Sony Xperia M5-6.0
G[b]	222	9	2	6, 30, 186	Huawei Y6-5.1
H[b]	154	5	3	33, 8, 47, 66	Samsung Galaxy J7-6.0.1
I[b]	41	3	0	41	Samsung Grand Prime-5.1.1

[a]N/A: not applicable.

[b]Patient not presented in the case series.

In the following, to shed some light on the technique and results, we present more detailed results for patients A-E as a case series, which were the patients to whom the PHQ-9 questionnaire was administered during routine appointments.

Detailed Analysis of Five Selected Patients

Patient A was a 56-year-old woman. She was diagnosed with recurrent depressive disorder and fibromyalgia. She was prescribed a daily oral medication of duloxetine 90 mg, quetiapine 150 mg, pregabalin 300 mg, and zolpidem 10 mg. She had regular bedtime and wake-up times during the study period. The clinical assessment of depression showed high scores of PHQ-9: 21 on April 6, 2017, and 25 on May 31, 2017. Unfortunately, this woman dropped out of medical follow-up, and there are no more clinical assessments.

She participated in the study from April 6, 2017 to February 28, 2018 and owned a Samsung Galaxy S7 that ran Android 6.0.1. Figure 3 shows that the MDL criterion selected 8 different clusters (ie, types of days or mobility patterns). We plotted the patient's inferred mobility patterns (in logarithmic scale), which are given by the mean of each cluster. For instance, profile 5 corresponded to a more active day and, on the days associated with this profile, the patient was more active between 9:00 and 16:00. Moreover, some of these profiles reported similar activity variations throughout the day. The sleep period was identified by a decrease in the activity between 1:00 and 6:00.

Figure 4 shows the output of the second step of the proposed method, the change-point detector; this figure displays the dates of the change-points (top) and the classification of each day given by the clustering technique and its temporal evolution (bottom). The algorithm identified a few dates as mobility pattern changes. Concretely, changes were noted on April 26, May 31, August 19, September 3, October 27, and November 5. These changes appeared when the probability (a portion of time) of each type of day varied.

Finally, we must point out that in Figure 4, where the temporal evolution of the types of days is shown, vertical light-blue rectangles indicate that the data corresponding to the marked days were completely missing. Even in these cases, the technique was robust enough to work properly.

Patient B was a 45-year-old woman. She was diagnosed with dysthymia and prescribed a daily oral medication of sertraline 100 mg. The clinical assessment of depression showed clinical improvement of depressive symptoms (June 7, 2017: PHQ-9=20; July 5, 2017: PHQ-9=8). Overall, medical records showed improvement during follow-up, explained by the participant as an improvement in cognitive performance, a decrease of death thoughts, and improvement of hedonic capacity.

She participated in the study from June 7, 2017 to January 30, 2018, and owned a Samsung Galaxy A5 running Android 6.0.1. In this case, the technique selected 5 different clusters. Figure 5 shows the patient's average mobility patterns. Figure 6 shows that our technique did not identify any change and that profile 4 was the most common, which was a low-mobility profile (there was not a single hour with >1 km). In this particular patient, clinical changes did not correlate with mobility as the main symptoms were expressed in cognitive and hedonic areas.

Patient C was a 40-year-old woman. She was diagnosed with a moderate depressive episode. She was prescribed a daily oral medication of paroxetine 20 mg, which was changed to vortioxetine 10 mg in August owing to the lack of improvement. Medical records showed an improvement after the change to vortioxetine.

This patient participated in the study from June 9, 2017 to February 28, 2018, and owned a Samsung Galaxy A5 that ran Android 7.0. In this case, the technique only considered 4 different types of days. Figure 7 shows the average distance traveled in each cluster, where we observed that the patient was more active after 7:00 in 3 out of the 4 profiles. Moreover, the remaining profile, profile 2, showed increased activity during the night, and profile 4 corresponded to a low-mobility profile. Figure 8 shows that the change-point detection algorithm detected only one change on December 9; after this date, the low-mobility profile began appearing more often, which possibly indicated a decrease of the patient's physical activity.

Figure 3. Distance traveled profiles of patient A.

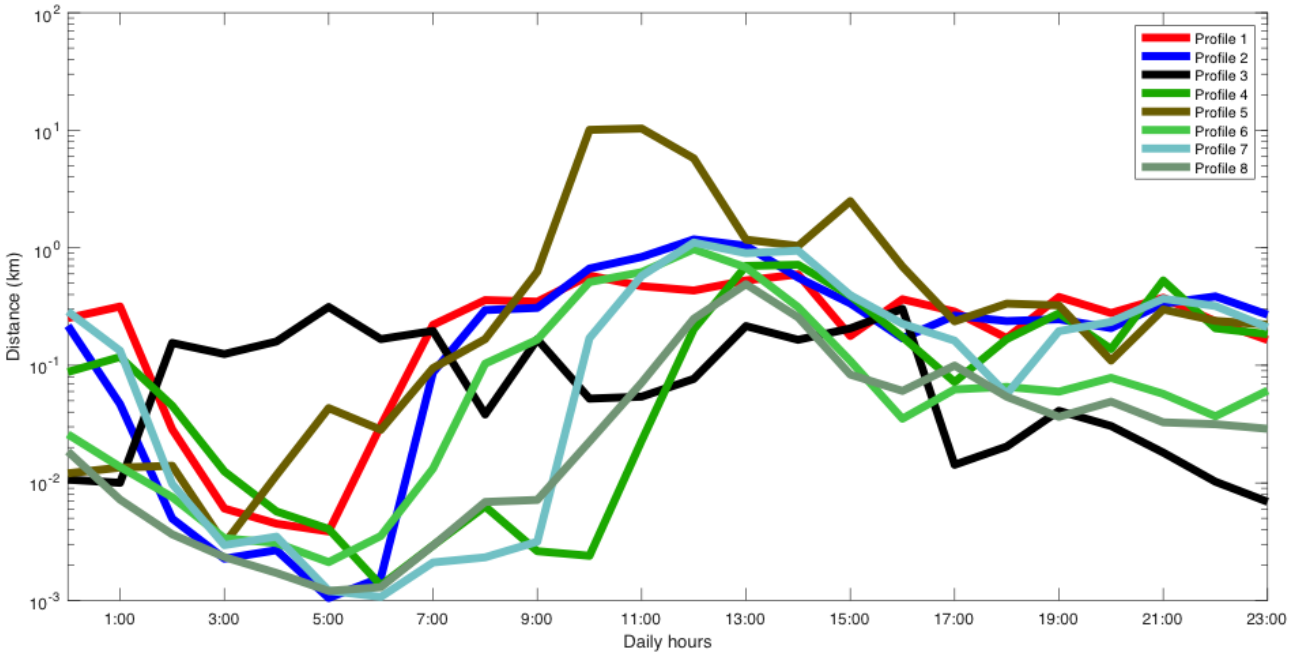


Figure 4. Representation of mobility pattern changes (upper) identified by the technique and corresponding patterns (lower) during study participation of patient A.

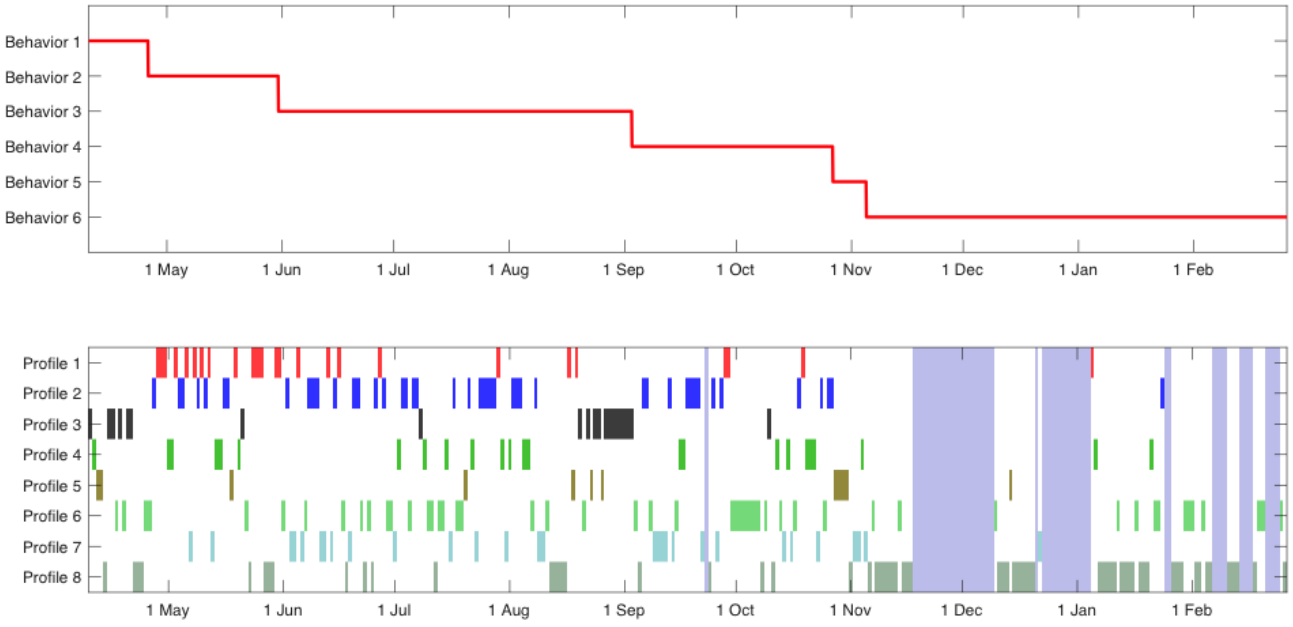


Figure 5. Distance traveled profiles of patient B.

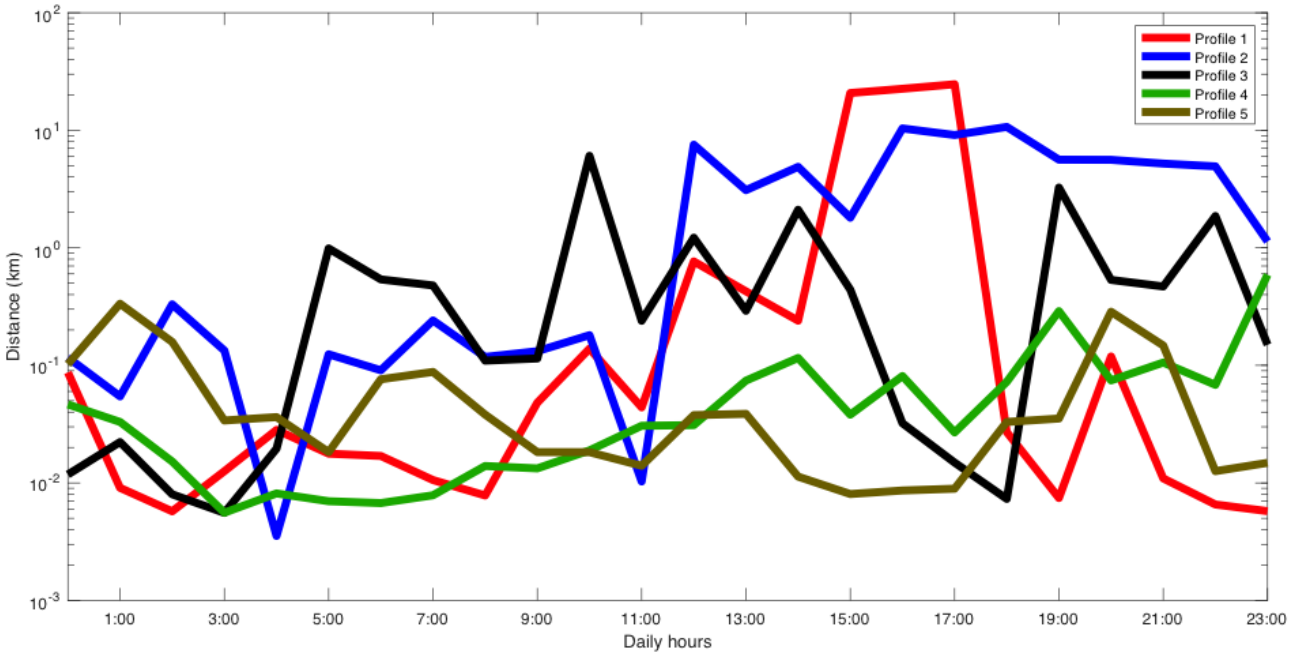


Figure 6. Representation of mobility pattern changes (above) identified by the technique and corresponding patterns (down) during study participation of patient B.

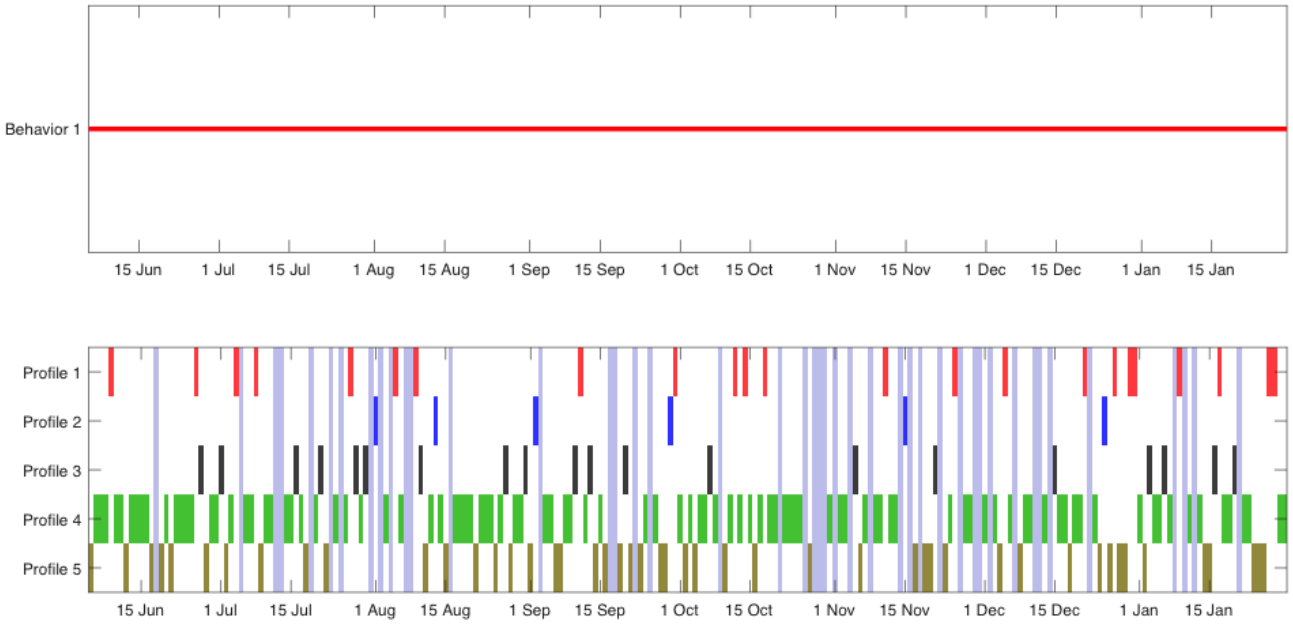


Figure 7. Distance traveled profiles of patient C.

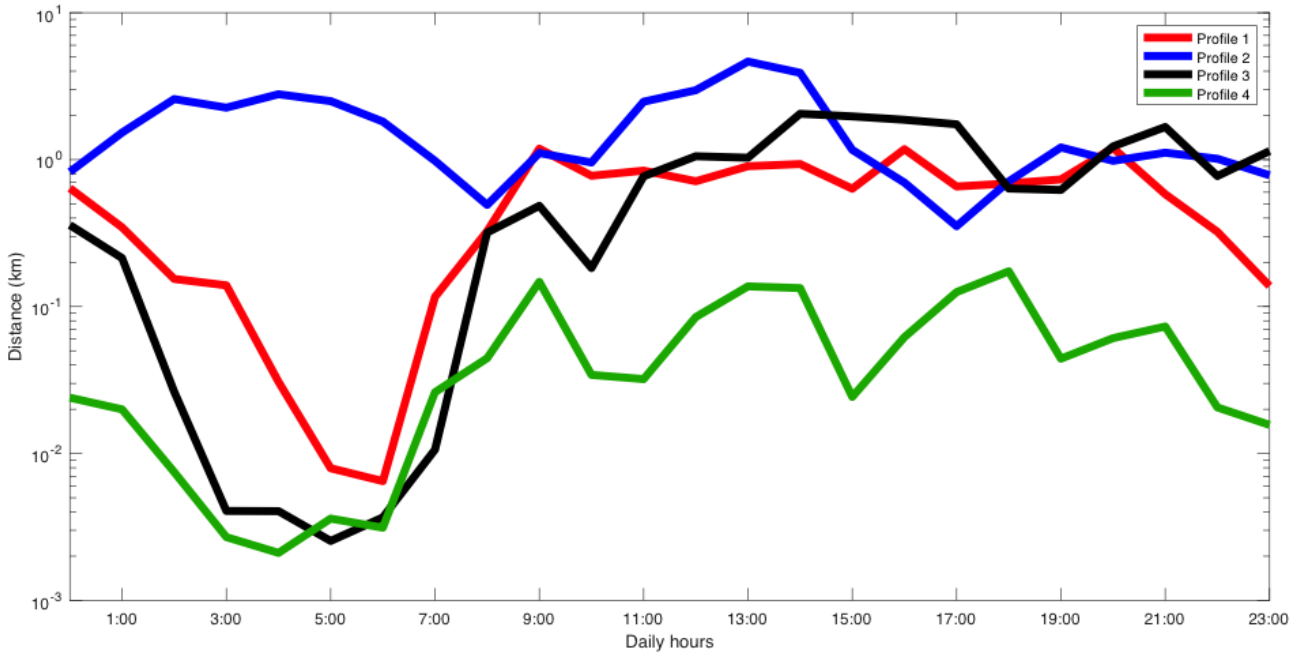
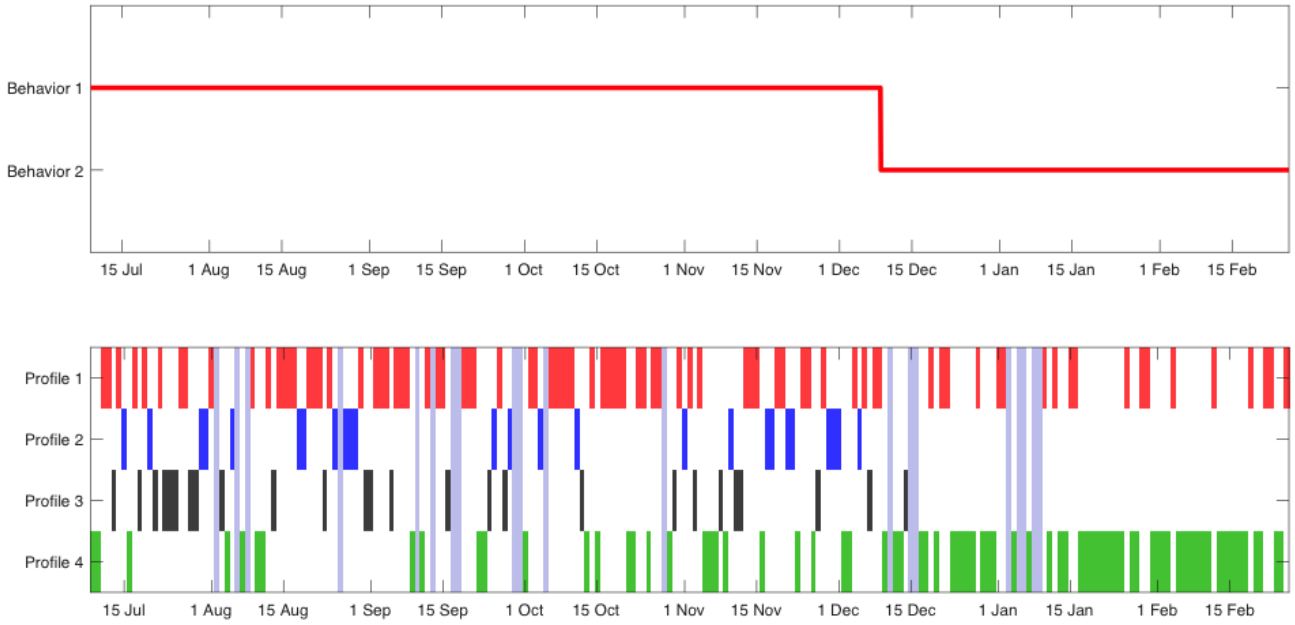


Figure 8. Representation of mobility pattern changes (above) identified by the technique and corresponding patterns (down) during study participation of patient C.



The clinical assessment of depression showed a decrease in depressive symptoms during the follow-up period (June 9, 2017: PHQ-9=22; Sept 9, 2017: PHQ-9=5; December 1, 2017: PHQ-9=4). Clinical improvement was associated with improved sleep time and sleep quality. A change of her work location led to less commuting, which can also explain the observed mobility patterns.

Patient D was a 36-year-old man. He was diagnosed with recurrent depressive disorder and prescribed a daily oral medication of venlafaxine retard 150 mg and lamotrigine 100 mg. He was included in the study after psychiatric hospitalization discharge, and clinical and functional remissions were observed in successive appointments in the outpatient setting. The clinical assessment of depression showed minor clinical improvement (March 17, 2017: PHQ-9=6; April 20, 2017: PHQ-9=2; May 24, 2017: PHQ-9=2; and June 26, 2017: PHQ-9=0).

He participated in the study from April 6, 2017 to August 11, 2017, and owned a Samsung Galaxy J7 running Android 6.0.1. Figure 9 shows that the number of profiles selected by the MDL criterion was 4. Profiles 1, 3, and 4 corresponded to typical urban mobility profiles. Some showed higher mobility during day or night, and some had peaks at commuting times (7:00 and 19:00). However, profile 2 corresponded very likely to a trip as the average movement per hour was around 100 km. Figure 10 shows the results of the change-point detector, which did not detect any change-point; this is coherent with the clinical evolution of the patient.

Figure 9. Distance traveled profiles of patient D.

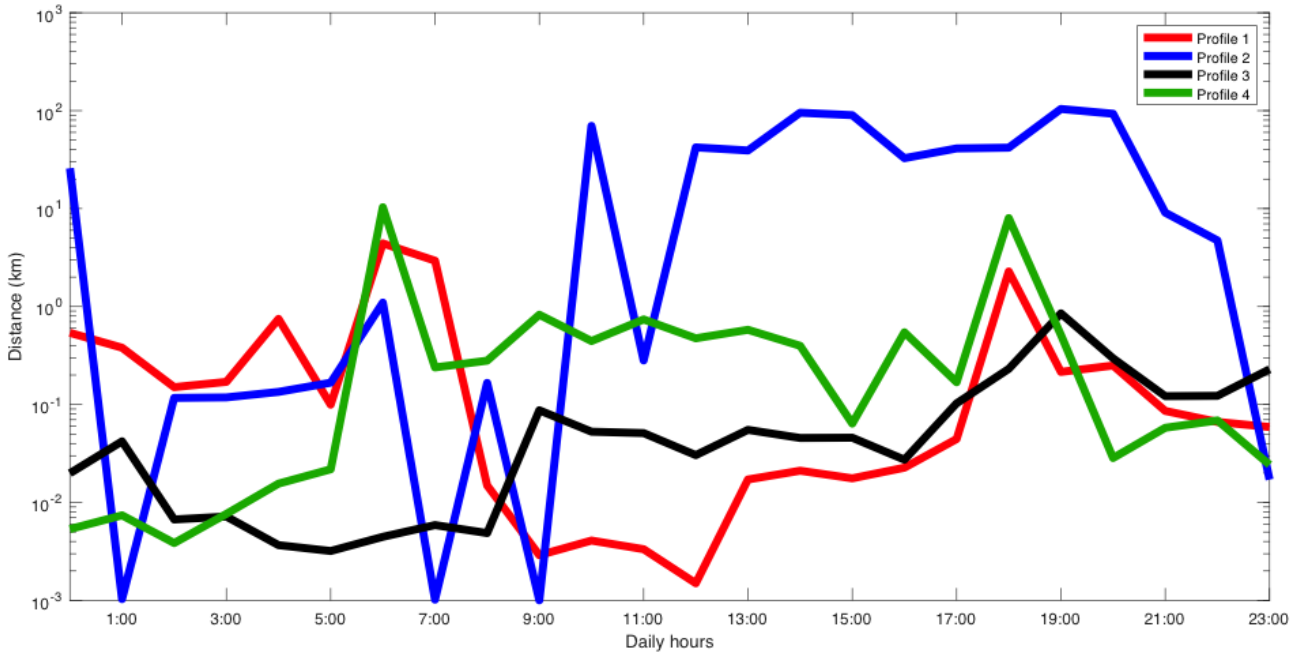
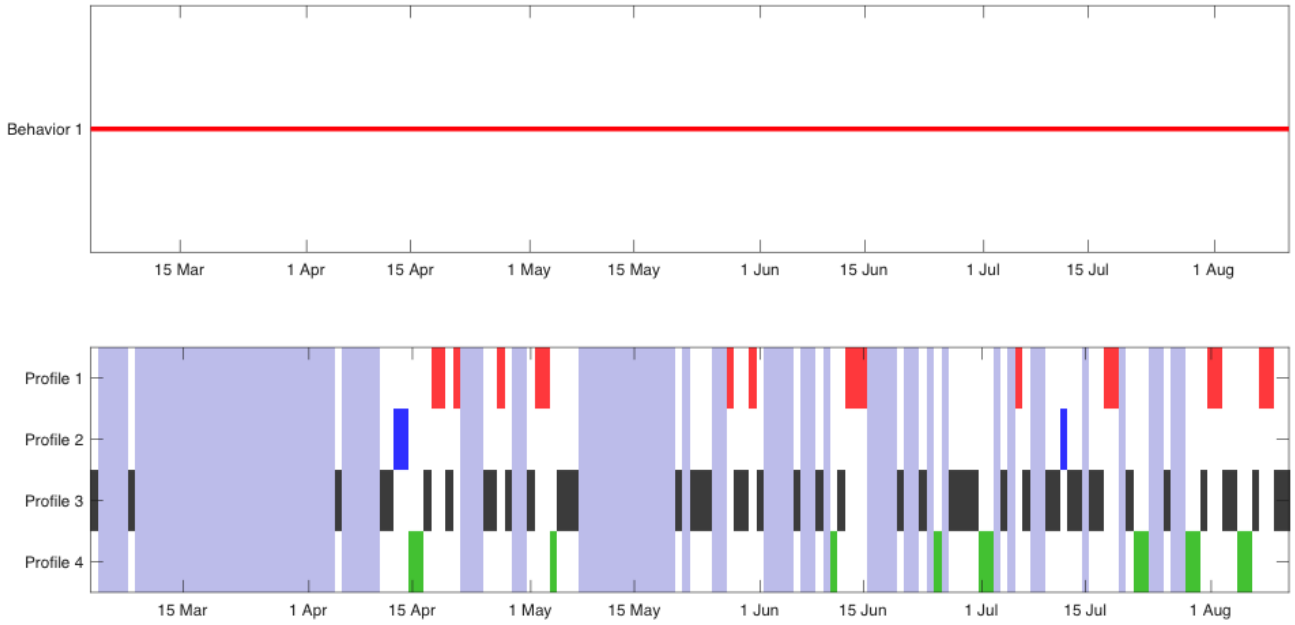


Figure 10. Representation of mobility pattern changes (above) identified by the technique and corresponding patterns (down) during study participation of patient D.



Patient E was a 42-year-old woman diagnosed with adjustment disorder with depressed mood and lumbar stenosis. She was prescribed a daily oral medication of escitalopram 15 mg, pregabalin 150 mg, and ketazolam 15 mg, besides antialgic medication. Fluctuations in the mood level were observed during follow-up in relation to back pain exacerbation.

This patient participated in the study from October 11, 2017 to December 21, 2017, and owned a BQ Aquaris M5 that ran Android 6.0.1. In addition, this patient showed improvement in depression scores during the study (June 23, 2017: PHQ-9=10; October 5, 2017: PHQ-9=6). In this case, as Figure 11 shows, the MDL criterion only selected 3 profiles, as the amount of data was rather small and, otherwise, would very likely have resulted in overfitting. Overall, 2 profiles corresponded to activity during the daytime, whereas profile 2 showed activity evenly distributed during the whole day. Figure 12 shows that the technique identified one change-point on November 25, 2017. Interestingly, this change-point appeared when profile 2 disappeared. The change-point coincided with an increase of painful osteoarticular symptoms.

Figure 11. Distance traveled profiles of patient E.

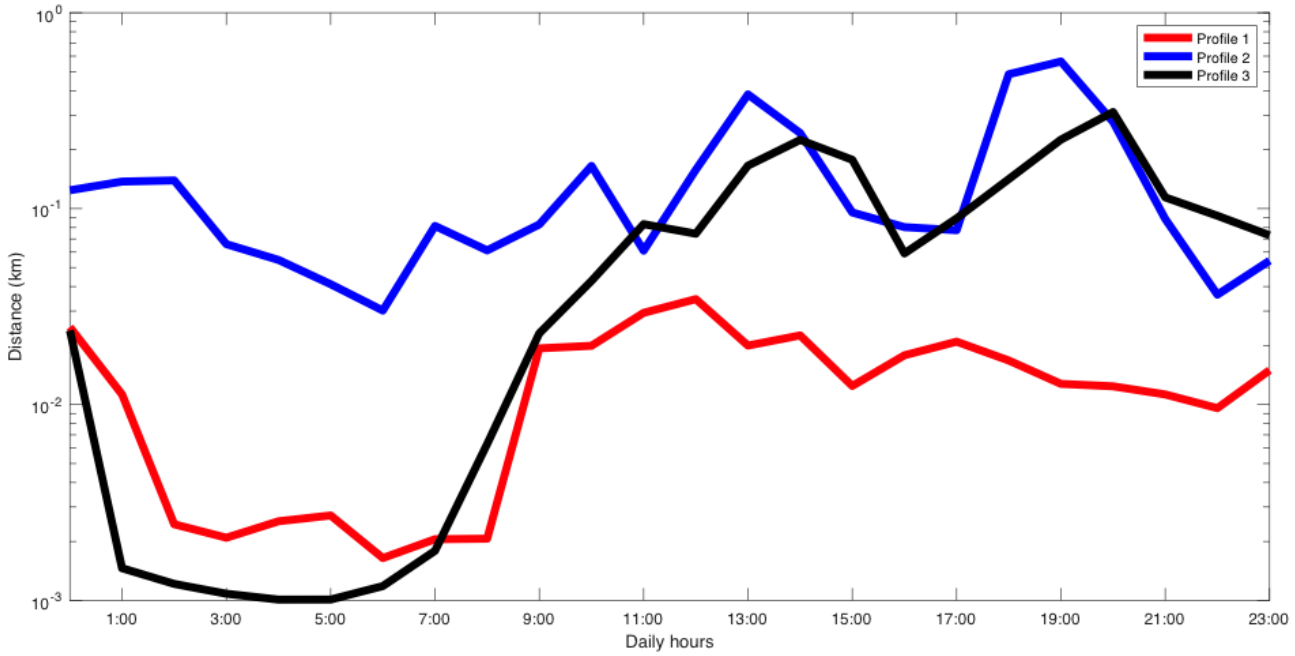
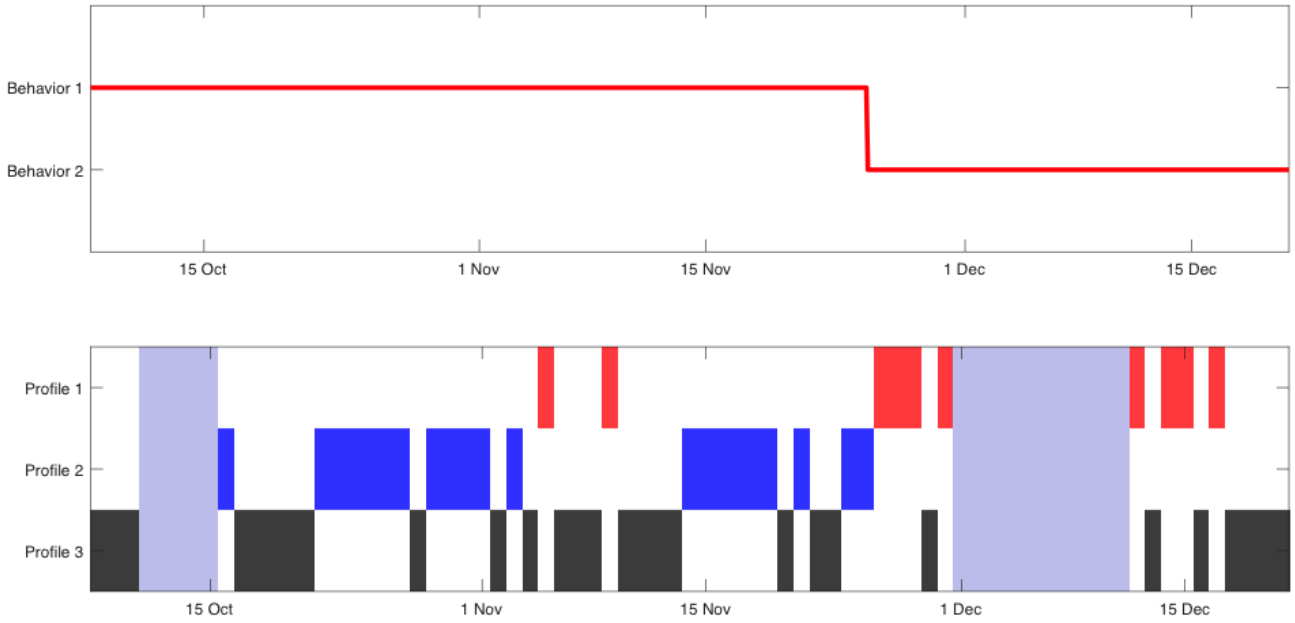


Figure 12. Representation of mobility pattern changes (above) identified by the technique and corresponding patterns (down) during study participation of patient E.



Discussion

Principal Findings

This study showed that the eB^2 system was capable of identifying mobility pattern changes, which may be used as proxies for behavioral changes and relapses. The technique was composed of 2 parts—a clustering algorithm to learn mobility profiles, which was based on a mixture of Gaussians model, and a change-point detector to identify probability changes of the mobility patterns. It is important to point out that detecting changes from one type of day to another does not suffice, what matters are the probability changes because we could have a type of day given by a typical workday and another one given by a typical weekend day; however, the change from the former to the latter (or *vice versa*) should not have been identified as a mobility pattern change.

This pilot study showed that the proposed technique could aid clinicians to detect relapses and other clinical changes. However, before its use in a clinical setting, the changes identified by the algorithm need to be interpreted. In this paper, we have shown the results from a few selected cases that may illustrate the potential applications of the eB^2 system in the outpatient follow-up of patients with depressive disorders.

The (possible) behavioral changes identification technique proposed in this study was based on the unsupervised processing of data from smartphone sensors. In particular, this work focused on detecting mobility pattern changes, which could be used as indicators of behavioral changes and only exploited the GPS

location data. The reasons were two-fold: (1) it yielded a relatively simple algorithm and (2) it admitted an easy clinical interpretation (more or less related to physical activity). However, as we have previously pointed out, the platform captured much more data, and the technique can be adapted also to exploit these additional data.

Our final goal would be the identification of more general behavioral changes (eg, Web-based social interaction) in outpatients, which has important applications for a wide range of chronic conditions, including mental health disorders. Apart from the continuous assessment of bioparameters themselves, smartphone-based monitoring would also allow researchers to gather information on context and environment, which may prove valuable for the interpretation of the monitored biomedical data (eg, information about weather conditions) and allow for a better interpretation of changes.

Clinical Contextualization of Smartphone Data

When the changes identified by eB^2 were contextualized in a given patient's routine, we were able to extract valuable information related to clinical changes. Thus, in our 5 selected patients, we identified different profiles of activity.

Interestingly, changes and different profiles represented different clinical scenarios. For instance, patients B and D showed no changes, whereas for patient A, the changes corresponded to a worsening. The algorithm detected this worsening on April 26, 2017 when the PHQ-9 depression score increased between April 6, 2017 and May 31, 2017. This participant did not show up for follow-up in September, although she continued using the eB^2 app and we cannot, therefore, establish clinical correlations from there on. Incidentally, a change-point was detected on September 1, 2017, which may be related to the drop-out from the follow-up. In patient D, the absence of changes reflected minimal clinical changes and stability in symptoms. However, patient B was an example in which mobility patterns were not useful for clinical purposes, as the proposed method did not identify any change, but there was, indeed, a clinical improvement. In this particular patient, the remaining data collected by the smartphone might be more useful, but this analysis is out of the scope of this work.

In addition, changes could represent both improvement and worsening, depending on the specific patient. On the one hand, the change identified on December 9, 2017 for patient C corresponded to a clinical improvement owing to the disappearance of increased activity during the night from that date onwards, reflecting a better night's sleep. In addition, a profile with low activity started to appear more often and, in fact, at this moment, the patient started to have a quiet lifestyle. In contrast, for patient E, a change represented a clinical worsening owing to the emergence of a profile of less activity and the disappearance of a profile of daytime activity. Both the emergence and disappearance of the above profiles indicated the worsening of the patient's condition owing to the exacerbation of her back pain.

Overall, these results highlighted that apps, such as eB^2, can be used for personalized psychiatry and that we are witnessing a paradigm shift from the traditional identification of shared factors in mental illnesses to individual and unique characteristics for each patient, that is, personalized medicine. A study presented EMA as the future of outpatient follow-up [18]. However, this technique strongly relied on patients' participation and was, therefore, prone to missing data [19].

Limitations

This study was conducted on a limited sample of patients with a limited time scale. Thus, it did not allow for the complete identification of ADLs; only mobility patterns could be identified. In addition, we did not have access to an ecological self-reported description of the patients' behavior. Ecological data are usually based on self-assessments and provide information that may be correlated with the digital phenotyping [19]. Ideally, we should have combined self-reported ecological data capture [20] with the results obtained by the eB^2 system to test whether the automatically detected changes correlated with the clinically diagnosed changes or data ecologically reported by patients. In this study, the algorithm detected changes in mobility patterns, which could be identified as behavioral changes. However, in this explanatory setting, we were not able to completely determine whether these behavioral changes identified by the algorithm corresponded to a clinical modification or the emergence of any normal or abnormal behavior. Moreover, we identified several factors that may explain the changes and which were not related to any modification in depressive symptoms. Furthermore, this study was based only on GPS data and many patients disabled this sensor; this is a problem that we will need to address in the future, and it is, therefore, important to convince patients not to disable the location in their smartphones. Nonetheless, location is not the only source of information, albeit it is important, and we should consider other types of data in future studies.

Data privacy is a serious concern in the eHealth research area. The eB^2 app captured data from smartphones, which possibly was a deterrent for patients to accept the app [21]. However, the selected patients were aware of the general approach of our method and were not very concerned about sharing their personal data as it was anonymized in the smartphone. Another major concern regarding personal electronic data is data security [22]. To uphold patients' privacy and reduce the risk associated with nonlegitimate access, all the sensible information stored in the eB^2 server was hashed and anonymized. Concretely, phone numbers, email addresses, Bluetooth, and Wi-Fi Media Access Control addresses were hashed using the SHA-1 algorithm, and the location was transformed using a noninvertible function. Specifically, we stored randomly rotated relative location coordinates, where the origin was the location that was most common during the first 3 days after installation (typically patients' home). Our app was (and still is) available through app stores, such as Google Play or App Store, which allowed us to continuously update and improve the app based on newly discovered bugs and also user feedback. For instance, we have improved the battery consumption, which should improve patient adherence in the future.

Future Application

Smartphone-based systems for managing and monitoring behavior present a highly promising field of innovation in health care. The normal use of a smartphone on a daily basis generates a larger amount of data than the amount that is typically collected in questionnaire-based studies or Web-based interventions; however, it requires that patients carry their smartphone most of the time. A smartphone sensor-based analysis already showed interesting results in the assessment of bipolar disorder [13], depression symptoms [8], prediction of schizophrenia [15], and sleep duration [23]. This work is in line with recent proposals of Torous et al., who established digital phenotyping as a promising method in the assessment of patients with mental health conditions [24].

We proposed a preliminary assessment of a method for patients with mental health conditions. Our system was able to identify changes in the mobility patterns of outpatients, which may correlate with behavioral changes and relapses. In the future, eB^2 may also be used for the assessment of physical activity in therapeutic programs or the identification of ADLs in the elderly [6].

Conclusions

We have developed a system that can capture data from the smartphone's native sensors and other wearables. The eB^2 system is composed of a smartphone app and a backend server. The preliminary results of the ongoing eB^2 study showed the feasibility of an unsupervised detection method for detecting mobility pattern changes, which we considered proxies for behavioral changes, in outpatients by exploiting the data acquired by the eB^2 app. So far, only location data were used, which resulted in relatively simple processing techniques and allowed for an easy clinical interpretation of the results. Of note, this method did not need intervention from patients. However, it was crucial that patients carried their phone all the time. With the development of the eB^2 system, we aimed to address most challenges raised by eHealth technologies in ecological monitoring.

Acknowledgments

This study was partly supported by Ministerio de Economía of Spain under project: AID (TEC2014-62194-EXP) and aMBITION (TEC2017-92552-EXP), the Ministerio de Economía of Spain jointly with the European Commission under projects ADVENTURE (TEC2015-69868-C2-1-R) and CAIMAN (TEC2017-86921-C2-2-R), the Comunidad de Madrid under project CASI-CAM-CM (S2013/ICE-2845), Fundación Jiménez Díaz Hospital, Instituto de Salud Carlos III (PI16/01852), Delegación del Gobierno para el Plan Nacional de Drogas (20151073), American Foundation for Suicide Prevention (LSRG-1-005-16), the French Embassy in Madrid, Spain, the fondation de l'Avenir, and the Fondation de France. The work of PMM has been supported by FPI grant BES-2016-077626.

Conflicts of Interest

None declared.

References

1. Wen CKF, Schneider S, Stone AA, Spruijt-Metz D. Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis. J Med Internet Res 2017;19(4) [FREE Full text]
2. Asselbergs J, Ruwaard J, Ejdys M, Schrader N, Sijbrandij M, Riper H. Mobile Phone-Based Unobtrusive Ecological Momentary Assessment of Day-to-Day Mood: An Explorative Study. J Med Internet Res 2016;18(3):A.
3. Varnfield M, Karunanithi M, Lee C, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779 [FREE Full text] [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
4. Hussain M, Li J, Brindal E, van Kasteren Y, Varnfield M, Reeson A, et al. Supporting the Delivery of Total Knee Replacements Care for Both Patients and Their Clinicians With a Mobile App and Web-Based Tool: Randomized Controlled Trial Protocol. JMIR Res Protoc 2017 Mar 01;6(3):e32 [FREE Full text] [doi: 10.2196/resprot.6498] [Medline: 28249832]
5. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, et al. Sleep Quality Prediction From Wearable Data Using Deep Learning. JMIR Mhealth Uhealth 2016 Nov 04;4(4):e125 [FREE Full text] [doi: 10.2196/mhealth.6562]
6. Qing Zhang, Karunanithi M, Bradford D, van Kasteren Y. Activity of Daily Living assessment through wireless sensor data. Conf Proc IEEE Eng Med Biol Soc 2014 Oct;2014:1752-1755. [doi: 10.1109/EMBC.2014.6943947] [Medline: 25570315]
7. Qing Z, Karunanithi M, Bradford D, van Kasteren Y. Activity of Daily Living assessment through wireless sensor data. Conf Proc IEEE Eng Med Biol Soc 2014;2014:1752-1755. [doi: 10.1109/EMBC.2014.6943947] [Medline: 25570315]
8. Saeb S, Zhang M, Karr C, Schueller S, Corden M, Kording K, et al. Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]

9. Bowie CR, Reichenberg A, Patterson TL, Heaton RK, Harvey PD. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. Am J Psychiatry 2006 Mar;163(3):418-425. [doi: 10.1176/appi.ajp.163.3.418] [Medline: 16513862]
10. Semkovska M, Bédard M, Godbout L, Limoge F, Stip E. Assessment of executive dysfunction during activities of daily living in schizophrenia. Schizophrenia Research 2004 Aug 1;69(2-3):289-300 [FREE Full text] [doi: 10.1016/j.schres.2003.07.005]
11. Silsbury Z, Goldsmith R, Rushton A. Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. BMJ Open 2015 Sep 15;5(9):e008430 [FREE Full text] [doi: 10.1136/bmjopen-2015-008430] [Medline: 26373402]
12. Reinertsen E, Clifford GD. A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses. Physiol Meas 2018 May 15;39(5):05TR01. [doi: 10.1088/1361-6579/aabf64] [Medline: 29671754]
13. Abdullah S, Matthews M, Frank E, Doherty G, Gay G, Choudhury T. Automatic detection of social rhythms in bipolar disorder. J Am Med Inform Assoc 2016 Dec;23(3):538-543 [FREE Full text] [doi: 10.1093/jamia/ocv200] [Medline: 26977102]
14. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, et al. Voice analysis as an objective state marker in bipolar disorder. Transl Psychiatry 2016 Dec 19;6:e856 [FREE Full text] [doi: 10.1038/tp.2016.123] [Medline: 27434490]
15. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
17. Murphy J. Machine Learning: A Probabilistic Perspective. USA: MIT Press; 2012.
18. van Os J, Delespaul P, Barge D, Bakker R. Testing an mHealth momentary assessment Routine Outcome Monitoring application: a focus on restoration of daily life positive mood states. PLoS One 2014;9(12):e115254 [FREE Full text] [doi: 10.1371/journal.pone.0115254] [Medline: 25513813]
19. Torous J, Onnela J, Keshavan M. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. Transl Psychiatry 2017 Dec 07;7(3):e1053 [FREE Full text] [doi: 10.1038/tp.2017.25] [Medline: 28267146]
20. Barrigón ML, Berrouiguet S, Carballo JJ, Bonal-Giménez C, Fernández-Navarro P, Pfang B, MEmind study group. User profiles of an electronic mental health tool for ecological momentary assessment: MEmind. Int J Methods Psychiatr Res 2017 Mar;26(1). [doi: 10.1002/mpr.1554] [Medline: 28276176]
21. Wu E, Torous J, Hardaway R, Gutheil T. Confidentiality and Privacy for Smartphone Applications in Child and Adolescent Psychiatry: Unmet Needs and Practical Solutions. Child Adolesc Psychiatr Clin N Am 2017 Dec;26(1):117-124. [doi: 10.1016/j.chc.2016.07.006] [Medline: 27837937]
22. Fernández-Alemán JL, Carrión Señor I, Lozoya P, Toval A. Security and privacy in electronic health records: a systematic literature review. J Biomed Inform 2013 Jun;46(3):541-562 [FREE Full text] [doi: 10.1016/j.jbi.2012.12.003] [Medline: 23305810]
23. Kaufmann C, Gershon A, Eyler L, Depp C. Clinical significance of mobile health assessed sleep duration and variability in bipolar disorder. J Psychiatr Res 2016 Dec;81:152-159 [FREE Full text] [doi: 10.1016/j.jpsychires.2016.07.008] [Medline: 27451108]
24. Torous J, Kiang MV, Lorme J, Onnela J. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Ment Health 2016;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]

Abbreviations

ADL: activities of daily living
API: application programming interface
eB ^{2}: Evidence-Based Behavior
eHealth: electronic health
EMA: ecological momentary assessment
GPS: global positioning system
MDL: minimum description length
PHQ-9: Patient Health Questionnaire-9

Edited by J Torous, R Calvo, M Czerwinski, G Wadley; submitted 04.01.18; peer-reviewed by S Hussain, CKF Wen, M Varnfield; comments to author 16.02.18; revised version received 18.05.18; accepted 10.09.18; published 10.12.18.

Please cite as:
Berrouiguet S, Ramírez D, Barrigón ML, Moreno-Muñoz P, Carmona Camacho R, Baca-García E, Artés-Rodríguez A
Combining Continuous Smartphone Native Sensors Data Capture and Unsupervised Data Mining Techniques for Behavioral Changes Detection: A Case Series of the Evidence-Based Behavior (eB2) Study
JMIR Mhealth Uhealth 2018;6(12):e197
URL: https://mhealth.jmir.org/2018/12/e197/
doi:10.2196/mhealth.9472
PMID:30530465

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org