Questioning the Reflection Paradigm for Diabetes Mobile Apps

Conference or Workshop Item

How to cite:

For guidance on citations see FAQs.

© [not recorded]

Version: Accepted Manuscript

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
Questioning the Reflection Paradigm for Diabetes Mobile Apps

Dmitri Katz, Nick Dalton, Simon Holland, Aisling O’Kane, and Blaine A. Price

dk5697@my.open.ac.uk, nick.dalton@northumbria.ac.uk,
simon.holland@open.ac.uk,a.okane@cs.ucl.ac.uk,
b.a.price@open.ac.uk

Abstract: Hundreds of diabetes self-management apps are available for smartphones, typically using a diary or logging methodology. This paper investigates how well such approaches help participants to make sense of collected data. We found that, while such systems typically support data and trend review, they are ill suited to helping users understand complex correlations in the data. The cognitively demanding user interfaces (UI’s) of these apps are poorly adapted both to the restricted real estate of smartphone displays and to the daily needs of users. Many participants expressed the desire for intelligent, personalized and contextually aware near-term advice. By contrast, users did not see tools for reflection on prior data and behavior, seen as indispensable by many researchers, as a priority. We argue that while designers of future mobile health (mHealth) systems need to take advantage of automation through connected sensors, and the increasing subtlety of intelligent processing, it is also necessary to evolve current graphs and dashboards UI paradigms to assist users in long-term self-management health practices.

Keywords: User interface (UI), diabetes apps, mHealth

1 Introduction

Among major health conditions, diabetes is one of the most common and costly. It is believed to affect 380 million people worldwide, and numbers are rising. Type 1 diabetes (T1D), which affects roughly 5-10% of people with diabetes, is an autoimmune disease that necessitates daily injections of the hormone insulin in order to control blood glucose levels. While short and long-term complications can be severe, diabetes can be successfully managed with careful attention to lifestyle and the correct use of medications. Self-monitoring and self-management practices are essential for good diabetes outcomes, as the majority of care is by necessity self-care [1]. However, glycemic control can be challenging, as it is a multivariate task affected not only by diet, exercise, and insulin dosages, but also hard to control factors such as stress, illness, and natural variability. Diabetes with its strong reliance on diverse data, dynamic treatment, and ability to quantify effects through blood glucose (BG) values, can also be viewed as an edge case that can give insights into the design of similar assistive mHealth technologies.
The paper-based daily diabetes logbook has long been a method of assisting the diabetes management process. The received wisdom is that by patients recording, reviewing, and analyzing tracked factors such as diet, medication dosages, exercise, and location, the patient will be more engaged, form better habits, recognize patterns, and thereby optimize their treatment decisions. This self-management approach has been updated for the ubiquitous smartphone, with hundreds of products allowing not only recording of diverse personal data, but also adding data visualization, data export, social support, food databases, and other functionalities.

However, in an earlier pilot study [2], we found relatively low adoption and retention rates of these diabetes related apps. Many participants were generally positive in assessment, but many viewed the apps as too much work for the benefits delivered. In principle, automation of data streams has the potential to reduce this effort; however, there is little research on the extent to which users are able to extract meaningful insights from this collected data.

Through analysis of user interaction sessions, we present findings that contribute to an increased understanding of: benefits and limitations of the use of data visualizations within diabetes apps; what users want from mobile diabetes apps; potential directions for research into future user interface paradigms and features to better serve user needs in the self-management of chronic conditions.

2 Related Work

The human computer interaction (HCI) community has long investigated the ability of digital technologies to encourage healthier behaviors. The UbiFit garden attempted to foster physical activity using a graphical garden metaphor on a mobile device [3], while Fish’n’Steps used a social mechanism, showing not only the owner’s pedometer movements but automatically sharing this information as a motivator [4]. Intille et al. [5] focused on the ability of precisely timed reminders to support healthful behavior change, while King et al. [6] promoted physical activity with early mobile device based exercise programs. The Quantified Self (QS) movement has led to much interest in the role of personal data for life optimization through personal informatics [7]. Li et al. [8] investigated which sorts of questions users seek to answer with collected data, and the shifting nature of their needs. Mamykina et al. [9–11] have written extensively on the use of computer-based systems to aid People with Diabetes (PWD). This work largely focuses on sensemaking, which Mamykina et al. define as the “perception of new information related to health and wellness, development of inferences that inform selection of actions, and carrying out daily activities in response to new information.” [11]. Kanstrup et al. [12] looked at the situational infrastructure of diabetes management in the home in order to enable participatory design of IT based systems for supporting daily life [13]. The Bant project [14] set out to iteratively develop a smartphone app which variously incorporated diverse elements into a single system: wireless data transfer from a blood glucose (BG) meter; gamification; decision support; and social elements. Storni [15] introduced an iPhone App TiY (Tag-it-yourself) which attempted to encourage reflective capabilities of diabetes monitoring. And Smith [16]
investigated the use of digital photography to aid and augment memory as a method to facilitate visualization, interpretation and reflection.

Much of this research has followed a ‘collect and reflect’ approach to behavior change, seeking to create tools to assist the user in greater self-knowledge. Many users desire automation of data recording [17]; however, this has the potential to reduce data engagement [8], leaving the insights derived from the collected data as the principal opportunity or intervention for self-management. However, successful multivariate data analysis is not a given [18], requiring what Kahneman [19] calls system 2 thinking, which is reflective and requires greater effort as opposed to system 1, which can be thought of as reflexive, intuitive and low effort. This is potentially problematic on a smartphone which is supplementing and changing human thought processes, to allow for people to avoid “effortful analytic thinking in lieu of fast and easy intuition” [20]. This suggests that smartphones, due to their size, public use, and portability could be thought of as primarily a System 1 device, with users expecting largely intuitive interactions.

The present paper describes a study undertaken to evaluate the effectiveness of current data visualizations used in popular commercial apps in order to assess their value in communicating diabetes specific information, and to better understand how this might impact long-term use, an area under-represented in current literature.

3 Methodology

There were n=13 participants in this user interaction study, with an age range from 25-45 years with a mean age of 34 years (SD±7.8). Time since diagnosis ranged from 2-26 years, with a mean of 13.5 (SD± 8.1). Three participants were female. Recruitment of people with T1D willing to undertake user studies proved to be challenging, resulting in recruitment taking longer than initially expected. Participants were primarily located through a Berlin-based diabetes and technology Meetup. We expected that this approach to selection would lead to a potential bias towards a technically literate and early-adopter test group, possibly biasing the findings towards the success of the technology. Overall, 11 of the 13 participants worked or studied in an area connected to an aspect of the study: diabetes; technology; graphic design or software design. All but one participant rated themselves highly comfortable with smartphones and 11/13 had previous experience with diabetes diary apps. Ethical approval was granted by the University Ethics Committee. All respondents were guaranteed confidentiality, and their faces were at no time recorded on video. There were no financial incentives offered.

The sessions began with a short profile questionnaire on personal characteristics, product choices, and previous and current patterns of diabetes app usage. This was followed by a semi-structured interview conducted while the participant reviewed a two-week data set, pre-entered into a range of diabetes apps. To provide authenticity and comparability, this was actual diabetes data from the lead author. This set was comprised of blood glucose levels, carbohydrate intake, exercise, and insulin dosages, for a total of 173 matched entries in each of the various apps. The exception among the apps was the Bant app, due to the restricted nature of this app, designed primarily
for the recording of BG levels. The interactions with the apps were captured with a fixed tethered camera attached to an iPhone 5s. This recorded the complete interaction of each participant with each app, along with an audio recording of the interviewer’s questions and their answers. We posit that this approach, while subject to some limitations, simulates a near future scenario of automated data entry, as well as giving methodological uniformity across apps and users.

The diabetes apps used in the study were chosen from several sources, as follows. Firstly, the study included the three diabetes logging apps most frequently mentioned by participants in a preceding pilot study: MySugr; SiDiary; and iBGStar. Three other apps were chosen from heterogeneous sources as follows: firstly an app called Bant [14], which was cited in an earlier study in an HCI context; secondly an app Roche, which has potential to be a part of an ecosystem of connected diabetes technologies; and thirdly Diabetik - a representative of a crowd funded, open source, patient initiated project. While this is a small sample of diabetes logging apps, and there are considerable variations in user interfaces, this selection is broadly representative.

4 Results

From the transcriptions of the recorded user interaction sessions, we analyzed the user’s comments to find dominant themes. We found that these apps did support broad overviews of BG management. However, this positive was tempered by, difficulties of in-depth understanding, the need for more actionable data, and in some cases, negative user-experiences related to interaction with sensitive health data.

Before the user interaction session, participants were asked on a 5-point Likert scale questions (1 for “strongly agree” to 5 for “strongly disagree”) to assess self-described attitudes towards data reflection and diabetes self-management. Participants rated themselves as generally friendly towards graphics, with a mean of 2.0 (SD 1.0) on “I like graphs and charts”, and 1.9 (SD 1.0) on “I enjoy solving puzzles”. Users rated themselves as confident on smartphone app usage with a mean rating of 1.25 (SD .6) on the statement “I am comfortable using smartphone apps in general.” While participants were reasonably positive on self-management skills, with a 2.1 mean rating (SD .7) on the statement “I am confident that I can troubleshoot my diabetes logs to understand problems.” There was a tendency toward dissatisfaction with personal control, with a 3.3 mean rating (SD .9) for the statement “I am in general satisfied with the level of my diabetes control.” Only P3 was a habitual daily user of a diabetes logging app at the time of the study, an app that he himself was developing. An additional two out of 13 participants used a logging app on occasion, one largely related to his work as a diabetes patient spokesman.
Figure 1: Accu-Chek Graph showing 14-day graph (left) and Daily diary (right)

Figure 2: mySugr diary

Figure 3: Bant app showing daily diary

Figure 4: Diabetik Journal and daily diary

Figure 5: iBGStar showing 14-day graph and table

Figure 6: SiDiary showing 14-day graph and Pie Chart
App-based visualizations offer value for general overviews

The graph functions of these apps were useful for assisting participants in obtaining an overview of glycemic control. Participants P9 and P12 both noted that a pie chart was useful for understanding the relative proportion of elevated BG values. P5 noted that the MySugr graph “…gives me a very quick graphic feedback on if am I doing good or not. If it’s flat, if it’s green… also by the steepness of the curve…(it) tells me if it’s a bad hypo or just a small hypo.” Participants were also capable of observing basic trends. Graphs were useful in this regard, and users were able to note details such as the general frequency of elevated, in range and below normal measurements. P6 reviewed the SiDiary pie chart noting elevated BG levels: “…this tells me I have to improve something, if every third test is really high blood sugar, I have to do something… only 37% is ok, and I think it should be much more.” P7 noted while viewing the iBGStar graph some elevated levels, and was able to compare them to the rest to conclude, “…they are doing alright, they have had some quite high points in some days, but generally they are ok.” However, P7 noted graphs are not useful as a daily tool to gain insights but rather to get a bigger picture of the data, “this stuff (graphs)… you would want to look at, but not on a daily basis… you want to reflect on the last week, or the last month.”

These graphs and their related diaries were seen to support an understanding of direct cause and effect, with all but one (12/13) participant readily reviewing data and engaging in basic cause and effect interpretation. For example, P3 was able to note a hypoglycemia event, leading to hyperglycemia, “…so it went down really low, then you corrected it with 45 grams of carbohydrates, then the next morning the sugar was really high.” P6, when looking at the Diabetik log, was asked what he would have done differently at night, suggested a solution based on this data representation: “…I would give (myself) more humulog, to get my blood sugar down…(then) I would test it two hours later… to control that everything is working.” P36 was able to trace through the course of a day using the Bant graph stating, “…the morning at 9:00 it starts with 14.7, its very high for a morning, then 12:00 9.0, it’s ok, 18:00 5.0 it’s very good, and in the evening 9.1 it’s a little high, but there are only four points… I think it’s not enough (measurements).”

The Challenge of In-Depth Understanding on the Smartphone Platform

Given the small dimensions of the smartphone, only limited information can be legibly shown on one screen. At the same time, displaying related information across multiple screens can be overly demanding. Many participants (7/13), noted difficulty comparing data across multiple days, as this often required viewing data on different screens or scrolling. P3 noted: “It’s not easy to compare two days, you always need to scroll up and down.” P3 was overwhelmed with the data during this comparison: “It’s really hard to compare two days….. too much information, too many numbers.” All participants found identifying meaningful connections and patterns between events challenging. For example, P10 while viewing MySugr was unable to recognize that an early meal was eaten without insulin, which caused elevated BG later. P10 reported understanding a pie chart displaying cumulative BG measurement, but couldn’t state how this data could be applied in a meaningful way, noting “One can see that values are bad, but not why.” P7 when asked to make a recommendation on what could have been done better on a specific day said that individual data points were not sufficient to answer this: “Again I have to look at each one individually… really individual measurements don’t tell you anything much because you have to look at them in aggregate, to make any decision, because there are so many variables
at play.” One app, Accu-Chek, has a logbook with relatively large and easy to read data entries, but this limits viewing to only one day per screen and does not support in-depth understanding, with P4 stating, “it’s very confusing, you check one day, but you don’t see the relations.”

Having excessive information on one screen can also inhibit understanding. When asked to correlate events on the iBGStar app, P2 noted “I’m not a doctor… I wouldn’t use this program… it’s too complicated.” Many participants found tables that simply listed data as especially hard to analyze, with P5 stating about Diabetik “I wouldn’t even consider using this, I would dismiss the whole app immediately.” However, seeking to reduce screen clutter by not having all essential information simultaneously visible also increases cognitive load, by overtaxing memory. P6 noted, “…it’s a bit confusing to not see it all at the same time, you always have to check what it is… if I ate something, or if I gave myself insulin, or if it’s my blood sugar.”

Five participants had clear difficulties processing all the needed information on the limited space of a smartphone screen. P9 noting how difficult it was to understand the Accu-Check graph data on such a small screen, stated, “… I think it would be more useful on a computer than on a smartphone.” P11 did not like the interaction requirements to access more information: “I have to swipe left and right…it’s just a bit too much…” Some participants were frustrated with the app’s capabilities on the smartphone as they had preferable experiences with the increased functionalities and display sizes supported by desktop applications. P5 explained, "Only by entering that data for a while, then importing it to the desktop version, then putting every day on top of the other, then realizing that 80% of mornings are too high, then your realize you really have a problem. So that makes sense, but this app doesn't do it."

Insufficient Contextual Information

Diabetes self-management requires rich knowledge of contextual information to inform everyday decisions. Many participants noted that the data did not have enough context to be useful, with P6 reporting “… here I see my high blood sugar, and… heavy exercise … what does it tell me… I am doing heavy exercise, or did I do some before, what did I do before at 2:00 p.m…. it was after lunch, and I had another 45 grams of (carbs)… it’s not helping me why my blood sugar was high in the evening, because I think I did everything right.” P11 said he couldn’t really understand data without knowing more about the situations that this data was embedded in, “ …I would need to remember what I did that day, how the weather was… I would need more context.”

Need for Actionable Information

While looking over collected data, 8/13 participants expressed the desire for help in understanding their collected data, and suggested that these methods of data visualization within these apps were not offering that assistance. For example P7 noted that the limited assistance from these apps still required manual effort comparable to their experiences with paper logbooks, saying, “I would have to work out everything manually myself. It is a good way of keeping my data to browse through it, but in terms of my ability to make decisions, it isn’t much better than writing it down on paper.” A common emergent theme was that reviewing of
previous data had limited application to current situations, a key concern of participants. For example, P5 stated, “Yesterday was yesterday, I don’t care, I care about today.” P5 continued about his need for intuitive advice in situ, “If this app would tell me why this value was too high, and what I did wrong, it would be really great, but it doesn’t. It tells me I had a high value here, and even after […] What I want is a very intuitive interface, that tells me what I did right... and graphically tells me what I did wrong.”

Some participants emphasized that what they really needed from a smartphone app was not the ability to review data, but rather the ability to understand how their choices would affect “what I can do better in the future” (P12). P5 acknowledged this lack of functionality and commented “Most apps haven’t helped me in looking forward because it is too complicated.” P7 gave a specific example of how the app could support decision making, but currently does not provide actionable information: “…the kind of things that are useful (are) over the last month you have consistently gone high in the morning, and then you might realize then you have to take more insulin, because it is the time of day that affects it.” When asked about how this data on this graph would help with better decisions, P7 noted the inability to translate this visualization into action: “…I would be pretty confused about how to improve it, to be honest. There is not really any indication about what to do to improve the situation. I definitely see its bad, but…” P12 when assessing the SiDiary Modal day graph said the app “…tells me you have to look better after your diabetes therapy, you have to be better.” But, when questioned as to whether it told her anything about what she had to do better stated, “No, nothing.”

Reflecting with Apps can be Emotional

Diabetes management can be frustrating, and interacting with undesired data can increase stress. P7 reflected on the Bant app, “this interface would stress me out, the red just makes you feel like oh s**t, this isn’t a good feeling when you are out of ideal range.” P10 on being asked to reflect on the data in the SiDiary app, had a visibly negative experience with the app: “I don't like it (the interface)...it's so negative... these red dots... demotivating... I didn't do a good job... because maybe I tried to have more green dots, but I failed... I think I just wouldn’t use it anymore... too much reflection of the job I did…” P9 went further and said it could be discouraging to be faced with bad numbers: “finding out that you are not doing as good a job as you can is always discouraging. It sucks finding out that over time you haven’t been doing what’s best for you.” P5 also had a negative experience with seeing the number of elevated BG levels, and replied on being asked how he would feel if the values shown were his own data stated, “S**t, I did s**t. Man you were bad. What have you done?”

5 Discussion

We have presented findings from a user study that, while limited in its scope, suggests that current generation diabetes diary apps are not adequately meeting user needs for understanding personal data. We argue that this failure is due at least
partially to an essential misunderstanding of the smartphone medium, as mobile
device users could be expecting low cognitive load [20]. We propose that for diabetes
diary apps to successfully integrate into people’s lives, these apps should stop trying
to convince users to work more, and accept that users want systems that reduce
effortful thought about their diabetes management. Perhaps as the IoT infrastructure
matures and as apps take advantage of the richer data streams, apps could move away
from a basis of retrospection and reflection, and evolve into easy to understand real-
time decision support systems, with clearly stated suggestions for actions.

**Limitations and Future Work**

As noted earlier, both the recruitment phase and methodology could have
introduced bias into some of our findings. In this section we explore some of the
limitations of this work as well as areas we have identified for future work.

*Limitations:* This study used only a limited number of apps. There may exist apps that
better support the issues raised. The focus of this study was on logging apps: other
paradigms might be beneficial in other ways not directly related to the user
understanding personal data - for example encouraging motivation. This study used
pre-entered data, whose context was therefore unfamiliar to participants. If
participants had been using their own data, it is possible that the memories attached to
data entry as well as increased familiarity with the respective apps, would have
allowed for improved sensemaking. On the other hand, we argue that the move
towards automating data streams could also reduce this familiarity. Lack of previous
experience with the apps may have influenced this study, as it is likely that with
extended use, participants would become more proficient.

*How To Filter Multivariate Data:* Health is determined by the interaction of many
interrelated factors. Ever-cheaper sensors and connectivity will lead to increasing data
streams from many potential sources. However, systems that rely on tables or graph
of such data risk overloading the user. UIs could help to meet this challenge by
supporting the user in focusing on the most pressing and essential factors, and
promoting understanding of the most important correlations between them.

*Glanceable Information for Short Interactions:* T1D demands frequent attention.
Therefore, the user should be able to obtain the desired data or advice as quickly as
possible. The UI should decrease the need for attention when the user does not need
help. While for many apps the duration of interaction could be a metric for product
success, in the case of health systems it might be a sign of failure: evidence of a
tendency to disrupt the daily routine.

*Immediately Actionable Information, Not Reflection on Data:* Many participants
indicated desire for health systems that assist with in-the-moment decision-making. It
is unclear from our participants to what degree retrospective displaying of data is
helpful for immediate decision making. Many users stated they wanted access to
information that was relevant to their current situation, not historical events to reflect
on. It might be of interest to explore the concealing or even elimination of past data,
and emphasize the best course of action for present situations by leveraging advances in data analytics and machine learning. Historic data might be visible on demand, or to illustrate a specific pattern, but otherwise largely hidden by default.

**Contextually and Emotionally Sensitive UIs for Everyday Self-Care:** As context is important for decision-making in healthcare self-management, systems could integrate location, time, and past patterns and events to get closer to predicting the current needs of the user. Sensitive situations where privacy might be a concern, such as at work, on public transit, or on a date, could affect how users interact with personal data and therefore UI design could dynamically adjust to such situations.

This paper contributes to understanding the strength, weaknesses and challenges of current generation smartphone app UI’s designed to support self-management of diabetes. This paper questions whether current paradigms for such apps match the everyday needs of the mobile user. In general, we found that expecting the user to engage in multivariate problem solving through displaying graphs and tables of collected data, was not well suited to the limitations of the smartphone. Mamykina [11] among others, have proposed that PWDs go through cycles of discovery and maintenance: it is not clear that these apps are adequate to support either phase. While much of current literature has focused on building tools to encourage and support the process of reflection and self-knowledge [9–11, 15, 21], it is unclear how much work users are willing to invest in such an approach on a routine basis. The need for understanding context to make sense of data has been previously highlighted [21]. And while this need could be supported by systems that encourage the recording of additional contextual clues, such as photos or tags, this could also increase workload and cognitive overload, creating additional barriers to habitual usage. We theorize that these apps fail to engage users, at least partially, because they fundamentally misunderstand the mobile user’s desire to reduce the need for cognitive effort [20, 22]. These diabetes management apps are primarily designed as System 2 systems, requiring careful precise logging and extensive time-consuming reflection, placing these products in direct conflict with the expectations and desires of the mobile user. Reflective, or System 2, thinking is also inherently effortful which when combined with stress provoking diabetes data, could create a major barrier to adoption.

### 7 Conclusions

We initiated and documented user interaction sessions to study how effectively diabetes logging apps helped participants understand diabetes related personal data. This was achieved by asking them to review diabetes related data, and share their observations and feelings. It appears that these apps would be well served by significant further development. Current diabetes apps offer some real benefits for data review and general overview of glycemic control and some users reported using diabetes diary apps for periodic specific problem solving. In the majority of cases, users were quite comfortable reviewing and engaging with logged data and were aware of its significance in diabetes management. However, participants struggled to
find useful correlations in collected data. We speculate that this is partially due to the limited screen dimensions of mobile devices in relation to the complexity of collected data, and that users expect that interaction with a smart phone app will not be cognitively demanding, but rather a smooth and intuitive experience. Health data can have emotional impact. Such potentially negative interactions might be especially challenging in the uncontrolled environment of public use. Automation by way of connected sensors could provide us with the tools to assemble higher quality data, but the gap between data collection and better decision-making is still large. Additionally, as there is reduced data engagement that comes with automation [8], new tools may need to be developed to bolster engagement.

Our interpretation of our participants’ responses suggests that rather than tools for interpreting and learning from past events, as the reflection paradigm implies, users want personalized and easy to understand advice for the near term. Such advice should be easily and quickly accessible. The smartphone, due to its screen size and mobile nature, seems ill suited for cognitively demanding analysis. We conclude that the emphasis should be placed on systems that largely remove the demands of reflection, and instead deliver meaningful analysis of collected data with clear suggestions for action.

8 References


