A Systematic Evaluation of Mobile Applications for Diabetes Management

C. Martin, D. Flood, D. Sutton, A. Aldea, R. Harrison, M. Waite

Oxford Brookes University, UK
{cemartin, derek.flood, dsutton, aaldea, rachel.harrison, mwaite}@brookes.ac.uk

Abstract. This short paper contains a summary of work that is currently in progress towards the development of an intelligent, personalised tool for diabetes management. A preliminary part of the development process has consisted of a systematic evaluation of existing applications for mobile phones.

Keywords: efficiency, effectiveness, keystroke level modelling, heuristics.

1 Introduction

The combined technologies of mobile phones and cloud computing have huge potential to improve health care services [1], but a crucial factor in the adoption of such services is the usability of the software [2, 6]. Previous studies have been conducted to evaluate the effectiveness of mobile phones for assisting with diabetes self-management, eg [12], and standards are beginning to emerge with regard to the design of mobile applications in general [14] as well as mobile medical applications in particular (eg ISO/IEC 62366), but research into the usability evaluation of applications for mobile devices is still a relatively new area [7]. Some issues that affect the usability of mobile applications include mobile context, connectivity, screen size and resolution, limited processing power and data entry methods [15], and these issues remain broadly the same despite recent advances in mobile phone technology. The purpose of this study is to uncover usability issues in order to know what pitfalls to avoid in the design of an application for diabetes management, and also to determine what functionality is typically included in such applications.

Type 1 diabetes (T1DM) occurs when the insulin producing cells of the pancreas are destroyed leaving the body unable to control its blood glucose levels. People with T1DM have to take insulin regularly to try to keep their glucose levels within a safe range. The vast majority of patients with T1DM in the UK administer their insulin through multiple daily injections, and the remaining proportion use insulin pumps. Modern insulin pumps include dose calculators to help patients determine how much insulin to administer, but people on multiple daily injections do not usually have this support, and tend to do the calculations themselves. The resulting need for electronic decision support, combined with the recent growth in smart phone use has led to the development of a plethora of diabetes management applications, despite the many barriers that have inhibited the adoption of electronic systems for diabetes management in the past [3,8]. One of the key barriers has been usability, and another

has been the economic implications for healthcare providers, but the standardisation of controls and rigid human interface guidelines imposed by the providers of most phone platforms has already had a huge impact on usability, and the ease with which software applications for mobile phones can now be developed and distributed helps to mitigate against the economic factors. The usability issues listed in [15], are all applicable to systems for diabetes management, but in addition, common complications of diabetes, such as retinopathy and peripheral neuropathy, could affect a patient's ability to view data or use a touch screen respectively.

2　The Evaluation

The protocol used here to analyse the applications is fairly general and has also been applied to mobile spreadsheets [9]. The results given below are limited to the iOS platform, but work is in progress to extend the evaluation to other platforms. The number of spreadsheet applications available for iOS is only a fraction of those for diabetes management, but some similar issues emerged with data entry and visualisation for example. The following steps were performed:

1. *Identify all potentially relevant applications:* this involved a search on the iPhone app store using the keyword *diabetes* which returned a total of 231 applications.

2. *Remove light or old versions of each application.* There were 9 light versions of applications which were removed from the list, leaving 222 apps.

3. *Identify the primary operating functions and exclude all applications that do not offer this functionality.* The key functionality was mapped to corresponding tasks which are shown in Table 1. Only 8 applications had all of the functionality.

   Table 1. Tasks to evaluate primary functionality.

Task	Description	Task	Description
1	Set measurement units	4	Log insulin dose
2	Log blood glucose level	5	Display data graphically
3	Log carbohydrate intake	6	Export data via email or similar

4. *Identify all secondary functionality.* A full list of additional functionality that was offered by the various applications was compiled to be assessed by patients and clinicians in the second phase of this study.

5. *Construct tasks to test the key functionality using each of the methods below:*
    a. *Keystroke level modelling* [4,13] was performed on the final 8 applications using the tasks from Table 1. The number of times the screen was touched in order to complete each task was recorded, where a touch might refer to a single touch or a swipe of the finger. The results are shown in Table 2.

Table 2. Results of KLM.

Application Task:	1	2	3	4	5	6	Total
RapidCalc Insulin Dose Manager	10	2	1	3	2	1	19
GluCoMo	5	5	4	4	1	3	22
Diabetes Diary	5	3	3	2	3	6	22
Diabetes Personal Manager	3	7	4	3	3	6	26
DiabetesPlus	2	7	6	6	1	6	28
LogFrog DB	6	5	6	5	1	7	30
Diabetes Buddy - Control your Blood Sugar	3	9	10	7	1	8	38
Diabetes Pilot	5	5	7	3	3	27	50

This analysis excluded the actual time taken to complete tasks, which will be measured in the study on patients. The KLM process highlighted a number of key differences between the applications that could have a significant effect on usability. For example, the data logging occurs so frequently that it is essential for it to be as efficient as possible, and yet a number of different methods were in use including sliders, pickers and various styles of keyboard. Some applications also made much more effective use of default settings than others.

b. Heuristics [10] were then used to identify more usability problems. The four applications that scored the best in the KLM were subjected to heuristic evaluation by five expert evaluators, using guidelines that take into account specific issues associated with mobile systems [3]. Each heuristic was ranked using Nielsen's five-point Severity Ranking Scale (SRS) [11].

Table 5. Average heuristic ranking.

Heuristic:	A	B	C	D	E	F	G	H	Tot
RapidCalc	1.1	0.9	0.4	0.4	0.8	1.3	1.3	2.0	8.2
GluCoMo	0.6	1.4	1.2	1.7	0.8	1.3	1.3	1.7	10
Diabetes Diary	0.5	0.8	0.9	0.5	0.9	0.7	1.4	0.8	6.5
Diabetes Personal Manager	0.6	0.8	0.9	1.1	1.1	0.9	1.6	2.3	9.3

Each evaluator provided evidence for each ranking, and this exposed some problems that had not been shown by the KLM. For example, there were problems with visibility of battery status, and limited ability to restore data if the device was lost. There were some navigation problems too, but some of these were only learnability issues. Some apps were over-crowded but most offered some degree of personalisation. There were some serious problems with error management and security: it was not easy to edit records on most apps, and data transfer was typically done via unencrypted email, which is not appropriate for healthcare information. Unfortunately it is beyond the scope of this paper to give a thorough analysis of the heuristic evaluation, but that will be the subject of a future paper.

3   Conclusion and Future Work

This short paper has described a systematic evaluation of diabetes management applications for the iOS platform. The KLM evaluation highlighted a number of usability issues associated with data entry and personal settings, and the heuristic evaluation exposed further problems associated with loss of the device, learnability, aesthetics, error management and security. Future work includes a usability study on patients using additional mobile platforms, and the ultimate goal is to design an application for diabetes management with a decision support engine and secure data transfer taking into account the associated legal implications. Further work is also planned on keystroke level modelling of mobile applications, in order to standardise measurements such as what constitutes a single touch or swipe.

4   References

1. Alagöz, F., Calero Valdez, A., Wilkowska, W., Ziefle, M., Dorner, S. & Holzinger, A. (2010). From Cloud Computing to Mobile Internet, From User Focus to Culture and Hedonism: The Crucible of Mobile Health Care and Wellness Applications. IEEE 5th International Conference on Pervasive Computing and Applications (pp.38-45).
2. Bellazzi, R., Telemedicine and Diabetes Management: Current Challenges and Future Resarch Directions. Journal of Diabetes Science and Technology, 2(1), pp.98-104 (2008).
3. Bertini, E., Catarci, T., Dix, A., Gabrielli, S., Kimani, S. & Santucci, G. Appropriating Heuristic Evaluation Methods for Mobile. International Journal of Mobile Human Computer Interaction, Vol. 1, Issue 1, pp. 20-41 (2009).
4. Card, S. K., Thomas, T. P. & Newall, A. The Psychology of Human-Computer Interaction., London: Lawrence Erbaum Associates. (1983)
5. Calero-Valdez, A., Ziefle, M., Firat, A., Holzinger, A. (2010). Mental Models of Menu Structures in Diabetes Assistants. In: Lecture Notes in Computer Science (LNCS 6180), (pp. 584–591). Berlin, Heidelberg, New York: Springer. (ISBN: 978-3-642-14099-0)
6. Greenhalgh, T., Hinder, S. et al. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. BMJ 341, p.c5814. (2010).
7. Kjeldskov, J. & Stage, J New techniques for usability evaluation of mobile systems. International Journal of Human Computer Studies, 60(5-6), pp.599-620. (2004)
8. Klonoff, D. Using Telemedicine to Improve Outcomes in Diabetes - An Emerging Technology. Journal of Diabetes Science and Technology, 3(4), pp.624-628. (2009)
9. Martin, C.; Flood, D; Harrison, R. A Protocol for Evaluating Mobile Applications. To appear in IADIS (2011).
10. Nielsen, J. & Molich, R. Heuristic evaluation of user interfaces. The Conference on Human Factors in Computing Systems: CHI90 (1990).
11. Nielsen, J. Usability Engineering. Morgan Kaufmann Publishers. (1994)
12. Preuveneers, D. and Berbers Y. Mobile phones assisting with health self-care: a diabetes case study. Proceedings MobileHCI '08 Proceedings of the 10th international conference on Human computer interaction with mobile devices and services (2008).
13. Schulz, T. Using the Keystroke-Level Model to Evaluate Mobile Phones. IRIS 31 (2008).
14. W3C. Mobile Web Application Best Practices. http://www.w3.org/TR/mwabp/
15. Zhang, D. & Adipat, B., 2005. Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. International Journal of Human-Computer Interaction, 18(3), pp.293-308.