Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

OXFORD
UNIVERSITY PRESS

Mattias Georgsson[1,2], Nancy Staggers[1,3]

ABSTRACT

Objective Mobile health (mHealth) systems are becoming more common for chronic disease management, but usability studies are still needed on patients' perspectives and mHealth interaction performance. This deficiency is addressed by our quantitative usability study of a mHealth diabetes system evaluating patients' task performance, satisfaction, and the relationship of these measures to user characteristics.

Materials and Methods We used metrics in the International Organization for Standardization (ISO) 9241-11 standard. After standardized training, 10 patients performed representative tasks and were assessed on individual task success, errors, efficiency (time on task), satisfaction (System Usability Scale [SUS]) and user characteristics.

Results Tasks of exporting and correcting values proved the most difficult, had the most errors, the lowest task success rates, and consumed the longest times on task. The average SUS satisfaction score was 80.5, indicating good but not excellent system usability. Data trends showed males were more successful in task completion, and younger participants had higher performance scores. Educational level did not influence performance, but a more recent diabetes diagnosis did. Patients with more experience in information technology (IT) also had higher performance rates.

Discussion Difficult task performance indicated areas for redesign. Our methods can assist others in identifying areas in need of improvement. Data about user background and IT skills also showed how user characteristics influence performance and can provide future considerations for targeted mHealth designs.

Conclusion Using the ISO 9241-11 usability standard, the SUS instrument for satisfaction and measuring user characteristics provided objective measures of patients' experienced usability. These could serve as an exemplar for standardized, quantitative methods for usability studies on mHealth systems.

RESEARCH AND APPLICATIONS

BACKGROUND AND SIGNIFICANCE

To assist patients with diabetes, different types of support systems for self-management have been developed using Information and Communication Technology (ICT) including mobile health (mHealth) technologies.[1,2] Studies show that mHealth in particular has helped improve chronic disease[3] and glucose management for these patients,[4,5] but the usability of mHealth interventions still needs to be addressed more fully.[6–8]

Ensuring adequate usability is of the essence for the individual patient and because of the worldwide penetration of mobile phones.[8] Mobile phone availability is approaching the 7 billion mark globally, with a recent figure showing an estimated 6.9 billion subscriptions.[9] An estimated 335 million wireless subscriber connections (mobile phones, tablets, etc.) are in the United States alone.[10] Given the volume of mHealth applications and insufficient usability assessments, the potential impacts on user interactions are not clear but are likely substantial. Attention is needed to performing standardized, systematic mHealth usability assessments.[11]

International standards and validated instruments are available to aid in the design of mHealth systems, although these are not yet widely used in health care. For example, an October 2014 review of

usability testing studies on mHealth technology for diabetes showed great variations in usability testing methods. Of the list of 23 applicable studies, only four used a validated instrument.[11] Only one study followed a completely standardized procedure including the use of validated instruments.[11]

Authors recommend a more widespread use of International Organization for Standardization (ISO) guidelines and techniques[12] to guide usability evaluations.[13] Expanding usability testing to be more systematic and complete can help build a science of mHealth usability. More standardized, comprehensive approaches would improve methodological consistency, making it possible to begin comparing findings across mHealth application evaluations.[11,12] Researchers need to assess the full set of recommended measures—effectiveness, efficiency, and satisfaction—to obtain a more thorough picture of the usability of any application.[11,14] In that vein, this paper employs ISO 9241-11 and techniques to address this gap for mHealth systems.

Studies also need to be conducted on the effects of relevant user demographics because they can affect requirements including the design of future decision aids for chronic disease.[15,16] Specific user characteristics can produce differences in task performance, but how and in what way is not completely understood for mHealth systems.[17]

Correspondence to Mattias Georgsson, MSc, RN, Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden; mattias.georgsson@bth.se; Tel: +46 455385000. For numbered affiliations see end of article.

In this study, our goal was to complete an evidenced-based approach to assessing a mHealth system. To address noted gaps in current mHealth usability testing, we employed a standardized method assessing the full set of major aspects of usability recommended by ISO 9241-11.[18] We used instruments with established psychometrics such as the System Usability Scale (SUS) instrument, and we measured relevant user characteristics associated with mHealth task performance to obtain a broader understanding of users, their tasks, and performance outcomes in mHealth interactions.

Overview of Usability Testing

Usability is defined as "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."[18] Usability testing is commonly used in the usability engineering and human-computer interaction disciplines, and more recently, in the health care arena.[19,20] The purpose of usability testing is to determine obstacles to effective and efficient product use as well as product acceptability and satisfaction for representative users as they interact with the intervention in a specific context.[18,19,21]

Usability can be measured by several different methods. One of the most common is heuristic evaluation, an expert-based approach originally introduced by Nielsen.[22,23] In contrast, user testing typically employs actual users such as health consumers.[22,24] A particular advantage to usability testing is that resource requirements can be fairly modest. The number of users needed to perform user tests varies, but a small sample of five to eight users can commonly identify 80–85% of usability problems.[25,26] User test sessions include representative tasks and scenario-based evaluation[22] in an environment suitable for conducting the testing. This can be either a laboratory environment under controlled conditions or in the field, either of which allows researchers to obtain an understanding about how the intervention is used in a specific context.[27] The observation and task/scenario-based process involves observing and assessing how users interact with the intervention and how well the interaction supports users in achieving their specific task goal(s).[22,28]

Measuring Usability

ISO 9241-11

One of the quantitative methods to determine usability is outlined in ISO 9241-11, a standard consisting of specific metrics about how well a user fulfills specific goals (see Figure 1).[18] This standard includes the main concepts of user-centered design (UCD), which entails incorporating users early and often throughout all steps of the design and development process.[29]
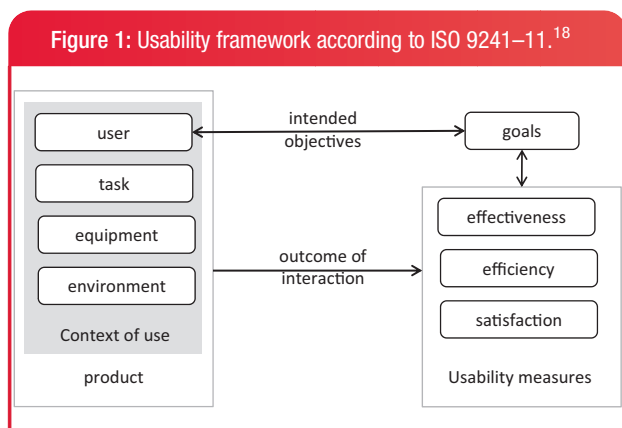
ISO 9241-11 describes in depth how users should interact with a product, employing hands-on methods to indicate its overall usability.[30,31] One common technique is to record users as they perform representative tasks during the interactions.[18] If the indicated measures of effectiveness, efficiency, and satisfaction are fulfilled adequately, the product can be considered to have attained an acceptable level of usability.[18] Appropriate terms are defined in Table 1.

Standards such as the ISO 9241-11 are especially suitable to apply to new technologies such as mHealth systems and applications. In support of this notion, Bevan concludes that these standards should be used more frequently in usability work as they define good practice, are objective, can ensure consistency in the work, and can provide benchmarks for intervention by designers.[12]

Measuring Effectiveness and Efficiency

Effectiveness is typically measured by task completion success and by counting the number of errors performed during the interaction. Efficiency includes the level of effort and resource use by the user to achieve usability goal(s) and is typically measured by timing each task and averaging times across users and/or tasks.[18]

Measuring Satisfaction

Satisfaction can be objectively measured with available instruments, including the SUS instrument used extensively outside health care. Developed and designed by Brooke, the SUS measures overall usability, allowing comparisons across a range of contexts and systems.[32] This 10-item Likert scale instrument is typically administered immediately after interaction, allowing users to record their initial feelings and responses.[32] Instrument items have a range of 0–4. The SUS has been evaluated for validity, reliability, and sensitivity.[32–35] SUS scores range from 0 to 100 providing an estimate of overall usability of the intervention in the minds of users.[32,35] Scores of above 70 are considered to be acceptable or good while scores of 85 or above indicate a high level of usability or excellent score. Scores of 50 or below indicate poor or unacceptable usability.[36]

MATERIALS AND METHODS

The measures for this work were determined by employing a standardized usability assessment along with patient characteristics. The evaluation process is described below.

Study Sample and Setting

The study received approval (HI-BEA-001#20111605) from the Western Institutional Review Board, Olympia, Washington, USA. Study participants were selected from a database of 2317 patients involved in a larger randomized controlled trial on a diabetes mHealth intervention. The larger study used a convenience sampling technique and

Figure 1: Usability framework according to ISO 9241–11.[18]

Table 1: Included attributes for usability according to ISO 9241–11.[18]

Effectiveness: To what extent the user can achieve a goal with accuracy and completeness

Efficiency: The level of effort and resource usage which is required by the user in order to achieve a goal in relation to accuracy and completeness

Satisfaction: The positive associations and absence of discontent that the user experiences during the performance

involved patients from 18 primary care clinics in the Salt Lake City metropolitan area, Utah.[37] Inclusion criteria for the larger study were (1) adults of 18 years old or older with (2) a type 2 diabetes diagnosis and glycated hemoglobin (HbA1c) ≥8%, (3) not pregnant, (4) with cell phone access, and (5) able to speak and understand the English language. Ten patients were randomly selected from the larger study database for inclusion in our usability study. These patients had no prior experience with the mHealth web system evaluated in our current study. Inclusion criteria for our usability study sample were (1) patients diagnosed with type 2 diabetes; (2) no cognitive impairment; (3) some knowledge and use of computers, the Internet, and cell phones; and (4) the ability to speak and understand the English language. After obtaining informed consent, patients were scheduled for individual usability sessions in a quiet room at HealthInsight, one of the US Beacon Communities in Salt Lake City, Utah.

System Description

Care4Life® is an interactive algorithm-based mobile short messaging service management system with an accompanying web portal for patients with diabetes (see Figure 2). The system was designed as a personalized, self-care management tool meant to function as input-driven coaching support for patients' self-care management and to aid in the collaboration between patients and their health care providers. Patients interact with the system by being prompted to send in their self-management measures via text messaging. Typical transmitted measures are blood glucose, blood pressure, weight, exercise, and medication adherence.

User Assessment Tasks

Representative user tasks consisted of (1) sending in measurement values to the system, (2) interpreting glucose measurements displayed in a graph, (3) recording glucose measurement values, (4) exporting glucose measurement value trends to a portable document format, (5) interpreting blood pressure measurement in a graph view, (6) setting and tracking goals regarding exercise and weight, (7) setting medication reminders, and (8) setting a physician appointment reminder. It was vital to include a variety of relevant tasks of different lengths and levels of difficulty to obtain representative and accurate performance. Tasks were based on real case scenarios to simulate how patients would interact with the system in a real-life situation according to the care and self-management process and were validated by a panel of three health care professionals and usability experts for content and context accuracy.

Figure 2: System overview.

Instruments and Outcome Measures

Demographics and IT/Computer Knowledge and Experience Questionnaires

Two short pretest questionnaires were administered to assess patients' background characteristics and information technology (IT) experience. These were divided to make them easy for patients to complete. Validation was performed by three health care professionals and usability experts. Patients were asked about their age, gender, educational background, how long they had had diabetes, their experiences and knowledge about computers/IT including self-assessed IT knowledge, and if and how often they used a computer as well as the Internet.

ISO Standard Usability Measures

ISO measures of effectiveness and efficiency were measured as follows: Effectiveness was determined by degree of task completion and the total number of errors per task. Task completion was coded using three categories: (1) completed with ease when the user was able to perform the task without any help from the test leader, (2) completed with difficulty when the subject achieved the task with minor difficulties and or with minor hints from the test leader, and (3) failed to complete when the subject was unable to complete the task, even with some minor hints. An error was coded when the subject performed errors they could not solve or committed errors preventing further progress.

Efficiency was determined by timing each individual task and computing the average time for each task across patients. Efficiency or time on task began when the user started performing the task and ended when they pressed the participant/home button. Time was deducted for prolonged loading and response time for the system.

Satisfaction was measured by administering the SUS. Scores were calculated according to Brooke's guidelines.[32] This consisted of summing the scores on each of the 10 individual items. For items 1, 3, 5, 7, and 9, one point was subtracted from the resulting score. For items 2, 4, 6, 8, and 10, five points were subtracted from the resulting score. The final sum of all scores was then multiplied by 2.5 to get the overall satisfaction value.

Study Procedure and Data Collection

After completing the informed consent, patients filled out the pretest surveys on demographics. Standardized training is recommended to decrease individual variability in task performance due to knowledge about a system.[38] Thus, standardized training was done to simulate an actual patient educational care process in the health clinic, and users interacted with the system on their own. Researchers explained the proposed evaluation process. Patients then interacted with the system using the described tasks to enter and retrieve values as well as read different graphs. Morae®[39] software was used to video- and audio-record patients interactions in the system. Subsequently, the recordings were carefully coded using the specific metrics as defined above on task success rate, errors, and time on task. The session ended with the administration of the SUS. The testing procedure took ~1.5 h per patient, and patients received a gift card for $20 after finishing the session.

Data Analysis

Results were calculated using a Microsoft® Excel®[40] spreadsheet. Descriptive statistics such as means and standard deviations were calculated in SPSS® version 22[41] on effectiveness, efficiency, and satisfaction results. Aggregated data were compiled for the specific tasks across patients. Descriptive statistics results were compared to the
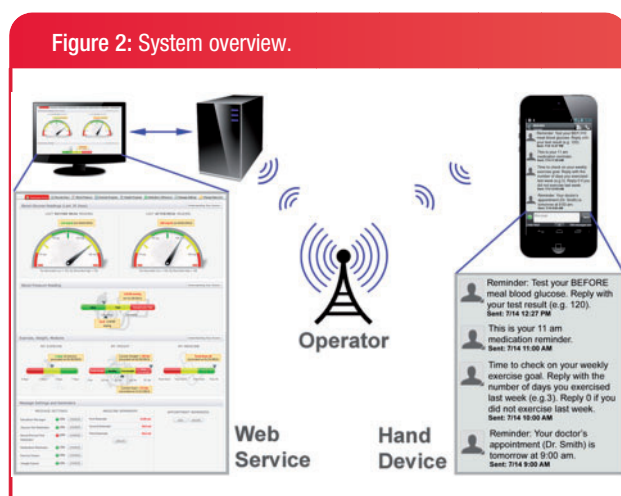
sociodemographic data to distinguish differences and similarities between users and their performance data. To make comparisons between user characteristics and user performance and satisfaction, we compared gender, age, education, diabetes diagnosis, and IT/computer knowledge and experience (based on rated experience vs rated inexperience) against effectiveness, efficiency, and SUS mean scores. Statistical correlations were not performed due to the small sample size.

RESULTS

Patient Demographics, IT/Computer Knowledge, and Use (Pretest Questionnaire)

The 10 patients had a variety of different backgrounds and characteristics. Six participants were female and four male. Most (70%) were older adults (50–69 years), while 30% were between the ages of 30–49. Most patients had college or university education (80%). Sixty percent were diagnosed with type 2 diabetes $\geq$5 years ago while 40% were diagnosed $\leq$4 years ago.

IT and computer knowledge use and preferences differed among patients. Half rated their IT knowledge level at medium and 30% at high while the remainder indicated little knowledge. A majority (80%) of patients reported using a computer and the Internet every day. Eighty percent also agreed or strongly agreed that they enjoyed using the computer in their work or leisure time.

Effectiveness

The task effectiveness results are presented in Figure 3. Tasks 3 (correcting a glucose measurement value) and 4 (exporting a glucose value) were the most difficult to complete with 30% and 40% failure rates, respectively. Tasks 1, 2, and 5 (send in measurement value, interpret glucose measurement value, and interpret blood pressure measurement value) were completed with ease by all.

The error rates mirrored task completion difficulty with Tasks 3 and 4 having the largest number of errors while Tasks 1, 2, and 5 were completed without any errors (see Figure 3). The largest number of errors was quite high at 9 and 13. The kinds of errors committed

included patients having difficulties in remembering steps, finding correct options, and performing the different steps in the process. Errors were especially prevalent with tasks related to (1) locating glucose values for correction, (2) exporting these data, (3) remembering how to delete values or select values, (4) navigating to the correct screens to accomplish these tasks, and finally (5) remembering how and where to add or export their glucose values.

Efficiency

As may be seen in Table 2, Tasks 3 and 4 consumed the longest amount of time, as might be expected given the difficulties with task success and errors mentioned above. On the other hand, Tasks 2 and 5 took the shortest times. Tasks of correcting and exporting values (3 and 4) had mean scores 2–3 times as long as those related to interpreting values.

Satisfaction

The average SUS score for the whole group was 80.5 (SD 11.47) indicating good satisfaction across these mHealth system users as seen in Figure 4. However, wide variations in scores existed with a low value of 62.5 and high score of 97.5 (a 35-point range). The highest ratings ranged from 87.5 to 97.5 or excellent (30% of the patient sample) to the lowest from 62.5 to 72.5 or "OK" to minimally acceptable (30% of the patient sample).

User Characteristics, IT Knowledge, and Internet Skills and Usability Metrics

User characteristics and objective data were assessed for additional insights (Table 3). Descriptive trends indicate a difference across genders, ages, diabetes diagnosis, and IT/computer knowledge and experience. Males in this sample had higher average success, lower error rates, and higher mean SUS scores than females. Younger patients also had higher average task completion rates on tasks, only one error average on tasks, lower task completion times, and higher mean satisfaction scores.

Patients with a more recent diabetes diagnosis had a higher average task success rate, fewer errors on tasks, shorter average time on tasks, and higher average SUS scores. High school or college/

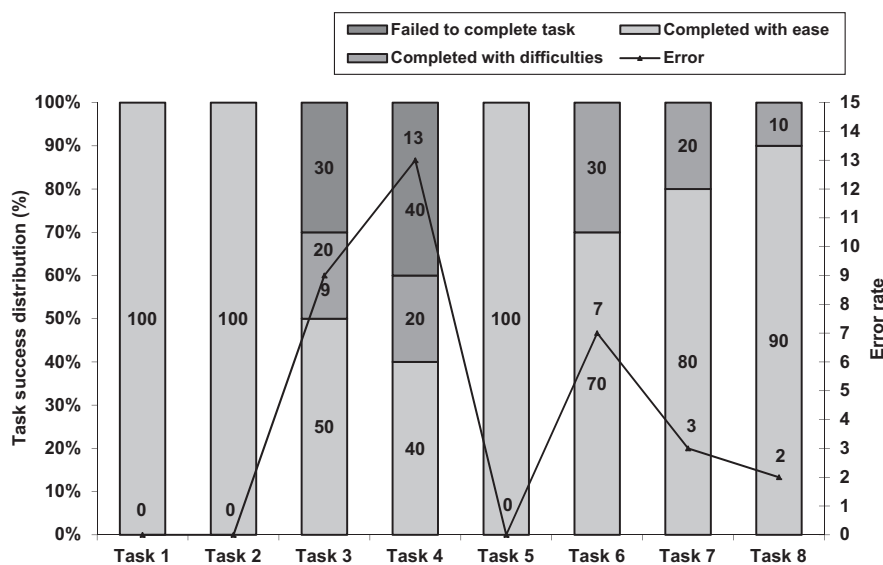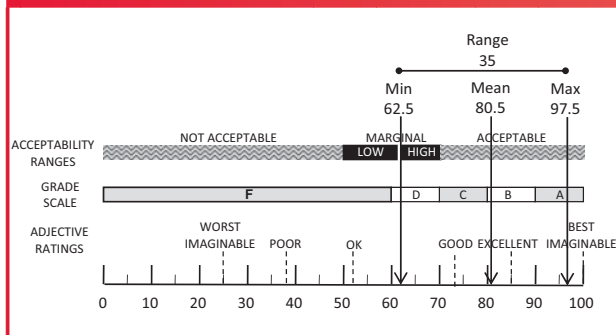**Figure 3:** Task success and error rate by task.

Table 2: Time on task per task									
		Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8
Time per task (min)	Mean (SD)	1.51 (0.48)	1.45 (0.45)	4.18 (1.55)	3.67 (1.71)	1.33 (0.37)	2.21 (0.72)	1.75 (0.56)	1.69 (0.57)
	Range	0.97–2.41	0.91–2.50	1.66–6.47	1.35–6.86	0.80–2.01	1.33–3.27	0.91–2.72	0.76–2.61

Figure 4: Overview of modified SUS rating table with inserted value ranges.[36]

university education did not have an influence on performance or satisfaction scores. More experienced IT/computer users performed tasks more quickly and with fewer errors than those less experienced. They also had higher SUS scores compared to the less experienced users.

DISCUSSION

This study demonstrates the application of systematic usability methods and how researchers may take into account relevant patient characteristics during mHealth system interactions. These considerations are consistent with the predicted growth in mHealth usage rates.[8] Our results show how the ISO usability standard may be employed to assess the set of effectiveness, efficiency, and satisfaction measures. The comprehensive set of variables allow increased understanding of system usability by users, their tasks, and their performance interaction requirements for mHealth systems. Moreover, this study shows how patients' characteristics can influence interaction performance and how developers might create more usable eHealth and mHealth systems.

Interpretation of Task Performance Results, Satisfaction, and Demographic Trends

Possible reasons that Tasks 3 and 4 (updating and exporting a glucose value) proved difficult might be related to the tasks, to user practice, and also to needed redesign. These two tasks were more complex, requiring several steps, while tasks 1 and 5 were more straightforward, with only one step each. Another possible reason is that this usability test assessed users while they were still practicing these new, more complex tasks.

The overall satisfaction results indicated good, although not excellent, usability. In fact, nearly one-third of the participants gave the system a rather poor usability rating ("OK"). The results indicate areas for system improvements. For example, users were confused about the multiple steps and current, nonintuitive steps in Tasks 3 and 4. Although this task ought to be straight-forward, the current design is not, requiring users to navigate to different areas of the system to delete and enter new values. Clearly, the multiple steps can be

streamlined, so this usability test presented objective data about these current, cumbersome tasks.

In this sample, demographic and performance trends indicated that males and those with more IT experience performed slightly better and had higher SUS scores. Younger patients also performed faster and had fewer errors. These trends point to problematic areas, especially for users with specific characteristics. Therefore, designers may also need to tailor interactions for targeted users; for example, older, less experienced users. This aspect of the study is congruent with other literature indicating mHealth interventions need to be better adapted to a wide variety of users to facilitate wider usage for a larger number of users.[42]

Results across the examined performance and satisfaction variables were congruent, although this is not always the case in usability studies. That is, participants who completed tasks more successfully were also faster and committed fewer errors. Those who performed better had higher satisfaction scores.

Contributions to the Literature

To our knowledge, this is the first usability study on mHealth diabetes interventions assessing usability effectiveness, efficiency, and satisfaction together by using validated measures. In addition, we used the ISO usability standard and SUS instrument to compare usability metric performance outcomes to pertinent patient user characteristics.

This study helps fill several gaps in the literature. Previous authors recommended more usability studies focused on patient engagement and product interaction[6–8,43] as opposed to exclusively medical outcomes.[4,5,44] Other authors have reported the need to explore the influence of demographic characteristics and technology on usability assessment scores to assist in designing future interventions targeted to specific populations.[15–17] Finally, previous authors argued about the importance of assessing effectiveness, efficiency, and satisfaction measures together to obtain a fuller picture of the usability as was done in the current study.[11,14] This study fulfilled these recommendations.

This study adds to the literature on mHealth interventions as an example of an evidence-based approach resulting in a more inclusive quantitative usability assessment. We employed actual users during testing, adhering to recommendations from the ISO usability standard. The results provide developers with specific data about where patients experience the most difficulty in the product and where tasks should be redesigned to accommodate a wider variety of users. This study could serve as an exemplar to others evaluating mHealth systems who want to generate benchmark data and provide reproducible comparable results.

Usability testing results do not have to show overall poor usability to be helpful. Moreover, we need positive examples of usability testing with mHealth systems. In the past, authors focused mainly on negative examples. This study uses a standard methodology to explicate the core usability issues that need to be addressed and provides specific techniques others may employ to evaluate and improve the usability of mHealth systems.

RESEARCH AND APPLICATIONS

RESEARCH AND APPLICATIONS

User characteristics	Success rate, mean (SD)	Error rate, mean (SD)	Time on task, mean (SD)	Satisfaction, mean (SD)
Gender				
Male	93.75 (12.50)	2.25 (2.63)	1.68 (0.39)	83.13 (2.39)
Female	89.58 (12.29)	4.17 (2.71)	2.58 (0.41)	79.69 (14.98)
Age (years)				
30–49	100 (0.00)	1.00 (1.00)	1.95 (0.27)	88.33 (8.04)
50–69	87.50 (12.50)	4.43 (2.57)	2.34 (0.68)	77.14 (11.50)
Diabetes (years)				
0–4	100 (0.00)	1.50 (1.14)	2.05 (0.30)	88.13 (6.57)
≥5	85.42 (12.29)	4.67 (2.73)	2.34 (0.74)	75.42 (11.56)
Education				
High school	93.75 (8.84)	3.50 (0.71)	2.39 (0.23)	81.25 (12.37)
College/university	90.63 (12.94)	3.38 (3.07)	2.18 (0.67)	80.31 (12.13)
IT/Comp. Experience				
Less experienced	87.50 (17.68)	6.50 (3.54)	2.71 (0.22)	76.25 (19.44)
More experienced	92.19 (11.45)	2.63 (2.07)	2.10 (0.61)	81.56 (10.43)
Sample ($n = 10$)	91.25 (11.86)	3.40 (2.72)	2.22 (0.60)	80.50 (11.47)

Table 3: Comparison between user characteristics and usability metrics

LIMITATIONS

Our study involved a smaller user sample from a specific diabetes population, which may make it difficult to generalize findings to the general population. Although these users were randomly selected from a larger database of a mHealth diabetes intervention study, the larger study used a convenience sampling technique. The sample size of 10 subjects is appropriate for usability testing and is greater than sample sizes recommended by Nielsen and Landauer[25], and Virzi.[26] However, these users may not be representative of all mHealth users of diabetes systems. In particular, the participants in this randomized sample were more highly educated than most samples of patients with diabetes, especially those in urban areas. Future research in the area could explore other mHealth technologies as well as repeating this study on a greater scale with larger, randomly selected user samples to determine performance and satisfaction outcomes. Other user characteristics could be compared to these performance measures. In summary, this system focused on diabetes specifically, but the findings may also have broader applications for product designs of other chronic disease applications and concomitant usability testing.

CONCLUSION

The purpose of this study is to act as an exemplar for systematic usability methods. While results showed good (not excellent) perceived usability satisfaction, nearly one-third of users rated the system as having rather poor usability. The results do show objective data for developers and point to needed corrections, especially for confusing tasks. The variations across users and tasks are notable. Objective data are important because designer and researcher perceptions about tasks and user performance may not be congruent with performance data. In addition, system change lists can then be data-based to determine priority corrections.

This study demonstrates the use of a thorough quantitative approach by taking into account varied needs of users who interact with mHealth systems for disease management. It also shows the usefulness of performing several different kinds of assessment measures.

The study used methods recommended in the ISO standard to assess effectiveness, efficiency and satisfaction with validated tools such as the SUS instrument. Together, these methods provide an increased understanding of system usability and could serve as an exemplar for methodological approaches by designers and researchers.

The present study addressed a gap in the literature by examining patients with varied characteristics and skill levels on a suite of performance outcomes. Trends such as those seen herein can help developers interpret user needs in designing more usable mHealth systems. Trends in participant socio-demographic and IT knowledge data also indicated that gender, age, disease duration, and IT experience level may influence interaction outcomes. To increase the scale of mHealth use to promote wider use and wider acceptance, these kinds of user characteristics will likely need to be considered more thoughtfully in system design in the future.

COMPETING INTERESTS

The authors have no competing interests to declare.

REFERENCES

1. Sutcliffe P, Martin S, Sturt J, et al. Systematic review of communication technologies to promote access and engagement of young people with diabetes into healthcare. *BMC Endocr Disord.* 2011;11:1.
2. Quinn CC, Royak-Schaler R, Lender D, et al. Patient understanding of diabetes self-management: participatory decision-making in diabetes care. *J Diabetes Sci Technol.* 2011;5(3):723–730.
3. Jones KR, Lekhak N, Kaewluang N. Using mobile phones and short message service to deliver self-management interventions for chronic conditions: a meta-review. *Worldviews Evid Based Nurs.* 2014;11(2):81–88.
4. Arsand E, Tatara N, Ostengen G, et al. Mobile phone-based self-management tools for type 2 diabetes: the few touch application. *J Diabetes Sci Technol.* 2010;4(2):328–336.
5. Kouris I, Mougiakakou S, Scarnato L, et al. Mobile phone technologies and advanced data analysis towards the enhancement of diabetes self-management. *Int J Electron Healthc.* 2010;5(4):386–402.
6. Mulvaney SA, Ritterband LM, Bosslet L. Mobile intervention design in diabetes: review and recommendations. *Curr Diab Rep.* 2011;11(6):486–493.
7. Tsai CC, Lee G, Raab F, et al. Usability and feasibility of PmEB: a mobile phone application for monitoring real time caloric balance. *Mobile Netw Appl.* 2007;12(2-3):173–184.
8. El-Gayar O, Timsina P, Nawar N, et al. Mobile applications for diabetes self-management: status and potential. *J Diabetes Sci Technol.* 2013;7(1):247–262.
9. International Telecommunications Union. *The World in 2014: ICT Facts and Figures.* Geneva: International Telecommunications Union (ITU); 2014.
10. CTIA -The Wireless Association. *Wireless Industry Survey 2013.* CTIA: The Wireless Association; 2014.
11. Lyles CR, Sarkar U, Osborn CY. Getting a technology-based diabetes intervention ready for prime time: a review of usability testing studies. *Curr Diab Rep.* 2014;14(10):534.
12. Bevan N. International standards for usability should be more widely used. *J Usability Stud.* 2009;4(3):106–113.
13. Svanaes D, Das A, Alsos OA. The contextual nature of usability and its relevance to medical informatics. *Stud Health Technol Inform.* 2008;136:541–546.
14. Frokjaer E, Hertzum M, Hornbaek K. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? In *Proceedings of the SIGCHI conference on Human factors in computing systems.* The Hague, The Netherlands: ACM; 2000:345–352.
15. Or C, Tao D. Usability study of a computer-based self-management system for older adults with chronic diseases. *JMIR Res Protoc.* 2012;1(2):e13.
16. Li LC, Adam PM, Townsend AF, et al. Usability testing of ANSWER: a web-based methotrexate decision aid for patients with rheumatoid arthritis. *BMC Med Inform Decis Mak.* 2013;13:131.
17. Coursaris CK, Kim DJ. A meta-analytical review of empirical mobile usability studies. *J Usability Stud.* 2011;6(3):117–171.
18. International Organization for Standardization. Ergonomic requirements for office work with visual display terminals (VDTs) Part 11 Guidance on usability. International standard, 9241-11. Geneve: ISO; 1998.
19. Kushniruk AW, Borycki EM, Kuwata S, et al. Emerging approaches to usability evaluation of health information systems: towards in-situ analysis of complex healthcare systems and environments. *Stud Health Technol Inform.* 2011;169:915–919.
20. Liljegren E. Usability in a medical technology context assessment of methods for usability evaluation of medical equipment. *Intern J Indust Ergonom.* 2006;36(4):345–352.
21. Yardley L, Morrison LG, Andreou P, et al. Understanding reactions to an internet-delivered health-care intervention: accommodating user preferences for information provision. *BMC Med Inform Decis Mak.* 2010;10:52.
22. Nielsen J. *Usability Engineering.* Boston: Academic Press; 1993:xiv, 358 p.
23. Nielsen J. Heuristic evaluation. In Mack RL, Nielsen J. eds. *Usability Inspection Methods.* New York: John Wiley & Sons inc; 1994: 25–62.
24. Nielsen J. *Why You Only Need to Test with 5 Users.* 2000 March 19 [accessed December 27, 2014]; http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/.
25. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 conference on human factors in computing systems.* ACM: Amsterdam, The Netherlands; 1993: 206–213.
26. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? *Hum Factors.* 1992;34(4):457–468.
27. Maguire M. Methods to support human-centred design. *Intern J Hum Comput Stud.* 2001;55(4):587–634.
28. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. *J Biomed Inform.* 2004;37(1):56–76.
29. International Organization for Standardization. Ergonomics of human-system interaction - Part 210: human-centred design for interactive systems. *International standard, 9241-210.* Geneve: ISO; 2010.
30. Jaja C, Pares-Avila J, Wolpin S, et al. Usability evaluation of the interactive Personal Patient Profile-Prostate decision support system with African American men. *J Natl Med Assoc.* 2010;102(4):290–297.
31. Bevan N. International Standards for HCl. In Claude G, ed. *Encyclopedia of Human Computer Interaction.* Hershey, PA, USA: IGI Global; 2006:362–372.
32. Brooke J. SUS: a "quick and dirty" usability scale. In Jordan PW, et al., eds. *Usability Evaluation in Industry.* London; Bristol: Taylor & Francis; 1996:189–194.
33. Lewis JR, Sauro J. The Factor Structure of the System Usability Scale. In *Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009.* San Diego, CA: Springer-Verlag; 2009: 94–103.
34. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Intern J Hum Comput Interact.* 2008;24(6):574–594.
35. Brooke J. SUS: a retrospective. *J Usability Stud.* 2013;8(2):29–40.
36. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud.* 2009;4(3):114–123.
37. Capozza K, Woolsey S, Georgsson M, et al. Going mobile with diabetes support: a randomized study of a text message-based personalized behavioral intervention for type 2 diabetes self-care. *Diabetes Spectr.* 2015;28(2):83–91.
38. Karsh BT, Weinger MB, Abbott PA, et al. Health information technology: fallacies and sober realities. *JAMIA.* 2010;17(6):617–623.
39. TechSmith Corporation. *Morae Usability Testing and Market Research Software.* Okemos, MI, USA: TechSmith Corporation; 2014.
40. Microsoft Corporation. *Microsoft Excel 2010 for Windows.* Redmond, WA, USA: Microsoft Corporation; 2010.
41. IBM Corporation. *IBM SPSS Statistics for Windows.* Armonk, NY, USA: IBM Corporation; 2013.
42. Czaja SJ, Sharit J, Lee CC, et al. Factors influencing use of an e-health website in a community sample of older adults. *JAMIA.* 2013;20(2):277–284.
43. El-Gayar O, Timsina P, Nawar N, et al. A systematic review of IT for diabetes self-management: are we there yet? *Int J Med Inform.* 2013;82(8):637–652.
44. Quinn CC, Shardell MD, Terrin ML, et al. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care.* 2011;34(9):1934–1942.

AUTHOR AFFILIATIONS

[1]Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah, USA

[2]Faculty of Computing, Blekinge Institute of Technology, Karlskrona, Sweden

[3]School of Nursing, University of Maryland, Baltimore, Maryland, USA

RESEARCH AND APPLICATIONS