SUMMARY

This article addresses how a mobile application can aid type 2 diabetics (T2D) to prevent health risks associated with nocturnal hypoglycemia. Hypoglycemia is defined as an unhealthy drop in blood glucose (BG) levels. This article examines how the output of a predictive algorithm can be visualized to signal a warning if a hypoglycemic event is expected to happen during night. The predictive algorithm is based on sensor technology which can track one's BG every five minutes. The purpose of the study is to evaluate how T2Ds can make use of a predictive application and how they perceive different design alternatives. The article also discusses barriers and enablers for developing a digital supporting tool targeted for users typically from an older segment. Furthermore, the article discusses how the user expectations align with the possibilities a predictive machine learning algorithm offers.

Methodologically we examine the topic through interviews (N=10) with people with diabetes supplemented with an interview with a doctor specialized in diabetes. Through thematic analysis we elicit knowledge about user needs and preferences, which is used to determine a best suitable design for a predictive application. With this knowledge we develop an Android prototype application in Kotlin. The application is evaluated in a think-aloud evaluation (N=5) with focus on how to take individual preferences into account and enabling the users to take appropriate action. Through deductive coding we establish knowledge about how the users respond to the prototype application.

The findings from both the interviews and the evaluation are summarized in a recommendations table with guidelines for creating a self-management application (SMA) for presenting AI-based predictions for people with T2D. Our findings generally align with previous research emphasizing the use of simplicity, familiarity, color cues and visualizations to display quantitative data. From our interviews we also find that the users ask for flexibility to personalize some of the settings for the prediction. The users also emphasize the need for avoiding user inputs when possible. Furthermore, we find it necessary to consider accessibility for the application both regarding cost of required technology and endorsement from treatment providers.

Our research contribution is in part that we provide contextual knowledge related to the domain we have worked within, as well as the type of users we have involved. We draw parallels to previous research recommending user-centered design and models for technology acceptance. However, we have also seen a shortage of research involving T2D in visualizing health predictions. Our research contribution is presented as recommendations for designing and developing predictive diabetes SMAs. We find that most participants would never have considered the idea of predicting hypoglycemia if we had not presented it to them, which is why we suggest striking a balance between designing for and with users. We find that some user requirements are not technically feasible and therefore we suggest close collaboration between health professionals, people with diabetes and researchers. Future research directions are suggested; these include evaluating in real-life settings with additional users as well as studying barriers for long term adoption.

LIST OF ABBREVIATIONS

BG	Blood glucose
CGM	Continuous glucose monitoring
SDCN	Steno Diabetes Center Nordjylland
SMA	Self-management application
T1D	Type 1 diabetes
T2D	Type 2 diabetes
TAM	Technology Acceptance Model
UTAUT	Unified Theory of Acceptance and Use of Technology

Alerting Type 2 Diabetics of Nocturnal Hypoglycemia Risk

Integrating user-centered design in a self-management app to present AI-based health predictions

ANDERS SOLSBÆK OTTOSEN, Aalborg University, Denmark
ANNE SCHACK GUSTENHOFF, Aalborg University, Denmark
MADS SIGH LUND ANDERSEN, Aalborg University, Denmark

Abstract

People with diabetes who are insulin treated are at risk of experiencing hypoglycemia – a condition of potentially dangerously low blood glucose levels. Predictive algorithms can aid these people in preventing hypoglycemia. In this two-part study we applied user-centered design to explore how a personalized nocturnal hypoglycemia risk prediction can be communicated in a mobile self-management app (SMA) for people with type 2 diabetes. In a preliminary understanding study, we conducted semi-structured interviews with potential users (N=10) and one endocrinologist to acquire domain knowledge, establish user needs, and identify preferences for visual presentation of hypoglycemia risk. Based on our initial findings, we developed a mobile application prototype which was evaluated in a think-aloud evaluation with prospective users (N=5) to validate user needs and verify design and functionality. Our findings generally align with previous research in emphasizing the significance of using color cues and simple visualizations supplemented with concise text or numbers. We also found that the SMA should provide flexibility as users need different levels of detail and that the user should be able to personalize the prediction configuration. We further identified barriers and enablers for adoption of mobile diabetes SMA and highlight the importance of collaboration between users, health professionals, and researchers. Based on our findings we draw recommendations for future work in this field.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: self-management app, mHealth, user-centered design, AI prediction, diabetes, hypoglycemia

1 INTRODUCTION

In Denmark, approximately 252.000 people are diagnosed with type 2 diabetes (T2D) and it is estimated that the number will rise dramatically over the coming years [12]. The average onset of T2D in Denmark is 63 years [18]. One of the health risks people with diabetes experience is hypoglycemia. Hypoglycemia is defined as a drop in blood glucose (BG) to unhealthy levels and can be a side effect of insulin treatment in diabetes patients [25]. It is a potentially dangerous condition for people with diabetes, and can lead to loss of consciousness, confusion, seizures, and in extreme cases, death. Additionally, hypoglycemia is often symptomless, and over half of extreme cases happen during sleep [25]. In a Danish study Dømgaard et al. [13] found that 35% of insulin treated T2D feared mild nocturnal hypoglycemia while 56% feared severe nocturnal hypoglycemia. The fear of hypoglycemia can influence self care, as some patients attempt to avoid low BG levels by having continuously high BG [13]. Hypoglycemia and hyperglycemia (high BG levels) and their adverse health effects necessitate that people with diabetes continuously monitor their BG levels. To assist people with diabetes in maintaining safe and healthy BG levels a wide range of solutions has been created. Still, people with diabetes often have to monitor their BG level several times a day and take action, such as adjusting insulin dose, or consuming a small amount of fast-acting carbohydrates, when hypoglycemia is detected [32].

In recent years development of predictive artificial intelligence has aimed to solve some of these issues focusing mostly on people with type 1 diabetes (T1D) [25, 32]. By leveraging health data, algorithms that can inform patients and help decision-making are developed to supplement self-management. With these predictive algorithms it is relevant to examine how individuals perceive, understand, and act on personalized predictions while also considering how interactive systems incorporating personalized predictions should be designed [11]. As Desai et al. [10] demonstrate, users may experience difficulty utilizing the self-monitoring data, if the data is presented in a way that is not suitable for the people with diabetes. This issue is further relevant as existing diabetes self-management applications (SMAs) are often not used [30]. Additionally, research is limited concerning integration of personalized predictive algorithms in SMAs. To leverage the full potential of a predictive algorithm, people with diabetes must be able to make sense of and act on the predictive result.

In this project, we work with an algorithm developed by Steno Diabetes Center Nordjylland (SDCN) and Aalborg University. The algorithm uses BG data as input and logistic regression to predict risk of hypoglycemia for the following night from midnight to 6 a.m. The algorithm is based on a data set from insulin treated T2D and is trained on a definition of hypoglycemia as three consecutive BG measurements below 3.0 mmol/L.

The aim of this study is to explore how the results of the algorithm can be communicated in a mobile SMA to alert T2Ds. Based on semi-structured interviews with potential users (N=10) and one endocrinologist, we acquire domain knowledge, establish user needs, and identify preferences for visual presentation of hypoglycemia risk. These findings inform the design of a mobile app prototype which is evaluated with potential users (N=5) to validate user needs and verify design and functionality. We identify preferences towards the use of color cues and simple visualizations combined with concise supplementary text or numbers to convey predicted hypoglycemia risk. We further identify a need for personalizing the prediction's definition of night and threshold for hypoglycemia. Our findings also indicate that users have very different preferences when it comes to the amount of detail needed to take appropriate action. We therefore suggest, designing for a flexible user interface which presents a simple visualization to all users and provides additional information through interaction. Finally, we identify barriers and enablers for adoption of diabetes SMAs. We highlight the need for collaboration between users, health professionals, developers, and researchers to overcome these barriers. Based on our findings, we list 12 recommendations for future work with diabetes SMAs and presenting personalized AI-based health predictions. These recommendations focus on both design and adoption and constitute the main contribution of this work.

2 RELATED WORK

Literature concerning diabetes self-management is comprehensive. Three main areas have been identified as significant for the research area addressed in this paper. These are related to 1) supportive tools for diabetes self-management, 2) presenting health data to users, and 3) user-involvement in diabetes app development.

2.1 Supportive Tools for Diabetes Self-Management

Numerous technologies for supporting diabetes management have been developed and improved over time. These address improvements in areas such as insulin delivery, detection of dysglycemia, and prediction of hypoglycemic events [34]. Continuous Glucose Monitoring (CGM) sensors have proved to be a revolutionizing technology making it possible to track data from people with diabetes to analyse BG levels and support decision making related to insulin administration [32]. This can happen retrospectively by health professionals who can review previous data but also in real-time to predict and prevent future events of dysglycemia by warning people with diabetes to take mitigating actions [32]. However, CGM sensors are expensive and unavailable to many people with diabetes. Additionally, they can be uncomfortable to wear, difficult to operate, and provide confusing outputs and false alarms [19, 23].

Many SMAs have been developed to help people with diabetes. In one review [14], 77 of 104 studies reported benefits of using digital technology to support self-management decisions for people with diabetes. However, a 2019 study found that only 18% out of 87 selected apps were research-based [3]. Additional studies have shown that personalized research-based applications often depend on manual inputs [11, 19]. As Årsand et al. [37] point out, however, avoiding manual inputs and utilizing automatic data transfer is very important for user acceptance, especially for applications developed for long term use. A 2017 study [30] further demonstrated that a large portion of diabetes SMAs are either not used or have little to no effect.

2.2 Presenting Health Data to Users

It is important to consider how data can be presented in a way that enables the users to clearly identify how they will benefit from the technology. In a study on visualizing BG level forecasts to T2Ds Desai et al. [10] argue that computational models can only reach their potential if they are understood by non-expert users. The same study describes a rich-simple design paradox where users seek solutions that are easily understood but also provide comprehensive information. Other studies point to how complex data and too much information can alienate users [14, 32]. Desai et al. [10] suggest that one way to reduce complexity in data representation is by incorporating visualizations, numbers, icons, and pictographs. Research has also described color cues as easily interpretable – as an example by replicating a traffic light scheme or having *red* associated with danger/stop [10, 15, 29, 36]. Research has further emphasized visual displays and graphs over table formats, especially for quantitative data, striving for simplicity and for the user to obtain the intended interpretation of the data [1, 36]. Furthermore, studies have described that supplementary text is often necessary [1, 3].

2.3 User Involvement in Diabetes App Development

A 2013 review [14] argues that user-centered and socio-technical design, as well as agile methods should be built into the planning, design, and implementation of digital interventions for diabetes self-management. This is to enhance usability and ensure a better fit between technology and user. One study from 2019 [2] argues that understanding the desires of people living with diabetes needs to be the first step in app development to ensure that the app provides features and benefits for the users. A 2022 review [27] calls for qualitative research on users' experiences with mHealth apps as a prerequisite for future mHealth app development. It further argues in favor of collaboration between manufacturers, health professionals, regulatory bodies, researchers, and people living with diabetes.

Designing SMAs for T2Ds is further complicated by the age of end users. Studies have shown that younger people are more likely to benefit from mhealth apps than older people [20, 29]. Fischer et al. [16, 17] have studied involvement of older users in technology design. In a 2020 article, they claim that *design agendas* – preconceived user goals for a particular technology – are set by designers representing a power asymmetry. They introduce the method of *unfettered design* which implores designers to have an increased focus on four concepts: ongoing reflection, retained impartiality, the participants' views, and remaining flexible throughout the design procedure. Unfettered design represents the authors' democratic ideal of equal collaboration with older adults. Yet, their practical guidelines for how to do unfettered design seem a little meager.

3 METHOD FOR PRELIMINARY UNDERSTANDING STUDY

In this section, we describe methods used to gain domain understanding and present users' reflections on user interface designs. The methods used address the research question: *What are users' needs in relation to managing a stable blood glucose and what are their preferences for visual presentation of hypoglycemia risk?*

3.1 Participants

The participants were recruited through Facebook, personal network, and through a contact person at SDCN. Interviews were conducted with ten participants and one endocrinologist, a 58-year old doctor specialised in diabetes. The intended users were insulin treated T2Ds. We did, however, have difficulties recruiting such persons. We therefore included T2Ds without insulin and T1Ds as they are insulin treated and have experience with hypoglycemia. A list of the participants can be seen in table 1.

Table 1. List of participants

Participant	Age	Gender	Occupation	Type	Years with diagnosis	Treatment
1	61	M	Warehouse Worker	1	55	Insulin
2	69	F	Retiree	2	21	Insulin
3	59	M	Self-employed	2	4	Metformin
4	23	F	Student	1	15	Insulin
5	69	M	Retiree	2	18	Insulin
6	67	M	Retiree	2	8	Metformin
7	69	M	Retiree	2	10	Metformin
8	69	M	Retiree	1	51	Insulin
9	51	M	Nurse	2	19	Insulin
10	67	F	Retiree	2	19	Insulin

3.2 Interviews

One purpose with the interviews was to understand daily challenges and use of existing technology amongst the participants to determine the scope of the application. We prepared an interview guide for the semi-structured interviews. The interview with the endocrinologist and a pilot interview with a T1D were used to sharpen the questions and avoid misunderstandings. The interviews were guided by four overall questions:

- How do participants cope with daily challenges related to their disease?
- Are the participants currently using any other diabetes technology?
- How much do the participants worry about hypoglycemia?
- Do the participants trust an algorithm to make predictions about their health?

3.3 Materials

A second part of the interviews involved reflections on potential visual designs for alerting the user of predicted hypoglycemia risk. We motivated the participants to draw initial sketches before presenting them with predeveloped mock-ups. The mock-ups were presented to support the dialog as suggested by Faiola et al. [15] and Vistisen [33]. They were exploratory as we sought to understand what the participants considered necessary information and how they would perceive alternative mock-ups. In total, 16 different designs were shown. Some of the mock-ups were intended as alternatives while others were supplementary. The mock-up designs were based on prior research and focused on three different topics: 1) visualizing the prediction, 2) additional information, and 3) becoming aware of risk. The mock-ups were inspired by an existing application, Glooko[1]. Mimicking this application was suggested by our stakeholder at SDCN as this was expected to be the platform on which the predictive algorithm would be implemented in the future. The mock-ups can be seen in Appendix A.

Visualizing the prediction – appx. fig. 11-14
Studies have shown, that graphical formats have higher appeal amongst users with low digital literacy and older adults with less digital experience [1, 29]. Our first designs aimed for simplicity with familiar visualizations and limited text and numbers. For this first topic we presented four different designs. One example is seen in figure 1. The smiley is a familiar design as seen in numerous studies (e.g. [10, 22]). The other designs for this topic were a pendulum, a person in distress, and a BG measuring device. The pendulum was mostly inspired by Desai et al. [10] with a clear distinction between the color cues red (bad) and green (good). The person in distress was a metaphor suggesting consequences if the warning was not acted on, and was influenced by Desai et al. [10] and Laughery and Wogalter [22]. The measuring device was a graphic representation of a real-life object as suggested by Laughery and Wogalter as well [22].

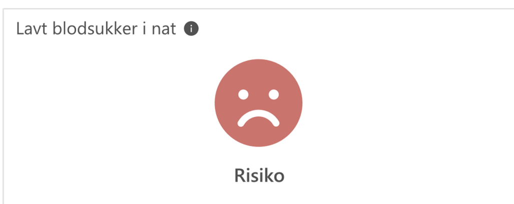

Fig. 1. Smiley

Additional information – appx. fig. 15-23
We wanted to explore what kinds of information the participants preferred in addition to a simple risk/no risk visualization. One example of additional information as seen in figure 2 was a graph giving more insights about forecasted BG level. The other designs had additional information related to the specific time of the predicted hypoglycemic event, the predicted BG value, prediction uncertainty, suggestions for action, details

[1] www.glooko.com

about how the prediction is made, and tendencies in hypoglycemic events over time. We used graphs and bar charts for displaying quantitative data as these are well known data-visualization graphics for depicting numeric values in a intelligible manner [1, 10, 36] as seen in figure 3, we added additional text advising how to act to prevent a hypoglycemic event.

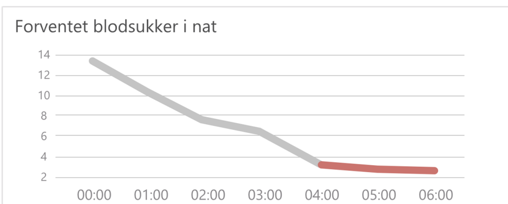

Fig. 2. Graph

Fig. 3. Smiley with suggestion for action

Becoming aware of risk – appx. fig. 24-25
To support awareness of predicted risk we designed a notification to be displayed on the smartphone lock screen as seen in figure 4. The designs concerning awareness draw inspiration mainly from the Glooko application. In general, however, we did not restrict the interviews to what would fit into the Glooko application but it provided context and scope for our additional features.

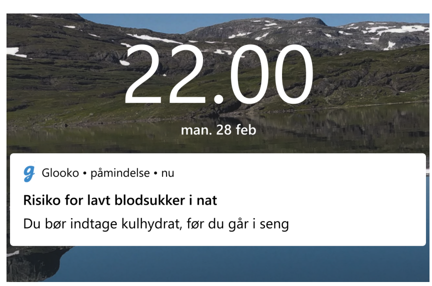

Fig. 4. Notification

3.4 Analysis

The interviews were transcribed and coded following the phases described in Braun and Clarke's [4, 5] reflexive thematic analysis approach. Reflexive thematic analysis is a flexible approach that acknowledges the researcher's position and prior knowledge as impactful but also adaptable to new knowledge when analyzing patterns within data [4, 5]. We developed themes in an iterative approach with data familiarization and coding of data as basis for searching for, reviewing, and defining themes. We coded all 11 interviews together and discussed our findings using a thematic map to gradually refine the themes.

4 FINDINGS FROM PRELIMINARY UNDERSTANDING STUDY

During analysis we developed two main themes: *adopting diabetes technology*, and *determinants for individual preferences*. The first theme concerns general challenges in managing diabetes and the participants' attitude towards digital supportive tools while the latter covers how alternatives in application designs address individual preferences. This section elaborates on these themes and further presents an overview of perceived strengths and weaknesses of our mock-up designs, based on the participants' reflections during interviews. This is seen in table 2.

4.1 Adopting Diabetes Technology

The endocrinologist pointed out that numerous diabetes SMAs already exist but are rarely adopted. Most participants (P2, P3, P4, P5, P6, P7, P8, P9) corroborated this by expressing that they were aware of the existence of diabetes SMAs but did not currently use any. From the interviews, we found that the adoption of self-management technology relates to the individual's attitude towards the technology and is impacted by three main factors: *challenges and needs*, *relationship with treatment provider*, and *knowledge about existing technologies*.

Challenges and needs

The participants experienced different challenges related to living with and managing their diabetes. Several participants (P1, P2, P4, P5, P10) expressed that unpredictability in their daily lives was a challenge. All participants had experienced events of spontaneous dysglycemia that seemed inexplicable at first but most likely related to minor, daily changes in behavior. In the words of P10:

> *If I do more or less the same things every day, then I can figure out what it* [BG level] *will be. But as people say to me "You're not a machine - you're a human being."*
>
> (P10)

Other challenges related to adhering to the recommended treatment, diet, and level of exercise as well as remembering to measure BG level. Several participants expressed skepticism towards mobile apps where the user must enter data about meals, insulin intake, BG measurements, etc., as they viewed this as tiresome and would not bother manually entering data. It became clear that the users had a desire towards more automated solutions that do not add extra workload on to the existing daily practices in managing their. Some of the user inputs can be relieved by adopting and connecting additional technologies such as CGM sensors, smart insulin pens, or activity trackers. Here though, the price point and lack of government subvention for these products was expressed as a barrier for many. Overall, we found that the individual's challenges and the perceived threat of their diabetes impact their attitude towards self-management technology.

Relationship with treatment provider

Another factor related to adoption of new self-management technology is the relationship between treatment provider and patient. Some participants expressed that a digital self-management solution could in some cases replace the treatment provider and help them manage their diabetes more independently. For others, the technology was viewed more as a supplement to the treatment provider. Affirmation, guidance, and even some disciplinary action was still wanted from the treatment provider. Almost all participants, however, expressed that they preferred and even expected their treatment provider to recommend technologies and vouch for their credibility. On this matter, however, the endocrinologist expressed that being able to recommend and vouch for digital self-management solutions like mobile apps is often a struggle as there are no national standards for these technologies.

Knowledge about existing products

When it comes to knowledge about existing technologies, most participants fell into two groups – they either had limited knowledge about existing products (P1, P10), or they were aware of the existence of current technologies, but felt overwhelmed by the quantity of them (P2, P3, P4, P5, P6, P7, P8, P9). In the words of P5:

> *I mean, I have done some searching but it's a jungle. I've mostly been guided by the outpatient clinic or the Diabetes Association. They are the ones who know the most about it.*
>
> (P5)

All participants had a finger-prick device to monitor their BG level and while multiple participants expressed that they would like to use additional technologies, only one participant (P4) did. Four participants (P5, P2, P9, P10) had previously been part of a study where they had to wear a CGM sensor. P4 had further researched technologies by herself and tried them out without endorsement from a treatment provider.

4.2 Determinants for Individual Preferences

Two design aspects that almost all participants alluded to were *familiarity* and *simplicity*. Yet, the interviews suggest that what one considers familiar and simple may differ from another. Several participants as well as the endocrinologist expressed that they are very aware of these individual differences. The amount of additional information needed, and the preferred style of presentation relates to individual context and preferences. From the interviews, we identified five main areas that impact the users' preferences: *knowledge and experience with diabetes*, *routines and habits in managing diabetes*, *level of concern*, *digital literacy*, and *situational circumstances*.

Knowledge and experience with diabetes
Most participants expressed that when diagnosed with diabetes, there is a significant amount of knowledge, skills, and behavior that needs to be learned. Some is acquired through the treatment provider, the internet, or communities like the Diabetes Association while some is internalized through experience. In relation to avoiding nocturnal hypoglycemia, there were noticeable differences between the participants' levels of experience. Three of the participants (P3, P6, P7) had never experienced hypoglycemia. Another three (P1, P2, P5) estimated that they experienced mild hypoglycemia once every two weeks. Three participants (P4, P8, P10) experienced mild hypoglycemia several times a week, and the last participant (P9) had previously experienced hypoglycemia but not for years. The way the participants experience hypoglycemia and how they define it also differed. In the words of P8:

> *Some are able to feel it at night – I am. You know, your sleep is restless and it's in the back of your mind that something is wrong so you get up and measure your blood sugar. But not everyone can* [feel it]. *They sleep through it and get low blood sugar.*
>
> (P8)

Most participants (P2, P4, P5, P8, P9, P10) would feel symptoms and discomfort at BG levels between 3.5 and 5.0 mmol/L. Two participants (P5, P8) pointed out that the ability to register hypoglycemia may diminish with age. By all accounts, the current BG threshold for hypoglycemia at 3.0 mmol/L in the algorithm was considered too low by all participants except P1. Most also preferred a customizable threshold, so the individual can fit the algorithm to their own definition of hypoglycemia.

Knowledge about treating and preventing hypoglycemia was also not consistent across the group. All insulin-treated participants expressed that drinking juice or eating dextrose tablets will treat acute hypoglycemia. But concerning prevention of predicted low BG level several hours away the answers were more muddled. The endocrinologist suggested eating slow-acting carbohydrates like bread as the most appropriate strategy as drinking juice (fast-acting carbohydrates) will increase BG levels temporarily but will not stabilize it over several hours. Some participants expressed awareness of this difference between fast- and slow-acting carbohydrates but others gave more ambiguous answers indicating uncertainty on the matter.

Routines and habits in managing diabetes
The interviews suggest that the participants' current routines and habits affect their preferences for SMAs. The way the participants managed their diabetes differed. Some measured their BG level several times a day. P4 could monitor her BG level on her lock screen continuously, while P7 would only measure BG level every third month when meeting with his treatment provider. Some participants were used to interpret data when managing their diabetes and they would prefer access to more detailed information. Others expressed that they would like the application as simple as possible. As P10 expressed:

> *I've got a comprehensive system to remind me* [about insulin intake]. *Yet I sometimes forget. I think that there shouldn't be too much* [information] *it can make one stress about it.*
>
> (P10)

Level of concern

The way participants perceived their own disease and their level of concern for nocturnal hypoglycemia also varied. Most participants (P1, P2, P4, P6, P7, P8) expressed that diabetes was a natural part of their lives and that they were not gravely troubled by it. When it comes to hypoglycemia specifically, only three participants (P5, P8, P9) reiterated this attitude. In the words of P5:

> *It's not something I fear. If it happens, it happens and then I know how to handle it.*
>
> (P5)

The others expressed more concern and took preventative measures such as relying on their spouse to wake them up at night if they detect symptoms. Some would make sure to measure their BG level before going to bed and eat accordingly. P1 mentioned never going to bed with a BG level under 4.0 mmol/L while P10 preferred going to bed with a BG level of 10.0 mmol/L. The more concerned generally preferred loud alarms for prevention while those less concerned preferred more unobtrusive ways of receiving the prediction of risk such as a simple notification.

Digital literacy

The interviews suggest that a participant's "digital literacy" affects how they perceive an additional digital supporting tool. It also affects their preferred level of detail. When presented with the graph visualization of BG levels for the entire night (see figure 2) P6 commented on these differences:

> *One of the strengths is that you are able to track it very accurately and see when it becomes critical. But the weakness is that the information is a little redundant to me. In general, you shouldn't create something that only appeals to "geeks".*
>
> (P6)

The individual's trust in the presented solution as well as trust in digital technologies in general also impacts their attitude and perceived benefit. Some believed that the solution could replace current practices of having alarms and manually measuring BG levels at night while others were more skeptical and imagined that they would continue their current practices as a back-up or accuracy-checking mechanism. The participant's level of trust was also an indication for how much information about the algorithm they found necessary. Only two participants (P4, P6) expressed an interest in knowing how the risk prediction is calculated and being presented with uncertainty in the prediction.

Situational circumstances

Most participants point to the fact that one of the main challenges of managing diabetes is that not all days look the same. Changing situational circumstances make treatment and self-management more difficult. Especially vacation, parties, and illness appear to be challenging circumstances for some participants. They report that they forget to monitor their BG level, forget insulin doses, and that alcohol intake or a fever can affect BG levels, making it unpredictable. At parties or on vacations, three participants (P4, P5, P7) would like to be able to mute or snooze alarms warning them about nocturnal hypoglycemia risk. Most participants, however, actually want the alert to be constant so they at least know that they will be warned even under changing circumstances.

4.3 Design Strengths and Weaknesses

In addition to the presented findings, the interviews also provided feedback on the visual designs for the application. We asked participants to sketch and elaborate on ideas for how they wanted the information presented. Only five participants made sketches and most of the participants had difficulties concretizing their requirements for the solution. We therefore highlight perceived strengths and weaknesses of our mock-up designs. They are grouped by topic and summarized in tables 2, 3, and 4. The tables refer to the mock-ups found in Appendix A.

Table 2. Visualizing the prediction

Visualizing the prediction	
Strong Elements	*Simple and familiar* Six of the ten participants preferred the smiley visualization (fig. 11) and praised it for its simplicity and familiarity. *Colors* Five participants emphasized the use of colors as an easy way to make a quick, effortless interpretation of the visualization.
Weak Elements	*Too simple, and too familiar* Several participants commented on the smiley visualizations (fig. 11) simplicity as a weakness. Five participants expressed a wish for it to be supplemented with numbers. Two other participants expressed that the smiley was "too familiar" and that they were tired of seeing it used "everywhere". *Unintelligible visualization* Two visualizations were found difficult to interpret. This was the case for five participants for the pendulum (fig. 12) which was considered complex and confusing. Four participants were skeptical of the BG measurement visualization (fig. 14) which was found non-intuitive due to its use of less than/greater than symbols. *Inappropriate metaphor* Five participants expressed dissatisfaction with the distressed person visualization (fig. 13). The metaphor for hypoglycemia was deemed childish or too dramatic.

Table 3. Additional information

Additional Information	
Strong elements	*Display of tendencies over time* All but one participant were positive towards being able to see tendencies over time (fig. 17) and suggested that the information could be used for tracking behavior and adjusting treatment or daily activities. *Details about time of risk* Four participants preferred having details about the predicted time of the hypoglycemic event (fig. 15 and 16) as this could be used for adjusting one’s preventative actions. *Specific suggestions for action* Four participants were pleased with specific suggestions for action such as "Eat a banana" (fig. 21). Three of them, however, had never experienced hypoglycemia.
	Continued on next page

Table 3 – continued from previous page	
Weak Elements	*Display of graph* Six participants expressed that the graph presenting forecasted BG values for the entire night (fig. 17) was unnecessary and not helpful when trying to prevent nocturnal hypoglycemia. *Display of uncertainty* Eight participants found the display of the model's uncertainty (fig. 18) unnecessary and too complex to interpret. *Suggestions for action* Five participants considered the suggestions for action (fig. 21) redundant as they believed that people with diabetes ought to know these things. They were also split between finding the suggestions either too general or too specific.

Table 4. Becoming aware of risk

Becoming aware of risk	
Strong elements	*Notification* All participants preferred receiving a notification to be made aware of a predicted risk. *Customization* Seven participants emphasized the possibility for customizing the notification. This was mostly related to varying bedtimes and preferences for time span to react to the risk before going to bed.
Weak Elements	*Disruptive alarm* Several participants called attention to how some people with diabetes are already bombarded by different alarms (BG measurement reminder, medication reminder, CGM registered drop in BG level, etc.). To them another alert could cause irritation and lead them to avoid alert technology of any kind.

5 METHOD FOR EVALUATION STUDY

In this section, we describe how we used our understanding of users' experiences and design preferences to develop an application prototype and evaluate it with potential users. This part of the project was guided by the research question: *How can we present a person's current nocturnal hypoglycemia risk in a diabetes SMA in a way that accommodates individual preferences and enables the user to take appropriate action?*

5.1 Creating the Prototype

Leaning on agile methodology and principles from user-centered design, we developed personas and user stories to aid the development process. We followed suggestions from Pichler [28] to describe personas. These fictional characters represented user archetypes and described which problems the application should address. For creating user stories, we followed a template suggested by Cohn [7] to ensure that all stories addressed questions of *who* desired *what* functionality and *why*. We created two personas which can be seen in Appendix B as well as 11 user stories of which two examples can be seen in table 5. The complete list can be found in Appendix C. We collectively prioritized the 11 user stories and used them to form our initial backlog for development. We created a visual prototype to illustrate the visual design and navigation of the intended application. The overall layout of the visual prototype was based on the existing Glooko app. The visual design for specific features related to our context was created by combining and refining some of the previously created design mock-ups from the understanding study. When developing the app, we adopted a peer-programming approach while continuously discussing and refining user story details. The application was programmed in Android (Kotlin). Screen shots from the application can be seen in figures 5, 6, 7, and 8.

Table 5. Examples of created user stories

#	**Story**
1	As Lise, I would like to be able to see a simple visualization of the predicted risk, so that I can act on it before I go to bed.
2	As Jan, I would like to see additional information indicating the time of predicted hypoglycemia, so that I can adjust my preventive actions.

5.2 Participants

We evaluated the application with five participants. Two of these were recruited from the first group of interview participants. Another two were recruited through the Diabetes Association in Aarhus while one participant was recruited through a contact person at SDCN. We selected these participants because they perfectly matched the intended user profile – insulin-treated T2D who had experienced hypoglycemia. An overview of the evaluation participants can be seen in table 6.

Table 6. List of participants

Participant	Age	Gender	Occupation	Years with diagnosis
2	69	F	Retiree	21
10	67	F	Retiree	19
11	59	F	Driving instructor	24
12	76	F	Retiree	22
13	66	F	Retiree	24

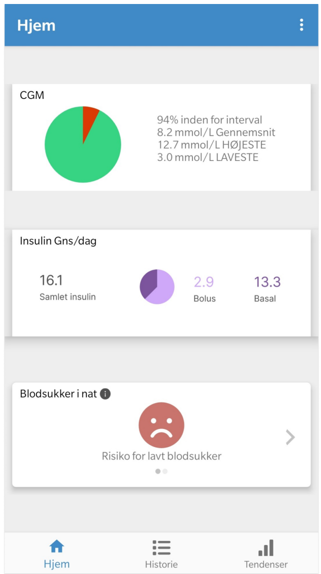

Fig. 5. 'Home' tab displaying CGM values, insulin intake, and predicted risk for nocturnal hypoglycemia.

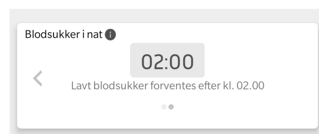

Fig. 6. Displaying expected time of the predicted hypoglycemia risk. Visible if user swipes left.

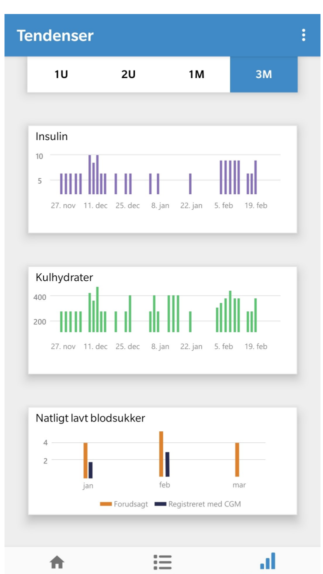

Fig. 7. 'Tendencies' tab displaying insulin intake, carbohydrate intake, and nocturnal hypoglycemia tendencies over the past three months.

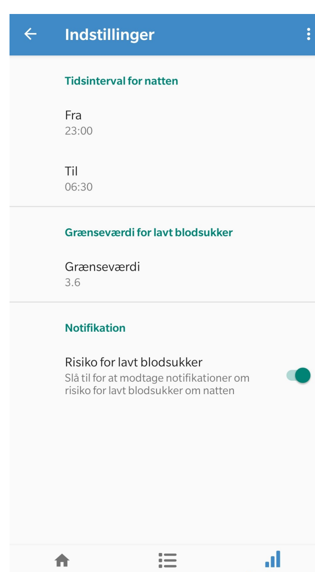

Fig. 8. 'Settings' tab displaying personalization options for nocturnal hypoglycemia risk prediction and alerts.

5.3 Evaluation Interviews

The purpose of conducting an evaluation was twofold. First, we wanted to validate the user stories and thus our understanding of the users' requirements for the application. Second, we wanted to verify the application's functionality and visual design to decide whether it satisfied the intention of taking individual preferences into account and enabling appropriate action. As Crispin and Gregory [8] describe, different kinds of tests and assessments should be used throughout an agile development project. We prioritized critiquing the product from a *user's* perspective by emulating the way a real user would work the application. The evaluation served as an early-stage user acceptance test to provide us with feedback for product adjustment and to enable us to discuss further implications for design.

The evaluation interviews were structured into two parts. First, we did a think-aloud evaluation with the participant and secondly, we interviewed the participant with prepared questions related to adopting the application. The first part was intended to identify bugs and usability problems and to validate our created user stories. The second part was intended to evaluate the prototype on a more abstract level. We sought to identify potential new user stories and to asses the perceived value of the application. We also probed for issues related to taking the application into use. This did not only concern design-related issues, but also questions about how the participant would discover and start using the application. In preparation for the think-aloud evaluation, we developed five scenarios which reflected key user stories. Three of these had elaborating sub-scenarios. An example of this is provided in table 7, the complete list can be found in Appendix D. During the evaluation, we provided the participant with a smartphone with the application installed on it. We then asked the participant to read the scenarios out-loud one at a time and subsequently interact with the application while describing their thoughts and actions verbally.

Table 7. Examples of developed scenarios

#	Scenario and task
3a	You have been using the app for a while, but you don't understand why it only rarely sends you a notification. You're wondering if the threshold for low blood glucose might be defined as too low in the app. Find out what the threshold for low blood glucose is defined as.
3b	Your threshold for low blood glucose is around 3.8 mmol/L. Change the app's definition to match yours.

5.4 Analysis

All evaluation interviews were recorded and transcribed for analysis. Each group member read through all interviews while noting down the participants' feedback. We performed concept-driven deductive coding with predefined codes inspired by our research question. We focused on feedback related to user needs and requirements, design and usability, as well as general feedback. We discussed our findings within the group and categorised them into five topics: *benefits of use*, *accommodating individual preferences*, *enabling appropriate action*, *adopting the app*, and *usability problems*. We included the topic *usability problems* to be able to separate usability-related problems from feedback on the other four topics.

6 FINDINGS FROM EVALUATION STUDY

Our findings generally aligned with the results from our preliminary understanding study. However, the context changed as the designs were now integrated in our imitation of the Glooko application. This section presents our findings.

6.1 Benefits of Use

The participants had a positive attitude towards the usefulness of the application. The general opinion was that the participants found the application interesting and useful if it would allow them to sleep without fear of hypoglycemia. As P2 expressed:

> *I think the application is great. I can use the functionalities. It is something I would use and it would provide me with some security that I can not only feel low BG levels but also see it.*
>
> (P11)

P13 emphasized a preference for data to monitor fluctuations in BG levels as it can be difficult for the diabetic to register:

> *When you have diabetes it can be difficult to see it. You need to measure something. I think it would be nice, if I could access information on my phone. I often look at my smart watch to see how I have slept during the night and to see my heart rate and blood pressure. It is exiting to track.*
>
> (P13)

P10 and P12, however, doubted whether they would actually make use of a such application if accessible. P10 expressed some confusion about how the application could improve her current practices, and what the goal of the application was. P12 did not experience hypoglycemia but saw a strong potential for other people with diabetes with less knowledge about how they should manage their disease:

> *There are many diabetics and the number is increasing. The more you can anticipate* [with supportive technologies]*, the better. Then everyone wins.*
>
> (P12)

One functionality that was highlighted by P2, P11, P12, and P13 was the tendencies tab (see figure 7) showing predicted and experienced hypoglycemic events for the previous weeks or months. It was particularly expressed by P2 and P12 that this overview could be used in dialogue with the treatment provider. P11 and P13 would also use tendencies to determine the intake of insulin or food at certain times.

6.2 Accommodating Individual Preferences

For selected functionalities we had designed the application to accommodate individual preferences. The application allowed users to adjust the definition of night by changing start and end time. The application also allowed users to change the BG threshold for hypoglycemia that would activate the notification. All participants except P2 wanted to change the threshold value from the default. All participants expressed that they would change the night interval setting from the default. P2, P10, and P11, however, expressed that they would not change the night interval frequently after that. As P10 expressed:

> *I don't know my bedtime before going to bed, and then it is a little too late to change it.*
>
> (P11)

Two participants (P2, P12) did not need any additional information apart from the notification on the lock screen. They would trust the notification itself. Two participants (P2, P13) further expressed that they would like access to a prediction during daytime even if it would be more uncertain.

6.3 Enabling Appropriate Action

One essential aspect was that the application should enable the participants to take appropriate action. All the participants agreed that a notification 30 minutes prior to their bedtime would allow sufficient time to do something to prevent a hypoglycemic event. P2, P11, P12, and P13 expressed they would eat slow-acting carbohydrates and P10 was unsure what she would do. However, P12 and P13 would still measure their BG level. P10 and P13 also expressed that they would like to get an alarm at the time of the predicted hypoglycemic event. This would allow them to check their BG level and eat food accordingly. As expressed by P11:

> *I arrange my day according to my insulin intake and then I would trust that I get an alarm if it* [hypoglycemia] *happens.*
>
> (P11)

Particularly P13 highlighted the time indication (see figure 6) and said she would use it to decide whether to eat slow-acting carbohydrates right away or set an alarm if the event was predicted to happen late during the night.

6.4 Adopting the App

Taking a step back from enabling action, the application should also be adopted and somehow accessible to the users. Therefore, it was also a talking point how the participants would become aware of such an application and if any initial guidance would be necessary. P2, P10, P11, and P13 wanted some kind of instruction manual, preferably a paper manual with guidelines. As P13 expressed:

> *I would need some small manual before I would press any buttons. Otherwise I would probably think that I wouldn't be bothered with it.*
>
> (P11)

All participants expected that they should be told about the application from either their treatment provider or through the Diabetes Association.

6.5 Usability Problems

All participants were able to navigate through the application even though some of the scenarios caused more confusion than others. All participants had difficulties finding the settings menu (see figure 8) and P2 and P10 found it challenging to change the time setting for the night. Another couple of minor problems were encountered, however, all of these usability problems were of lower severity and all of the participants expressed that they had to learn how to use the application the first time and then they would find it more intuitive.

7 DISCUSSION

In this study, we pursued a user-centered approach to enhance usability and improve the fit between technology and user as well as capture key context for intended use of the application. This aligns with previous research arguing in favor of user involvement when developing diabetes SMAs [2, 14, 21, 27]. The preliminary understanding study was designed to capture domain knowledge and to help us understand the scope of the problem through in-depth interviews with potential users. The evaluation study was designed to validate our understanding of users and their requirements as well as evaluate the proposed design and functionality of the prototype app. Users thus played an important role in the development of our prototype. Their inputs enable us to discuss implications for designing diabetes SMAs that present AI-based health predictions and to discuss barriers and enablers for adoption. Table 8 summarizes our recommendations and thus our contribution to HCI research. It builds on and extends previous research by presenting a collective overview of recommendations that consider both design and adoption of diabetes SMAs that make use of AI-based health predictions. In the following section we elaborate on these recommendations and comment on considerations for doing user-centered design in the context of diabetes SMA.

7.1 Designing Diabetes SMAs and Presenting AI-based Health Predictions

This study corroborates existing research and design guidelines that suggest using color cues [10, 15, 36] and simple visualizations [1, 10, 36] to present health data. Our findings can further be related to what Desai et al. [10] refer to as the rich-simple paradox. Throughout the preliminary understanding study, it became apparent that participants often sought simplified but rich solutions to understanding their health. Our findings support the claim of Desai et al. that the most effective designs are those that balance ease of interpretation with thoroughness. We assert that using a simple, familiar visualization combined with concise supplementary text or numbers can be one way of resolving the paradox.

Another implication we identified was the need for personalization. We found that a solution for predicting nocturnal hypoglycemia ought to accommodate individual preferences for BG threshold for hypoglycemia, sleeping patterns, and alert noise and timing. This aligns with research advocating for tailoring diabetes information to the individual's personal concerns so that they may be motivated to treat the disease seriously [24]. We further identified five main aspects which impact users' preferences for the application design and call for flexibility in the solution: 1) knowledge and experience with diabetes, 2) routines and habits in managing diabetes, 3) level of concern, 4) digital literacy, and 5) situational circumstances. These findings are consistent with research by Owen et al. [26]. In their attempt to understand requirements for take-home diabetes management technologies they arrive at the hypothesis that diabetes self-management constitutes a cycle of five critical factors: monitoring, habits, confidence, unusual situations, and concerns. They argue that these factors form decision-making in diabetes self-management. We argue that they may also impact preferences for the SMA design. In the prototype, we chose to address these differences by grouping users into two different types. These two user types required different levels of detail which we supported by presenting all users with the same simple interface and then "hiding" additional information behind button clicks and swipe-motions. Our evaluation results suggest that this approach was successful for supporting both types of users and providing flexibility.

7.2 Barriers and Enablers for Adoption

Throughout this study, a challenge became apparent – none of our target users utilized any diabetes SMA. This is in accordance with previous research done by Trawley et al. [30] highlighting that many apps simply do not reach the users. Building upon our findings, we have identified several barriers and enablers for adoption of diabetes SMA. Our recommendations for exploiting the enablers and overcoming the barriers are included in table 8.

One barrier we found was high reliance on manual input of user data. With this we echo previous research who stress the importance of using automated data gathering and data transfer [37]. We also found that the cost of required technology can act as a barrier in diabetes SMA. The technologies needed for using diabetes SMA without significant user input such as CGM sensors, smart insulin pens, insulin pumps, and activity trackers are often expensive, and without government subvention many people with diabetes do no have access to these technologies. This aligns with previous research on technology adoption in older adults suggesting that the impact of cost is often neglected [6].

Additionally, the treatment provider can be either a rather large barrier or an important enabler. We echo Årsand et al. [37] in stressing the importance of having health care professionals involved in the establishment, adoption, and use of mobile-based SMA. Without any involvement or endorsement from health care providers, unlocking the full potential of diabetes SMA will be very difficult and the lack of involvement is the largest barrier for the participants in our study. Having involvement from treatment providers or other prominent organizations such as the Diabetes Association would be a vitally important enabler as all our users reported relying fully on these parties to recommend treatment options. Along with the involvement from health care providers, the use of a guide or user manual was also an enabler among our users, as the amount of information and navigation in SMAs can often be overwhelming.

These findings draw parallels to previous research on adoption of diabetes SMAs. Zhang et al. [35] applied the Unified Theory of Acceptance and Use of Technology (UTAUT) model to identify determinants of patients' intentions to use diabetes SMAs. UTAUT, proposed by Venkatesh [31], integrates elements from eight different models for technology acceptance – including Davis' [9] foundational Technology Acceptance Model (TAM). Zhang et al. [35] found that the most important determinants were *performance expectation* and *social influence* while *facilitating conditions* and *effort expectancy* played a smaller role. Zhang et al. also added *perceived disease risk* and *perceived privacy risk* to the existing UTAUT framework. Their updated research model is depicted in figure 9. Our findings suggest that performance expectation (how may I benefit from this app?) is indeed important. We argue that this expectation is impacted by the individual's challenges and needs and the perceived role of the treatment provider. Their level of concern or perceived disease risk in the words of Zhang et al. also has an impact. However, our findings suggest that the two most impactful elements are effort expectancy and facilitating conditions. The level of digital literacy among the participants in our study was highly varied which impacted their effort expectancy. They also highlighted manual inputs as something that would impact effort expectancy negatively. Facilitating conditions such as SMA endorsement and guidance from treatment provider or the Diabetes Association were very important for the participants in this study.

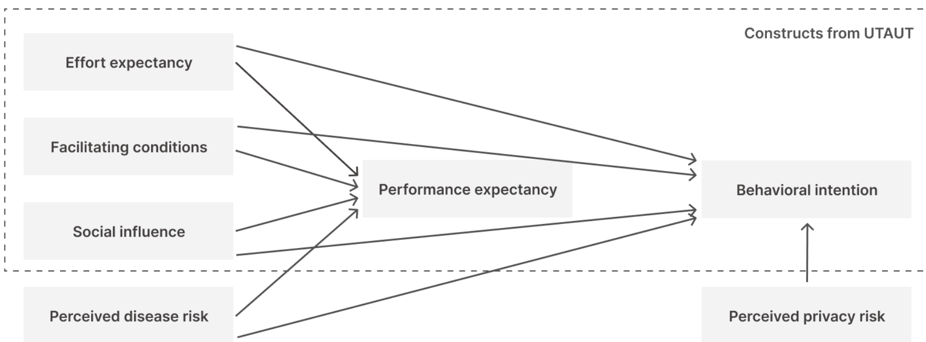

Fig. 9. Relationship between determinants and behavioral intention to use technology. From Zhang et al. [35].

7.3 Considerations for Doing User-centered Design

Throughout this study, we sought to follow suggestions from Fischer and Östlund [16] in constantly reflecting on our own role as designers and remaining impartial as the participants elaborated on their lived realities and perspectives. In the preliminary understanding study, for example, we made sure to question participants about their challenges with hypoglycemia and focusing on their suggestions for a solution before presenting our own designs. It can, however, be argued that we imposed a problem or design agenda on the participants. From the start, we already had part of a solution before even talking to potential users. The algorithm for predicting nocturnal hypoglycemia and its motivation was handed to us by SDCN. They also suggested mimicking the Glooko app to present the prediction. This initial approach therefore also contrasts with what Kanstrup et al. [21] refer to as *patient innovation*. They emphasize a participatory approach to patients as key innovators in diabetes information technology and argue in favor of patient innovations and their originality. Yet, the interviews in the understanding study suggest that the participants would not have considered the idea of predicting hypoglycemia risk using CGM data if we had not presented it to them. Many of them were frustrated with the uncertainty of going to bed with risk of hypoglycemia, but they did not think there could be an alternative to their current practices. Our findings therefore suggest that developing innovative and successful diabetes SMA's requires a balance between designing *for* and *with* users.

Another implication of doing user-centered design is that user-proposed requirements may not be technically feasible. For example, we found that participants in this study preferred being able to customize night time interval and BG threshold for hypoglycemia, and being able to see the predicted time of risk. Considering how the algorithm currently works, none of this is possible. The logistic regression model merely classifies the entire night as risk/no risk, it does not output predicted time of risk. Customizing time interval and BG threshold would further demand retraining the model to determine feature weights. With this, we reiterate previous research that argues for collaboration between health professionals, people with diabetes, and between researchers [27] – and from different disciplines we might add.

Table 8. Recommendations for creating SMAs presenting AI-based predictions for people with T2D

	Recommendation	**Description**
1	Collaboration	Include users, health professionals, and researchers from different disciplines in the design process.
2	Color cues	Use color cues to convey how prediction information should be interpreted (e.g. traffic light scheme).
3	Visualizations	Use visualizations or metaphors rather than numbers, tables, or graphs to display prediction results.
4	Simple and familiar	Use familiar and simple visualizations that the user can easily interpret.
5	Supplementary text and numbers	Consider placing supplementary text or numbers in close proximity of the visualization to convey more detail.
6	Personalization	Allow for individual customization in the interface and app/prediction model settings.
7	Flexibility	Be attentive towards different types of users. Look for differences in knowledge and experience with diabetes, routines and habits in managing diabetes, level of concern, digital literacy, as well as situational circumstances and design for flexible use.
8	"Hide" information	Do not overwhelm the user – "hide" additional information behind interaction.
9	Automated data	Avoid relying on user inputs. Exploit existing technologies to automatically gather and transfer data.
10	Required technology	Consider accessibility and cost of technology that your SMA demands to be operable (e.g. CGM, smart insulin pen, etc.).
11	Endorsement for adoption	Team up with treatment providers or prominent health organizations such and have them vouch for the SMA and commit to supporting adoption.
12	User manual	Consider creating a user manual to support adoption and sustained use.

8 LIMITATIONS AND FUTURE WORK

This study had several limitations. We had a small sample size, and the participants were not selected based on any generalization criteria, meaning our results may not be generally applicable. Additionally, a number of our participants were referred through another study at SDCN, which might indicate that many of our participants were more open towards health interventions and generally open to technology.

For this study, the primary users were insulin treated T2Ds with CGM sensors. This type of person is, however, very rare in Denmark as there is no government subvention for CGM sensors for T2Ds. We chose to ignore the CGM requirement when recruiting participants. Thus, eight of the participants had to imagine what it would be like to wear a CGM sensor and have seamless access to BG data while four participants only knew because they had temporarily worn a CGM sensor as part of another study. Further, the evaluation did not resemble a real-life setting. We did not use the participant's own BG data and we did not ask them to evaluate the prototype on their own in the evening as intended for the application. We suggest that further research on take-home self-management solutions should focus on evaluating the product in real-life settings.

As highlighted in our discussion, there are barriers for adoption of our app and diabetes SMA in general. Further research into mitigating these issues is critical for enabling the utilization of diabetes SMAs. Lastly, when working with AI in a diabetes SMA context, it is important to strike a balance between implementing user feedback, while being realistic about the feasibility of user feedback. For future research it is important to examine how a symbiosis between the development of the algorithm and implementation of user feedback can be supported.

9 CONCLUSION

In this study, we examined how users prefer to have visualizations of predicted hypoglycemic risk presented in a mobile diabetes SMA, and how different user needs impact their preferences. We found that personalization of the prediction model's definition of hypoglycemia and night is necessary. Additionally, the SMA should provide flexibility as different user preferences impact the level of information needed. We argue that these preferences are impacted by knowledge and experience with diabetes, routines and habits in managing ones diabetes, level of concern, digital literacy, and situational circumstances. We found that lack of endorsement from treatment providers acts as a barrier for adoption of diabetes SMA, and highlight the importance of collaboration between users, health professionals, and researchers. Based on our findings, we outline 12 recommendations for designing and developing diabetes SMA that present AI-based health predictions. We suggest that further research focuses on how barriers for adoption can be mitigated, and how to utilize enablers. Additionally we suggest that future work in mobile diabetes SMAs evaluate in real-life settings.

ACKNOWLEDGEMENTS

We are thankful to all participants for their time and effort and for the Diabetes Association in Aarhus and SDCN for facilitating contact. Additionally, we thank SDCN for providing the case and collaborating with us. Finally, we thank Niels van Berkel for supervising this project.

REFERENCES

[1] Jessica S. Ancker, Yalini Senathirajah, Rita Kukafka, and Justin B. Starren. 2006. Design Features of Graphs in Health Risk Communication: A Systematic Review. *Journal of the American Medical Informatics Association* 13, 6 (2006), 608–618. https://doi.org/10.1197/jamia.M2115

[2] Shaira Baptista, Steven Trawley, Frans Pouwer, Brian Oldenburg, Greg Wadley, and Jane Speight. 2019. What Do Adults with Type 2 Diabetes Want from the "perfect" App? Results from the Second Diabetes MILES: Australia (MILES-2) Study. *Diabetes Technology and Therapeutics* 21, 7 (7 2019), 393–399. https://doi.org/10.1089/DIA.2019.0086

[3] Anne Meike Boels, Rimke C. Vos, Lioe Ting Dijkhorst-Oei, and Guy E.H.M. Rutten. 2019. Effectiveness of diabetes self-management education and support via a smartphone application in insulin-treated patients with type 2 diabetes: Results of a randomized controlled trial (TRIGGER study). *BMJ Open Diabetes Research and Care* 7, 1 (2019), 1–10. https://doi.org/10.1136/bmjdrc-2019-000981

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[5] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[6] Ke Chen and Alan H.S. Chan. 2011. A review of technology acceptance by older adults. *Gerontechnology* 10, 1 (2011), 1–12. https://doi.org/10.4017/GT.2011.10.01.006.00

[7] Mike Cohn. 2004. Guidelines for Good Stories. In *User stories applied: for agile software development.* Addison-Wesley, Boston, MA, Chapter 7, 75–83.

[8] Lisa Crispin and Janet Gregory. 2009. The Purpose of Testing. In *Agile testing: a practical guide for testers and agile teams.* Addison-Wesley, Boston, MA, Chapter 6, 97–108.

[9] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease Of Use, And User Acceptance of Information Technology. *MIS Quarterly: Management Information Systems* 13, 3 (1989), 319–339. https://doi.org/10.2307/249008

[10] Pooja M. Desai, Matthew E. Levine, David J. Albers, and Lena Mamykina. 2018. Pictures worth a thousand words: Reflections on visualizing personal blood glucose forecasts for individuals with Type 2 Diabetes. *Conference on Human Factors in Computing Systems - Proceedings* CHI 2018 (April 2018), 21–26. https://doi.org/10.1145/3173574.3174112

[11] Pooja M. Desai, Elliot G. Mitchell, Maria L. Hwang, Matthew E. Levine, David J. Albers, and Lena Mamykina. 2019. Personal health oracle: Explorations of personalized predictions in diabetes self-management. *Conference on Human Factors in Computing Systems - Proceedings* CHI 2019 (2019), 1–13. https://doi.org/10.1145/3290605.3300600

[12] Diabetesforeningen. 2022. *Diabetes i Danmark.* Diabetesforeningen. https://diabetes.dk/forskning/viden-om-diabetes/diabetes-i-danmark Retrieved 25-02-2022.

[13] Mikala Dømgaard, Malene Bagger, Nicolai Alexander Rhee, Christopher Malcolm Burton, and Birger Thorsteinsson. 2015. Individual and societal consequences of hypoglycemia: A cross-sectional survey. *Postgraduate Medicine* 127, 5 (2015), 438–445. https://doi.org/10.1080/00325481.2015.1045815

[14] Omar El-Gayar, Prem Timsina, Nevine Nawar, and Wael Eid. 2013. A systematic review of IT for diabetes self-management: Are we there yet? *International Journal of Medical Informatics* 82, 8 (2013), 637–652. https://doi.org/10.1016/j.ijmedinf.2013.05.006

[15] A. Faiola, H. Vatani, and M. Agarwal. 2019. Hypoglycemic Detection by Human Breath: A Mobile Health App that Alerts Diabetics of Low Blood Glucose. *ICST Transactions on Ambient Systems* 6, 18 (2019), 162220. https://doi.org/10.4108/eai.23-3-2018.162220

[16] Björn Fischer and Britt Östlund. 2020. Technology development with older people: The role of “unfettered design”. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 12207 LNCS (2020), 18–33. https://doi.org/10.1007/978-3-030-50252-2_2

[17] Björn Fischer, Alexander Peine, Britt Östlund, and Patricia C. Heyn. 2020. The Importance of User Involvement: A Systematic Review of Involving Older Users in Technology Design. *Gerontologist* 60, 7 (10 2020), E513–E523. https://doi.org/10.1093/GERONT/GNZ163

[18] Videncenter for Diabetes. 2022. *Alder, køn og etnicitet*. Videncenter for Diabetes. https://videncenterfordiabetes.dk/viden-om-diabetes/type-2-diabetes/aarsager-og-arvelighed/alder-koen-og-etnicitet Retrieved 19-05-2022.

[19] Weixi Gu, Yuxun Zhou, Zimu Zhou, Xi Liu, Han Zou, Pei Zhang, Costas Spanos, and Lin Zhang. 2017. SugarMate: Non-intrusive Blood Glucose Monitoring with Smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–27. https://doi.org/10.1145/3130919

[20] Can Hou, Ben Carter, Jonathan Hewitt, Trevor Francisa, and Sharon Mayor. 2016. Do mobile phone applications improve glycemic control (HbA_{1c}) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. *Diabetes Care* 39, 11 (2016), 2089–2095. https://doi.org/10.2337/dc16-0346

[21] Anne Marie Kanstrup, Pernille Bertelsen, and Christian Nøhr. 2015. Patient innovation: An analysis of patients’ designs of digital technology support for everyday living with diabetes. *Health Information Management Journal* 44, 1 (2015), 12–20. https://doi.org/10.12826/18333575.2014.0013

[22] Kenneth R. Laughery and Michael S. Wogalter. 2006. Designing Effective Warnings. *Reviews of Human Factors and Ergonomics* 2, 1 (4 2006), 241–271. https://doi.org/10.1177/1557234X0600200109

[23] Martin Maritsch, Simon Föll, Vera Lehmann, Caterina Bérubé, Mathias Kraus, Stefan Feuerriegel, Tobias Kowatsch, Thomas Züger, Christoph Stettler, Elgar Fleisch, and Felix Wortmann. 2020. Towards Wearable-Based Hypoglycemia Detection and Warning in Diabetes. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA ’20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382808

[24] James Milewski and Hector Parra. 2011. Gathering Requirements for a Personal Health Management System. In *CHI ’11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI EA ’11)*. Association for Computing Machinery, New York, NY, USA, 2377–2382. https://doi.org/10.1145/1979742.1979881

[25] Omer Mujahid, Ivan Contreras, and Josep Vehi. 2021. Machine learning techniques for hypoglycemia prediction: Trends and challenges. *Sensors (Switzerland)* 21, 2 (2021), 1–21. https://doi.org/10.3390/s21020546

[26] Tom Owen, George R. Buchanan, and Harold W. Thimbleby. 2012. Understanding user requirements in take-home diabetes management technologies. In *BCS HCI*. BISL, Birmingham, 268–273.

[27] Margaret O’Neill, Catherine Houghton, Geraldine Crilly, and Maura Dowling. 2022. A qualitative evidence synthesis of users’ experience of mobile health applications in the self-management of type 2 diabetes. *Chronic Illness* 18, 1 (3 2022), 22–45. https://doi.org/10.1177/1742395320983877

[28] Roman Pichler. 2016. Strategy Development. In *Strategize: Product Strategy and Product Roadmap Practices for the Digital Age*. Pichler Consulting, Wendover.

[29] Da Tao, Juan Yuan, and Xingda Qu. 2018. Presenting self-monitoring test results for consumers: The effects of graphical formats and age. *Journal of the American Medical Informatics Association* 25, 8 (2018), 1036–1046. https://doi.org/10.1093/jamia/ocy046

[30] Steven Trawley, Shaira Baptista, Jessica L. Browne, Frans Pouwer, and Jane Speight. 2017. The Use of Mobile Applications among Adults with Type 1 and Type 2 Diabetes: Results from the Second MILES - Australia (MILES-2) Study. *Diabetes Technology and Therapeutics* 19, 12 (12 2017), 730–738. https://doi.org/10.1089/DIA.2017.0235

[31] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems* 27, 3 (2003), 425–478. https://doi.org/10.2307/30036540

[32] Martina Vettoretti, Giacomo Cappon, Andrea Facchinetti, and Giovanni Sparacino. 2020. Advanced Diabetes Management Using Artificial Intelligence and Continuous Glucose Monitoring Sensors. *Sensors* 20(14) (2020), 1–18.

[33] Peter Vistisen. 2016. What is Sketching. In *Sketching with animation: using animation to portray fictional realities aimed at becoming factual* (1st ed.). Aalborg University Press, Aalborg, Chapter 1, 11–24.

[34] Ester Yeoh and Pratik Choudhary. 2015. Technology to reduce hypoglycemia. *Journal of Diabetes Science and Technology* 9, 4 (2015), 911–916. https://doi.org/10.1177/1932296815574547

[35] Yiyu Zhang, Chaoyuan Liu, Shuoming Luo, Yuting Xie, Fang Liu, Xia Li, and Zhiguang Zhou. 2019. Factors Influencing Patients’ Intentions to Use Diabetes Management Apps Based on an Extended Unified Theory of Acceptance and Use of Technology Model: Web-Based Survey. *Journal of medical Internet research* 21, 8 (8 2019), 1–17. https://doi.org/10.2196/15023

[36] Brian J. Zikmund-Fisher, Aaron M. Scherer, Holly O. Witteman, Jacob B. Solomon, Nicole L. Exe, Beth A. Tarini, and Angela Fagerlin. 2017. Graphics help patients distinguish between urgent and non-urgent deviations in laboratory test results. *Journal of the American Medical Informatics Association* 24, 3 (2017), 520–528. https://doi.org/10.1093/jamia/ocw169

[37] Eirik Årsand, Dag Helge Frøisland, Stein Olav Skrøvseth, Taridzo Chomutare, Naoe Tatara, Gunnar Hartvigsen, and James T. Tufano. 2012. Mobile health applications to assist patients with diabetes: Lessons learned and design implications. *Journal of Diabetes Science and Technology* 6, 5 (9 2012), 1197–1206. https://doi.org/10.1177/193229681200600525

A DESIGN MOCK-UPS

We present the images in the order they were presented during the interviews with our users. First a mock-up of the Glooko app was presented.

Fig. 10. Mock-up of the Glooko app

Visualizing the prediction

Next we presented the different options for displaying whether a hypoglycemic event was predicted. The options shown here also had a "no risk" counterpart, where the color green was used instead.

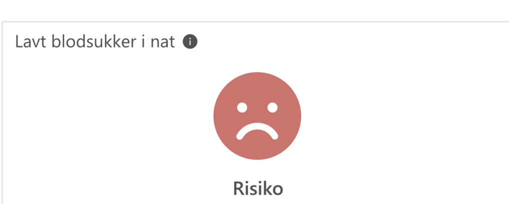

Fig. 11. Smiley

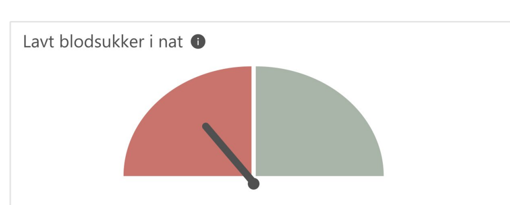

Fig. 12. Pendulum

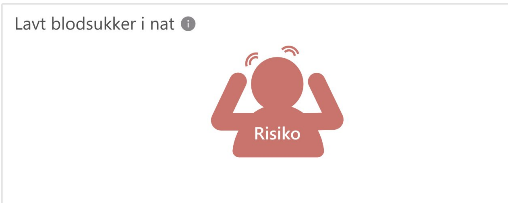

Fig. 13. Person in distress

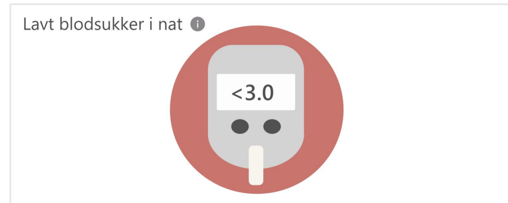

Fig. 14. BG measurement device

Additional information

After presenting the different options for displaying the prediction, we presented options for further information about the prediction. First we presented a clock, where one shows the time interval and the other shows the precise time the algorithm has predicted the hypoglycemic event will take place.

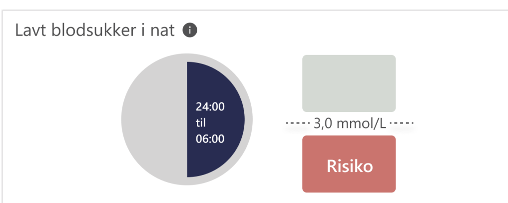

Fig. 15. Clock with interval

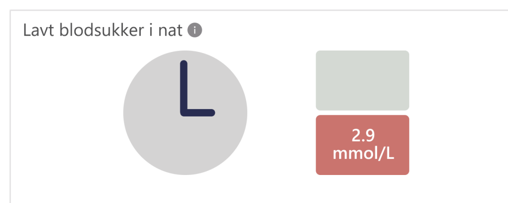

Fig. 16. Clock with precise time and BG value

The following images shows the algorithm's predicted BG level throughout the night in a graph format. The first image presents the most likely graph, while the second shows the uncertainty.

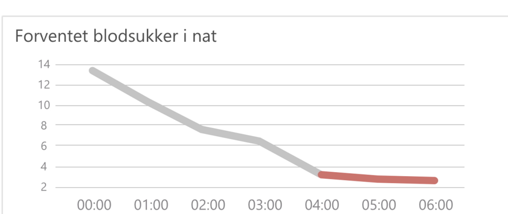

Fig. 17. Most likely BG level

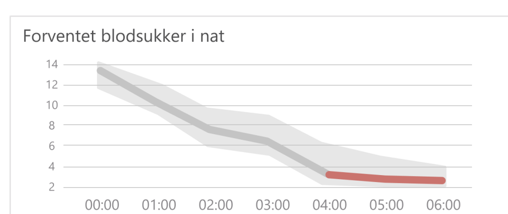

Fig. 18. BG level with uncertainty

These images show the tendency in hypoglycemic events and explanatory text accompanying the tendencies. The tendencies could be viewed on a weekly or monthly basis.

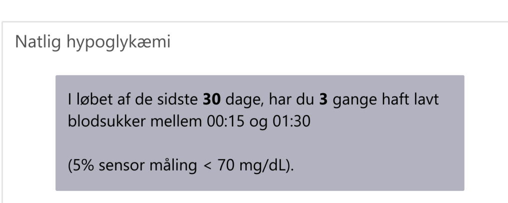

Fig. 19. Explanatory text accompanying tendencies

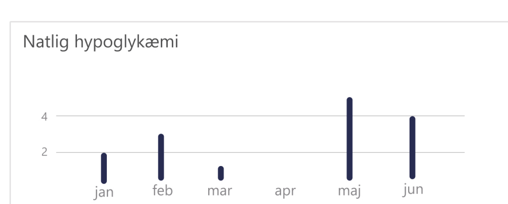

Fig. 20. Hypoglycemic tendencies

Expanding on the visualizing prediction these images show actionable instructions or text, meant to help the user understand how they should react to the algorithm predicting hypoglycemia.

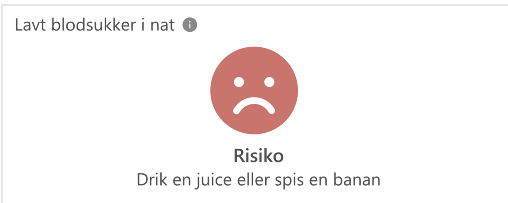

Fig. 21. Smiley with suggestion for action

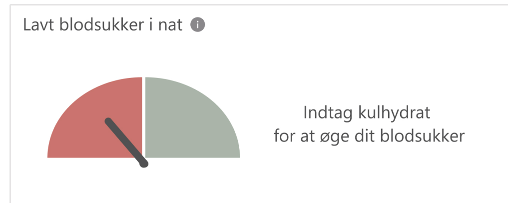

Fig. 22. Pendulum with suggestion for action

On the images for visualizing the prediction an info button can be seen next to the header text. This info button shows the user additional information about the algorithm. The extra information can be seen in the image below.

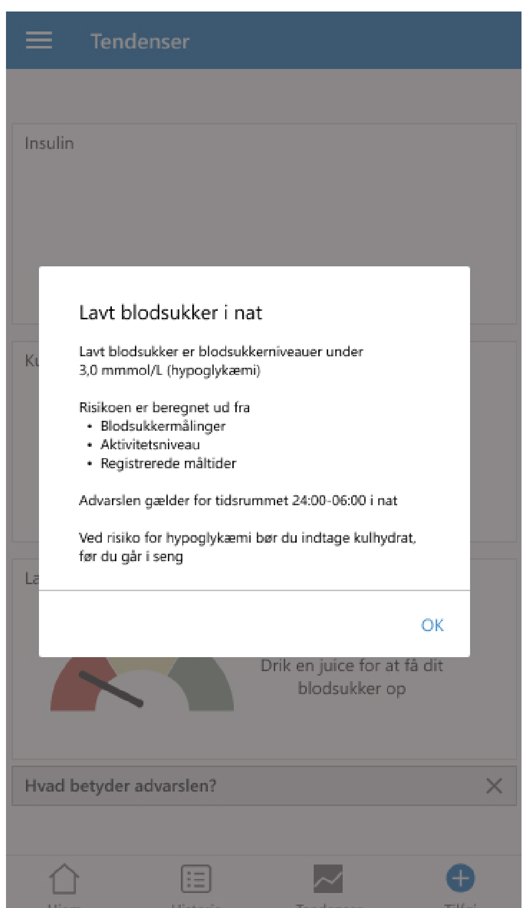

Fig. 23. Info box

Becoming aware of risk
Lastly we showed options for allowing notifications when the algorithm makes a prediction. On the images below the option menu screen can be seen along with how the notification will look on the lock screen.

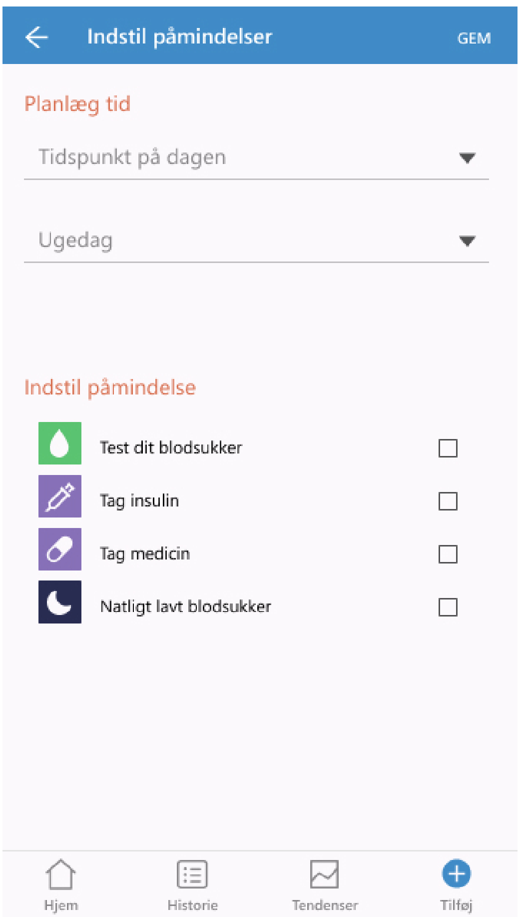

Fig. 24. Option menu with hypoglycemia notification

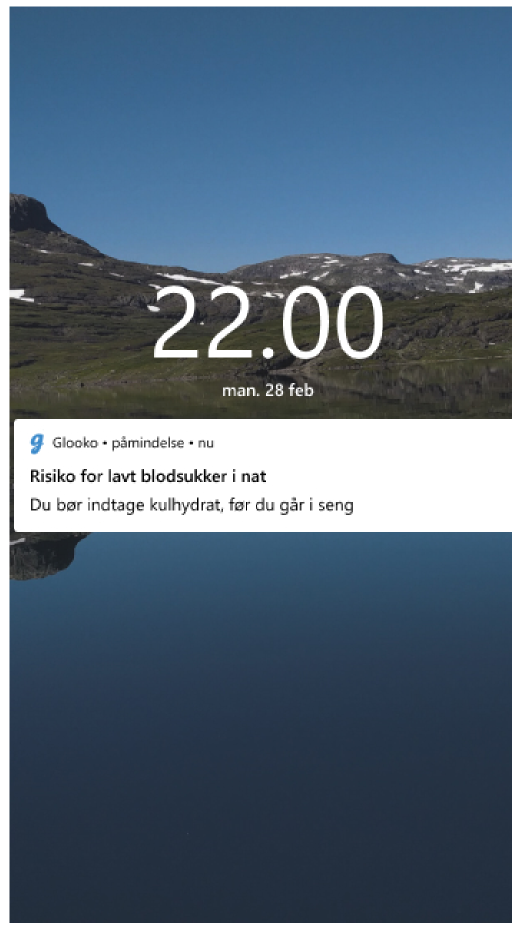

Fig. 25. Notification for risk of hypoglycemia

B PERSONAS

Fig. 26. Description of persona 1 – Lise

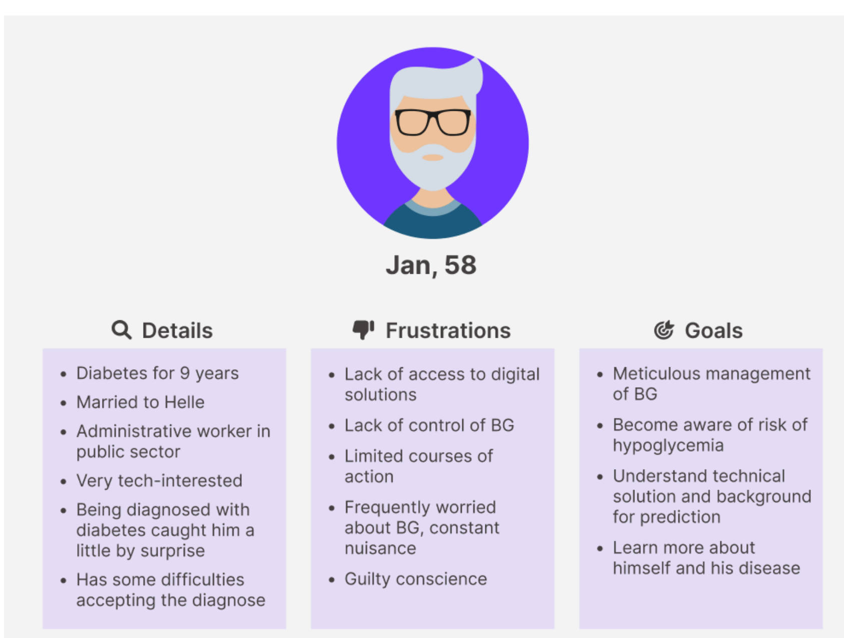

Fig. 27. Description of persona 2 – Jan

C USER STORIES

#	Story
1	As Lise, I would like to be able to see a simple visualization of the predicted risk, so that I can act on it before I go to bed.
2	As Jan, I would like to see additional information indicating the time of predicted hypoglycemia, so that I can adjust my preventive actions.
3	As Jan, I would like to have an understanding of how the algorithm works, so that I can better trust it.
4	As Lise, I would like to get an alarm with a distinctive sound in case of hypoglycemia risk, so that I am reminded to act on it.
5	As Jan, I would like to get a notification in case of hypoglycemia risk, so that I am reminded to take action before going to bed.
6	As Lise, I would like to be able to customize the app's definition of "night", so that it fits my routines.
7	As Lise, I would like to be able to customize the app's threshold value for hypoglycemia, so that it fits my experience and threshold for low blood glucose.
8	As Jan, I would like to be able to see tendencies of hypoglycemia over time, so that I can work more proactively with my diabetes.
9	As Lise, I would like to be able to customize the time of the alarm, so it fits my routines.
10	As Jan, I would like to be able to customize the time of the notification, so it fits my routines.
11	As Lise, I would like to be able to snooze the alarm, so it fits my current situation (e.g. vacation).

D SCENARIOS

#	Scenario and task
1	It is 11:30 PM on an ordinary weekday. You are about to go to bed. You know that you will receive a notification from the app 'P10' if there is risk of low blood glucose tonight. If you receive the notification, use the information to try to prevent low blood glucose.
2	Once again, it is an ordinary weekday. It is 9 PM and you are about to go to bed because you have to wake up early in the morning. Find out if there is risk of low blood glucose tonight.
3a	You have been using the app for a while, but you don't understand why it only rarely sends you a notification. You're wondering if the threshold for low blood glucose might be defined as too low in the app. Find out what the threshold for low blood glucose is defined as.
3b	Your threshold for low blood glucose is around 3.8 mmol/L. Change the app's definition to match yours.
4a	You have been using the app for a while but you are a little annoyed that whenever it sends you a notification, it is always at 11:30 PM. You usually sleep from 11 PM. Find out when the app thinks you are sleeping.
4b	Sleeping from midnight until 6 AM does not work for you. You usually sleep from 11 PM until 7 AM. Change the app's settings to match your sleeping pattern.
5a	Lately, you have been experiencing low blood glucose more often than usual and are wondering how often and when it happens. Find out how many times in the past week you have had low blood glucose levels.
5b	You have received a notification from the app several times during the past month. Yet, you feel like you have been able to prevent low blood glucose at night. Find out how many times in the past month that the app predicted low blood glucose levels, but it did not happen according to your CGM sensor.