Original Research

DIGITAL HEALTH

DIGITAL HEALTH
Volume 9: 1–11

DOI: 10.1177/20552076231186515
journals.sagepub.com/home/dhj

Sage

Usability evaluation of mobile phone technologies for capturing cancer patient-reported outcomes and physical functions

Ingrid Oakley-Girvan[1,2], Reem Yunis[1], Stephanie J Fonda[3], Elad Neeman[4], Raymond Liu[4,5], Sara Aghaee[5], Maya E Ramsey[5], Ai Kubo[5] and Sharon W Davis[1]

Abstract

Background: By eliminating the requirement for participants to make frequent visits to research sites, mobile phone applications ("apps") may help to decentralize clinical trials. Apps may also be an effective mechanism for capturing patient-reported outcomes and other endpoints, helping to optimize patient care during and outside of clinical trials.

Objectives: We report on the usability of Digital BioMarkers for Clinical Impact (DigiBioMarC™ (DBM)), a novel smartphone-based app used by cancer patients in conjunction with a wearable device (Apple Watch®). DBM is designed to collect patient-reported outcomes and record physical functions.

Methods: In a fully decentralized "bring-your-own-device" smartphone study, we enrolled 54 cancer patient and caregiver dyads from Kaiser Permanente Northern California (KPNC) from October 2020 through March 2021. Patients used the app for at least 28 days, completed weekly questionnaires about their symptoms, physical functions, and mood, and performed timed physical tasks. Usability was determined through a subset of the Mobile App Rating Scale (MARS), the full System Usability Scale (SUS), the Net Promoter Score (NPS), and semi-structured interviews.

Results: We obtained usability survey data from 50 of 54 patients. Median responses to the selected MARS questions and the mean SUS scores indicated above average usability. The NPS from the semi-structured interviews at the end of the study was 24, indicating a favorable score.

Conclusions: Cancer patients reported above average usability for the DBM app. Qualitative analyses indicated that the app was easy to use and helpful. Future work will emphasize implementing further patient recommendations and evaluating the app's clinical efficacy in multiple settings.

Keywords

BYOD, decentralized clinical trial, ePROs, wearable sensors, connected sensors, neoplasm, cancer, mHealth, patient-reported outcomes, mobile application, usability, smartphone

Submission date: 22 September 2022; Acceptance date: 20 June 2023

Introduction

Early detection methods and treatments tested in clinical trials have helped to increase the number of cancer survivors from 14 million in 2012 to a projected 18 million in 2022.[1,2] However, adult clinical trials often have difficulty achieving trial recruitment and retention goals,[3] and thus reach only a small portion of potentially interested participants.[4] Limited availability of trials at a patient's medical care facility is a major barrier to clinical trial participation[4]

[1]Strategy and Science Departments, Medable Inc., Palo Alto, CA, USA
[2]The Data and Technology Proving Ground, The Public Health Institute, Oakland, CA, USA
[3]Estenda Solutions, Inc., Wayne, PA, USA
[4]San Francisco Medical Center, Kaiser Permanente Northern California, San Francisco, CA, USA
[5]Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA

Corresponding author:
Ingrid Oakley-Girvan, Strategy and Science Departments, Medable Institute, Medable Inc., 525 University Avenue, Suite A70, Palo Alto, CA, 94301, USA.
Email: ingrid@medable.com

and is one reason trial populations rarely reflect actual patient demographics.[5] Recently, the COVID-19 pandemic caused 60% of research programs to temporarily halt screening and/or enrollment for oncology clinical trials, further limiting trial availability.[6]

Decentralized clinical trials (DCTs) represent an opportunity to reach more trial participants.[7] This can be achieved partly by shifting many trial tasks from in-person clinics to digital platforms using web-based or smartphone applications.[8] Smartphones are nearly ubiquitous in the United States, even among older- and lower-income populations,[9] and can be used for many tasks including online screening and recruitment, electronic informed consent, remote monitoring of participants through patient-reported outcomes (PROs) and digital clinical outcome measures derived from the smartphones themselves, or wearable sensors.[8] The FDA has recognized the increasing use and importance of PROs in clinical trials.[10] Moreover, integrating PROs into clinical research has the potential to detect and better manage adverse events and toxicities.[11,12] There is tremendous promise for remote monitoring by combining digital data from sensors with electronically collected PROs (ePROs), but these need to be user friendly to be embraced by participants and clinical teams.

We developed the smartphone app Digital BioMarkers for Clinical Impact (DigiBioMarC™ (DBM)) for the iPhone® and integrated with the Apple Watch® to enable patient screening and participation in DCTs. We included electronic informed consent and ePROs, digital monitoring of clinical outcome measures, and provided patient resources. We recruited a sample of cancer patients receiving intravenous (IV) chemotherapy or immunotherapy to use the app regularly for a period of approximately 28 days. We focused on this population because patients receiving these therapies tend to experience substantial symptom burdens.[13] To be eligible, cancer patients had to have an informal caregiver. Both the cancer patient and their caregiver were consented and enrolled. The DBM app was used by patients while the informal caregivers used a different app called TOGETHERCare. This article reports on the usability of the DBM app from the perspective of these patients as determined by usability questionnaires and semi-structured interviews.

Methods

Application development

The initial version of the DBM app was developed with funding from the National Institutes of Health using a three-phase development process, which several of the authors established and have used in the past,[14] that involves patients, caregivers, and healthcare professionals. During Phase I, we conducted unstructured interviews by phone and in-person with stakeholders to understand the current care gaps and needs in cancer care outside of the clinical setting and did internal development and testing (Alpha test). In Phase II (Beta test), we administered semi-structured interviews with three patients, three caregivers, and three clinicians, and we also collected feedback from seven patients from academic cancer centers at Duke University and Stanford University, who used version one of the app for 28 days. The app was iterated based on that feedback, and version two was developed. In Phase III, in collaboration with Kaiser Permanente Northern California (KPNC) and reported in this manuscript, we tested the usability of version two in a larger sample of cancer patients.

Application content and functionality

The DBM app uses functions built into the iPhone (bring your own device (BYOD) only) and Apple Watch (BYOD or study provisioned) to collect biometric data (e.g. step count, distance walked, etc.) and to perform standardized gait and sit-to-stand tests. The ePROs and feedback surveys are built into the application using questionnaires presented at specific intervals. The DBM app consists of three main sections: ***Profile*** for the study title, informed consent documents, and Health Insurance Portability and Accountability Act (HIPAA) authorization; ***Tasks*** for Apple Watch health permissions, active tasks, and surveys including demographics and ePROs (Table 1); and ***Resources*** for links to contact the research team and, because the study was launched at the onset of the recent pandemic, to access a KPNC COVID-19 informational website for patients.

Participant recruitment

We identified potential participants using the KPNC electronic health record (EHR) system. The KPNC serves over 4.5 million members in northern California with an integrated healthcare delivery system. Recruitment was initiated by email invitations after the approval, or the absence of disapproval, from the patients' primary oncologists. Recruitment activities were conducted from October 2020 through March 2021. Recruitment and consent were completed remotely through a televisit or phone call, with no need for the patients to come to the clinic. Inclusion criteria required patients to be adults 18 years of age or older, to be KPNC members with a cancer diagnosis, to be receiving IV chemotherapy or immunotherapy treatment, to be English speakers, to own an iPhone 6 or higher, and to have an eligible family member or a friend who acted as the primary caregiver and was willing to participate by evaluating the TOGETHERCare™ app for cancer caregivers.[14] Patients with severe mental illness or insufficient cognition to consent were excluded from the study. The study participants were asked to use the app for at least 28 days, during which the mobile app delivered surveys and activity requests (see Table 1). Recruitment and

Table 1. DBM application: patient-reported outcome surveys, tasks, frequency, and completed percentage.

Survey/Task	Scientific/Tool name(s) (maximum number of questions[a])	Frequency (*n* times)	Completed (%)
About You	Cancer health literacy test-6 (CHLT-6)[15] and demographics (number of questions = 13[a])	Baseline (once)	100
Short Symptom Report	National Cancer Institute Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)[16]–subset of symptoms selected by subject matter experts, Self-rated Eastern Cooperative Oncology Group performance status (ECOG-PS)[17], and Sleep and fatigue questions (number of questions = 44[a])	Weekly (5 times)	100
Fast 4	Patient-Reported Outcomes Measurement Information System (PROMIS) Item Bank v2.0–Physical Function Short Form 4a[18] (number of questions = 4)	Weekly (5 times)	99.6
COVID-19, Your Input	COVID-19 testing, behavior, impact on care, and knowledge questions (number of questions = 35[a])	Baseline Study end (twice)	100
Quick Check-in	Financial and social distress, four patient engagement questions, and the four-item Patient Health Questionnaire for anxiety and depression (PHQ4)[19] (number of questions = 11)	Baseline Study end (twice)	100
Gait Pace	Gait and balance predefined active task[20] (20 steps plus standing still for 30 s)	Weekly (5 times)	92.8
Sit and Stand	Sit to stand predefined active task[21] (over a 30-s duration)	Weekly (5 times)	94.8
App Feedback	System Usability Scale (SUS)[22,23], the Net Promoter Score (NPS)[24], Mobile App Rating Scale (MARS)[25]–five questions about functionality and visual appeal (number of questions = 16)	Baseline Halfway Study end (3 times)	100

Note: For each task, the proportion of completions was calculated as follows: 100 (the number of total task completions/(number of participants multiplied by the number of expected times the task was to be completed during the study period)).
[a]The maximum number of possible questions because in some surveys, additional questions were predicated on prior responses or allowed additional free text input.

compliance were monitored through an associated site application. Patients who completed the surveys in the app and two semi-structured virtual interviews received a $100 Amazon gift certificate.

Ethics approval

A KPNC research associate explained study procedures to potential participants during an eligibility confirmation phone call. Eligible and interested patient–caregiver dyads then received an email with the informed consent and HIPAA documents and reviewed the consent and HIPAA form in a subsequent video or phone call with the research associate. Informed consent was collected verbally during the second interaction, and those who elected to enroll in the study downloaded the app and provided written informed consent electronically within the app prior to proceeding with the study. The app was housed on the patient's iPhone, which was designed by Apple to be password protected. Moreover, the app also had a password associated with it, thus providing another layer of security. The app, once downloaded on the participants' phone, could be opened by either a passcode, facial recognition, or fingerprint depending upon how the user preferred to set it up for controlled access. These biometrics were not available to KPNC or Medable but were stored on the participant's own phones as part of the Apple iPhone operating system.

The Medable platform use was assessed by Kaiser IT and technology teams along with security and privacy for all data captured by the iPhone, and the study as a whole was reviewed and approved by the KPNC Institutional Review Board. All data collected using the app was

stored in a secure HIPAA compliant cloud that was accessed and controlled by the KPNC project staff.

Procedures

After signing the informed consent, participants were prompted by the app to complete a series of HIPAA-related tasks and to choose whether to allow phone-based health data collection, and this concluded the study enrollment process. Participants who did not already own an Apple Watch were loaned one for the duration of the study; all were instructed to pair their borrowed or personal Apple Watch with the DBM app and to activate a health data permissions feature. Participants were then asked to use the app and Apple Watch until all surveys were completed over the course of approximately 28 days. At the end of the study period, the app was disabled and all health data collection was ended.

Measures

The DBM app collected quantitative usability measures at the baseline, midway at approximately 14 days post enrollment, and at the study end at approximately 28 days post enrollment. Measures used included five questions from the Mobile App Rating Scale (MARS),[25] the entire System Usability Scale (SUS),[22,23,26] and the Net Promoter Score (NPS).[24] The five MARS questions pertain to functionality and visual appeal and were selected to keep response time to a minimum. The possible responses are on a scale of 1–5, with 1 being the worst and 5 being the best.[25] The SUS has 10 Likert-scaled items ranging from "strongly agree" (1) to "strongly disagree" (5). The SUS scores range from 0 to 100, with a higher score indicating better product usability. A total SUS score above 68 is considered above average.[27] The SUS also has two sub-scores, one for general usability and one for learnability. The NPS is derived by asking respondents to indicate the likelihood that they will recommend a product to a friend or colleague using a 0–10 scale[24] and is used broadly in product development and market assessment efforts. The percentage of detractors (defined as a score of 0–6) is subtracted from the percentage of promoters (scores of 9–10), resulting in the final NPS. Quantitative usability measures used for this analysis were calculated from the third round of surveys at the study end of approximately 28 days, when participants had the most experience with the app. Demographics of the study participants (e.g. age at informed consent, sex, cancer type) were self-reported as part of the "About You" task or were derived from the KPNC EHR.

We conducted semi-structured interviews with participants on approximately the 7th day of app use and after participants completed the final app surveys, using a secure videoconferencing program. An approximately 15-min Day 7 interview was completed by 54 patients, which provided an opportunity to discuss useful or confusing elements of the app, including the frequency of the surveys, the length of time it took to complete them, whether the app included information that the patient would like their clinical team to know, app rating scales, and whether there were other features not present that they would like to see. An approximately 30-min final interview was completed by 50 patients, and included topics from the Day 7 interview, a detailed explanation of the ultimate goals of the app, mockups of additional feedback screens based upon specific participant input, and questions about the usefulness of possible future features as part of our ongoing iterative development efforts.

Analysis

We collected descriptive statistics of study participant quantitative data, including patient background characteristics, number of completed in-app tasks, and responses to the end of the study MARS, SUS, and NPS surveys. For analysis of the semi-structured interviews, all were first transcribed verbatim, and the resulting transcripts were then analyzed by a member of the research team to identify patterns or themes.[28] Theme coding was then reviewed by a second researcher (SWD). Any discrepancies were reviewed by additional members of the team and resolved with input from the lead researcher (IOG). All Day 7 interviews were coded. After coding and carefully analyzing half of the final interviews, we determined that saturation had been reached and that the themes expressed were consistent throughout the study period. Thus, the interview analysis presented here is based on the common themes and content from the Day 7 interviews and the coded final interviews.

Results

Recruitment and retention

A total of 2155 potential patients received recruitment emails containing brief study details and eligibility requirements that instructed patients interested in participating to reach out to the study coordinator to learn more. Of the 2155 patients contacted, 247 responded. From the respondents, 166 were determined to be ineligible (physician indicated a contraindication to participation ($n = 42$), invalid emails ($n = 41$), no iPhone ($n = 33$), no eligible/willing caregiver ($n = 24$), no scheduled IV therapy during the study period ($n = 13$), not English speaking ($n = 5$), deceased ($n = 2$), ineligible unspecified ($n = 6$)), 20 declined to participate before eligibility could be confirmed, and 7 decided not to participate after learning more details about the study. A total of 54 patients were enrolled, with 50 of these ultimately completing the full study period, yielding a high retention rate with only 7.4% attrition[29]

(three patients elected to stop participating after enrolling due to their health and one passed away). Twelve of the 50 participants had their own Apple Watch; all other participants used a loaned device mailed to them at no cost. The analysis dataset comprises data from the 50 participants completing the full study period.

Background characteristics of respondents

The mean age of participants at the time of consent was 59.9, with a standard deviation (SD) of 11.6, range 31–83. A majority (86%) of participants were not of Hispanic or Latino descent, and 68% of patients identified as White (Table 2). Most (78%) participants were female. Almost all (92%) participants had completed at least some college. About half (52%) were retired or unemployed, with 32% self-employed or working full- or part-time. Almost all (98%) had adequate cancer health literacy as indicated by scores of 5 or 6 on the Cancer Health Literacy Test (CHLT-6).[15] Breast cancer patients represented about one-third of the study participants (34%), with gynecological (24%), gastrointestinal (24%), thoracic (12%), and other (6%) making up the remainder.

Task completion

Study participants completed 97.7% of all app tasks expected per the study protocol. The percentage of tasks completed ranged from 92.8% for the Gait Pace task to 100% for most other tasks (Table 1).

Mobile app rating scales

Feedback survey scores are shown in Table 3. The mean MARS functionality sub-score was 4.3 ± 0.8 out of a maximum of 5. Visual appeal was rated 3.2 ± 1.0, which is consistent with comments from qualitative interview responses about the lack of colors and graphics (described below). The total SUS (76.6 ± 18.5), usability (74.0 ± 19.7), and learnability (86.8 ± 21.5) sub-scores were above average.[30] For the last in-app NPS, the number of "detractors" (42.9%) outnumbered "promoters" (24.5%). However, the NPS score from the last qualitative interviews was substantially higher at 24 (see Table 3).

Semi-structured interviews

Most patients said that they found the app easy to use and helpful. They appreciated the prompts to increase their activity levels, which were a feature of the Apple Watch. Patients indicated that they wanted to receive more feedback from the app, and had suggestions for enhancing the content, frequency, and timing of the survey questions. While patients appreciated the simplicity and clinical appearance of the app, some suggested enhanced visuals and features. They also suggested notifications for when tasks were scheduled in the app, and suggested clarifying instructions for specific tasks. Table 4 describes identified themes and concepts and provides some specific comments.

Discussion

Principal findings

The DBM smartphone app is designed to collect PROs as well as physical function data from cancer patients. This mixed-methods usability study found that the DBM app worked well and was easy to learn. The SUS scores indicated an above average rating. The SUS score on the learnability questions showed high approval for the ease with which participants were able to learn to use the app, and this result was consistent with the results of five MARS questions pertaining to functionality and visual appeal.

The low in-app NPS score ("How likely is it that you would recommend DBM to a friend or colleague, on a scale of 10 with 0 being not at all likely and 10 being extremely likely?") was inconsistent with the SUS and MARS score data. This may have been partially attributable to the fact that many participants expressed negative feelings about their disease, and many indicated that they did not have any friends or colleagues with cancer. While it is possible that the higher NPS obtained from the interviews may have been influenced by participants wishing to please the interviewer, we believe it was partially attributable to a better understanding conveyed by the interviewer to participants that the NPS question was a general one referring to potential users and did not require the participant to have friends or family with cancer, and to the inclusion of mock-ups of data visualizations based upon previously provided specific participant input. Higher NPS scores provided during the interviews may also have resulted from improved participant knowledge about the purpose of the app that had not previously been described. Specifically, during the final interview, participants were told that the ultimate intention for the app was to enable remote patient monitoring with real-time data sharing and feedback to participants and clinical teams. Once the purpose was provided, multiple participants agreed that the app would be helpful; for example, "Now that I know kind of what the purpose of this is, it could be really valuable. There's a thing called white coat syndrome; sometimes when you go to the doctor, you don't always say everything that's on your mind or you don't remember, especially after a bunch of treatment, you start getting kind of fuzzy because it works on your brain too." Other participants indicated if they knew the data was being shared with their clinical team, they would be certain to enter the data in the app.

Table 2. Participant background characteristics at the time of consent ($N = 50$).

	Number (%)
Gender	
Female	39 (78%)
Male	11 (22%)
Ethnicity	
Hispanic or Latino	4 (8%)
Not Hispanic or Latino	43 (86%)
Missing	3 (6%)
Race[a]	
White	34 (68%)
Black	4 (8%)
Asian	3 (6%)
Native Hawaiian or Pacific Islander	2 (4)
Multiracial	3 (6%)
Other	3 (6%)
Missing	1 (2)
Educational attainment	
Less than 9th grade	1 (2%)
High school/GED	3 (6%)
Some college, no degree	11 (22%)
Associate's degree	6 (12%)
Bachelor's degree	14 (28%))
Master's degree	7 (14%))
Doctorate or other professional degree	8 (16%)
Employment status	
Unemployed	5 (10%)
Disability leave	5 (10%)
Homemaker	3 (6%)
Retired	21 (42%)
Self-employed	2 (4%)
Working part-time	3 (6%)
Working full-time	11 (22%)
Cancer Health Literacy Test (CHLT-6)	
Likely limited literacy	1 (2%)
Likely adequate literacy	49 (98%)
Cancer type	
Breast	17 (34%)
Gynecological	12 (24%)
Gastrointestinal	12 (24%)
Thoracic	6 (12%)
Other	3 (6%)

[a]Participants that indicated more than one race are listed as "Multiracial." No participant has been identified as an American Indian or an Alaska native.

Improving engagement

From a product design perspective, the NPS score at the final interview was 42 points higher than the score from the last in-app NPS. Final interviews took place after app usage was completed, when we described, for the first time, the ultimate purpose of the app. Because participants had previously indicated in interviews a desire to see the data they had entered, we mocked up patient-focused data visualizations and shared those during the final interview, but focused on data that would not cause unblinding or bias in a clinical trial setting. The observed 42-point difference demonstrates how valuable it is to obtain participant input and to implement suggestions such as data visualizations and providing information about the apps' intended use when developing a mobile tool. Inclusion of decentralized elements, like remote data collection with regularly scheduled tasks, will be further enhanced if we listen to what helps motivate people to stay engaged and use the product as intended. For instance, participant comments about timing notifications for when tasks are due is a critical aspect and is now reflected in the revised app. Other elements such as clear task instructions that may address technical literacy or feelings of inadequacy, a colorful

Table 3. Study participants' MARS, SUS and NPS final survey scores.

MARS	Mean ± SD, Median
How accurately/fast do the app features (functions) and components (buttons/menus) work?	4.2 ± 1.0, 4.0
How easy is it to learn how to use the app; how clear are the menu labels/icons and instructions?	4.2 ± 0.8, 4.0
Is moving between screens logical/accurate/ appropriate/uninterrupted; are all necessary screen links present?	4.3 ± 0.9, 5.0
Are interactions (taps/swipes/pinches/scrolls) consistent and intuitive across all components/screens?	4.4 ± 0.8, 5.0
MARS Functionality sub-score	4.3 ± 0.8, 4.8
How good does the app look?	3.2 ± 1.0, 3.0
SUS 10[a]	Mean ± SD, Median
Total SUS score[31]	76.6 ± 18.5, 78.8
Usability sub-score	74.0 ± 19.7, 78.1
Learnability sub-score	86.8 ± 21.5, 100.0
Total SUS score percentile rank range[30]	70–79
NPS in app[a]	Number (%)
0–6 (Detractors)	21 (42.9%)
7–8 (Passives)	16 (32.7%)
9–10 (Promoters)	12 (24.5%)
Net Promoter Score (Promoters % – Detractors %)	−18.4
NPS from the final semi-structured (qualitative) interview	Number (%)
0–6 (Detractors)	14 (28%)
7–8 (Passives)	10 (20%)
9–10 (Promoters)	26 (52%)
Net Promoters Score (Promoters % – Detractors %)	24

MARS: Mobile App Rating Scale; NPS: Net Promoter Score; *n*: number; SD: standard deviation; SUS: System Usability Scale.
[a]One participant did not respond to one or more SUS questions, so percentages are based on 49 participants for this measure.

appearance, and a clear purpose with clinical utility also appeared critical or very important for engagement.

Limitations

The current study was not without limitations. First, most patients in the study were female with high education and health literacy levels and we only included participants who could speak English. While the overall population served by KPNC broadly represents the corresponding local demographics, the participants from the convenience sample represent those with an interest in the study, an iPhone, an eligible informal caregiver, and those undergoing IV therapy. The findings from these patients may not be generalizable to populations that are most underrepresented in clinical trials. Recruitment across a wider demographic is essential to further evaluate DBM across various cultures and backgrounds. An Android version of DBM has now been developed and will be used in future studies with a focused effort to enroll across all

Table 4. Qualitative analysis themes with illustrative comments.

Themes	Concepts	Selected Comments
1. The app is easy to use	Most patients said that the app was easy to use and had clear instructions about what to do next.	"...it's just very clear, and uncluttered. Um? You just click on the thing you want, and it tells you what to do. So, there's no mystery to it." "It's been really easy. I haven't had any problems with the app at all."
2. The app is helpful	Patients found the app and the Apple Watch helpful for prompting them to increase their level of movement during the day.	"Because it keeps you motivated and keeps you active by requesting you to do those things." "Well, yes. I think it is helpful. I mean, to be keeping track of this every week to see how I'm feeling or how I'm describing I'm feeling or how my health is improving or decreasing, and someone is keeping track."
3. Additional feedback from the app would be useful	A recurring theme was the patients' desire for feedback, so that they could gauge their progress over time.	"OK, perhaps feedback for the week or whatever..." "I guess it would be nice to see progress or regress over the months that I had it for."
4. The content, frequency and timing of some of the survey questions could be improved.	Respondents offered several suggestions based on their experience with the app, such as including open-ended questions, greater clarity on the time frame for which their responses were applicable, a wider possible scale for their responses, and more specific, less general questions. They also suggested timing the frequency of the surveys to how far along they were in their treatment plan.	"It's kind of vague as far as like pain. Like, I had multiple types of pain. I had leg cramping, but I also had headaches so maybe a little bit more specific on the type of pain, OK?" "I mean, to me, if you're asking me how many times did I have diarrhea last week, it's a lot easier for me to say three times as opposed to have to think, "Well, I had diarrhea three times. So would that be often or very often?" Who knows what those words mean, really (laughter)?" "But if you were going through those harsher [treatment] cycles, your answers are vastly different."
5. Enhance visuals in the app and add reward features	Although they appreciated the simplicity of use and clinical appearance, most of the patients suggested that the app interface needed more color, or to include features such as "gamification," where they could receive points for completing tasks.	"It's a nice clean screen. Nothing flashing. The information you need presented in a clean manner. I like that." "There's no color. You know everybody loves color." "Kind of like oh, I got an award for. I got a response from making a. You know it would make me want to come back more often."
6. Notifications/Instructions	Almost all patients said they wanted more reminders and notifications when their tasks were due. Notifications should be coordinated with days when tasks were scheduled. Participants also indicated that the instruction for movement tasks (e.g.: sit and stand) needed to be clearer.[a]	"So, I didn't get any notification that I needed to look at my tasks, but I looked at it during the week to see if I needed to." "Actually, where it says tasks, when it was blank, I didn't know if that was a mistake. So maybe it should say no tasks today or something." "Because to open it and having nothing there. It's kind of like opening your Christmas and having it empty three days and three days in a row of that, it's, you know. Not as motivating on the 4th day to open it right?"

[a]We used standard publicly available instructions for established tasks, indicating a gap that needs to be addressed for better participant understanding and high-quality adherence.

demographics. Second, the study was not designed at this point to provide the collected information in real time directly to the clinical team; this may have affected usage and participant concerns about feedback. Future studies that provide real-time, clinically relevant ePROs, and digital data directly to clinical teams will most likely reduce these concerns. Third, in some instances, it is difficult to distinguish features of the app itself from those of Apple Watch. For instance, "motivation" to move may come from the app or the Apple Watch's push notifications. Finally, although our sample size of 50 participants is considerably larger than the <12 participants in many published mHealth app usability studies,[32] these study results are not necessarily generalizable to those without a caregiver or across the general population.

Strengths

Despite some limitations, our study had many advantages. Using KPNC's comprehensive EHR, we were able to efficiently and rapidly identify dyads that were potentially eligible to participate in the study. We were then able to demonstrate the feasibility of a fully remote approach using a smartphone app in an active treatment cancer population that reflects real-world conditions. While there have been other apps that have been used to obtain PROs from cancer patients with a specific cancer or undergoing a specific treatment,[33–38] there are not many apps that have been shown to be feasible or usable in a broad range of cancers across multiple treatments in a real-world setting. The tested app is also unique in its integration of a digital activity monitor (Apple Watch). The willingness of users to complete multiple surveys over the course of the study on their own smartphone highlights the benefits of bringing your own device[39] and is encouraging for the expansion of DCTs, which require less travel time for participants. Moreover, testing the DBM app helped to elucidate the potential for the development of in-app alerts sent to the study team if distress measures completed by participants exceeded a predetermined threshold, thus enabling remote symptom monitoring of cancer patients enrolled in clinical trials or standard care. An associated site application enabled the research associate to monitor both recruitment and compliance with in-app task completion; this could be beneficial to improving data capture as well as limiting missing data and ultimately regulatory submissions for new therapies.

Future directions

We are further refining DBM for to help improve screening of potential candidates for clinical trials. Screening will include the ability to tailor the product to include questionnaires and active tasks for specific therapeutic areas and study protocols. Further evaluations of DBM should include an assessment of whether it is usable from the standpoint of clinical trialists. Determining where in the workflow it fits, and whether it improves benchmarks such as trial speed, efficiency in monitoring adverse events, and study endpoint data collection within the clinical trial protocol are areas for further investigation.

Conclusions

Our study confirms the usability of the DBM app, a novel BYOD smartphone-based app used by cancer patients in conjunction with a wearable device (Apple Watch) using a best practices approach described previously.[14] DBM is designed to collect PROs and record physical functions. Cancer patients who used the DBM app over the course of approximately a month reported above average usability. Qualitative analyses indicated that the app was easy to use and helpful. DBM appears to be feasible to use during standard treatment to provide more in-depth reporting of mental and physical health over time and could also be used for remote clinical trial activities and symptom monitoring. In addition, remote evaluation of patients for eligibility could help reduce both the study site and participant burden and improve the study efficiency while driving faster study timelines. Future work will emphasize implementing the patient recommendations and evaluating the app's efficacy for symptom monitoring across a variety of therapeutic areas.

Acknowledgments: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This study could not have been done without the able assistance of Anushka Gupta, who conducted initial data analysis, Vandana Shah and Yasamin Miller, who coded the semi-structured interviews and contributed to the thematic analysis, and Elaine M. Kurtovich, KPNC Research Project Manager.

Author contributions: IOG and SWD researched literature and conceived the study. EN, RL, MER, and AK were involved in protocol development, IRB approval, and patient recruitment. SA prepared data files and data dictionaries. SJF provided data analysis with input and direction from IOG and SWD. IOG and RY completed the semi-structured interviews. SWD wrote the first draft of the manuscript and IOG provided significant edits and input. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests: The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: IOG, SWD, and RY are employed by Medable Inc., which developed the DBM app with funding from the National Institutes of Health through the Small Business Innovation Research program with the National Cancer Institute. SJF is subcontracted to work with Medable to analyze the study data.

Ethical approval: The KPNC Institutional Review Board approved this study. We obtained informed consent from participants prior to any data collection.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article. This work was supported by the HHS NIH (grant number HHSN261201800010C).

Guarantor: IOG.

ORCID iDs: Ingrid Oakley-Girvan https://orcid.org/0000-0003-0328-921X
Raymond Liu https://orcid.org/0000-0002-4667-9632
Sara Aghaee https://orcid.org/0000-0003-4912-5069

References

1. de Moor JS, Mariotto AB, Parry C, et al. Cancer survivors in the United States: prevalence across the survivorship trajectory and implications for care. *Cancer Epidemiol Biomarkers Prev* 2013; 22(4): 561–570.
2. American Cancer Society. *Cancer facts & figures, 2022*. Atlanta, Georgia: American Cancer Society, 2022.
3. Chaudhari N, Ravi R, Gogtay NJ, et al. Recruitment and retention of the participants in clinical trials: challenges and solutions. *Perspect Clin Res* 2020; 11(2): 64–69.
4. Unger JM, Vaidya R, Hershman DL, et al. Systematic review and meta-analysis of the magnitude of structural, clinical, and physician and patient barriers to cancer clinical trial participation. *J Natl Cancer Inst* 2019; 111(3): 245–255.
5. Feyman Y, Provenzano F and David FS. Disparities in clinical trial access across US urban areas. *JAMA Netw Open* 2020; 3(2): e200172.
6. Waterhouse DM, Harvey RD, Hurley P, et al. Early impact of COVID-19 on the conduct of oncology clinical trials and long-term opportunities for transformation: findings from an American Society of Clinical Oncology Survey. *JCO Oncol Pract* 2020; 16(7): 417–421.
7. Khozin S and Coravos A. Decentralized trials in the age of real-world evidence and inclusivity in clinical investigations. *Clin Pharmacol Ther* 2019; 106(1): 25–27.
8. Dorsey ER, Kluger B and Lipset CH. The new normal in clinical trials: decentralized studies. *Ann Neurol* 2020; 88(5): 863–866.
9. Pew Research Center. Mobile fact sheet. In. Demographics of mobile device ownership and adoption in the United States. Vol 9/8/2021: pewresearch.org; 2021.
10. U.S. Food and Drug Administration. Patient-reported outcome measures: use in medical product development to support labeling claims. Guidance for Industry. December, 2009. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims (accessed 3 May 2023.
11. Gotay CC, Kawamoto CT, Bottomley A, et al. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol* 2008; 26(8): 1355–1363.
12. Mierzynska J, Piccinin C, Pe M, et al. Prognostic value of patient-reported outcomes from international randomised clinical trials on cancer: a systematic review. *Lancet Oncol* 2019; 20(12): e685–e698.
13. National Cancer Institute. Severe side effects of cancer treatment are more common in women than men. News and Events. March 15, 2022. Available at: https://www.cancer.gov/news-events/cancer-currents-blog/2022/cancer-treatment-women-severe-side-effects (accessed 3 May 3 2023).
14. Oakley-Girvan I, Davis SW, Kurian A, et al. Development of a mobile health app (TOGETHERCare) to reduce cancer care partner burden: product design study. *JMIR Form Res* 2021; 5(8): e22608.
15. Dumenci L, Matsuyama R, Riddle DL, et al. Measurement of cancer health literacy and identification of patients with limited cancer health literacy. *J Health Commun* 2014; 19(Suppl. 2): 205–224.
16. Basch E, Reeve BB, Mitchell SA, et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J Natl Cancer Inst* 2014; 106(9): dju244.
17. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982; 5(6): 649–655.
18. Bevans M, Ross A and Cella D. Patient-reported outcomes measurement information system (PROMIS): efficient, standardized tools to measure self-reported health and quality of life. *Nurs Outlook* 2014; 62(5): 339–345.
19. Kroenke K, Spitzer RL, Williams JB, et al. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics* 2009; 50(6): 613–621.
20. Apple Research Kit. Active Tasks. Research Kit. http://researchkit.org/docs/docs/ActiveTasks/ActiveTasks.html. Published 2018 (accessed 8 October 2022).
21. Centers for Disease Control and Prevention. Stopping elderly accidents, deaths and injuries (STEADI). https://www.cdc.gov/steadi/. Published 2021 (accessed 26 July 2022).
22. Bangor A, Kortum PT and Miller JT. An empirical evaluation of the system usability scale. *Int J Human–Comput Inter* 2008; 24(6): 574–594.
23. Sauro J. Measuring usability with the system usability scale (SUS). In: *Measuring usability with the system usability scale (SUS).* Vol 2022. Measuring U.2011.
24. Reichheld FF. The one number you need to grow. *Harv Bus Rev* 2003; 81(12): 46–55.
25. Stoyanov SR, Hides L, Kavanagh DJ, et al. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015; 3(1): e27.
26. Usability.gov (archived). System Usability Scale (SUS). Available at: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html (accessed 3 May 2023).
27. Lewis JR and Sauro J. The factor structure of the system usability scale. Paper presented at: International conference on human centered design 2009.
28. Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3(2): 77–101.
29. Pratap A, Neto EC, Snyder P, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit Med* 2020; 3: 21.

30. Sauro J. 5 ways to interpret a SUS score. In: 5 ways to interpret a SUS score. Vol. 2022. Measuring U.2018.
31. Smyk A. The system usability scale and how it's used in UX. In. Vol 2022. XD.Adobe.2020.
32. Maramba I, Chatterjee A and Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. *Int J Med Inform* 2019; 126: 95–104.
33. Min YH, Lee JW, Shin YW, et al. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. *J Med Internet Res* 2014; 16(5): e135.
34. Teckie S, Solomon J, Kadapa K, et al. A mobile patient-facing app for tracking patient-reported outcomes in head and neck cancer survivors: single-arm feasibility study. *JMIR Form Res* 2021; 5(3): e24667.
35. Zini EM, Lanzola G, Quaglini S, et al. A pilot study of a smartphone-based monitoring intervention on head and neck cancer patients undergoing concurrent chemo-radiotherapy. *Int J Med Inform* 2019; 129: 404–412.
36. Benze G, Nauck F, Alt-Epping B, et al. PROutine: a feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. *Ann Palliat Med* 2019; 8(2): 104–111.
37. Trojan A, Huber U, Brauchbar M, et al. Consilium smartphone app for real-world electronically captured patient-reported outcome monitoring in cancer patients undergoing anti-PD-L1-directed treatment. *Case Rep Oncol* 2020; 13(2): 491–496.
38. Kneuertz PJ, Jagadesh N, Perkins A, et al. Improving patient engagement, adherence, and satisfaction in lung cancer surgery with implementation of a mobile device platform for patient reported outcomes. *J Thorac Dis* 2020; 12(11): 6883–6891.
39. Mowlem FD, Tenaerts P, Gwaltney C, et al. Regulatory acceptance of patient-reported outcome (PRO) data from bring-your-own-device (BYOD) solutions to support medical product labeling claims: let's share the success stories to move the industry forward. *Ther Innov Regul Sci* 2022; 56(4): 531–535.