Fighting the Hydra: Optimizing treatment for type 2 diabetes
Simon, A.C.R.

Citation for published version (APA):
Safety and usability evaluation of a web-based insulin self-titration system for patients with type 2 diabetes

Airin C R Simon, Frits Holleman, Wouter T Gude, Joost B L Hoekstra, Linda W Peute, Monique W M Jaspers, Niels Peek

Artif Intell Med. 2013 Sep;59(1):23-31
Abstract

Objective
The rising incidence of type 2 diabetes mellitus (T2DM) induces severe challenges for the health care system. Our research group developed a web-based system named PANDIT that provides T2DM patients with insulin dosing advice using state of the art clinical decision support technology. The PANDIT interface resembles a glucose diary and provides advice through pop-up messages. Diabetes nurses (DNs) also have access to the system, allowing them to intervene when needed. The objective of this study was to establish whether T2DM patients can safely use PANDIT at home. To this end, we assessed whether patients experience usability problems with a high risk of compromising patient safety when interacting with the system, and whether PANDIT’s insulin dosing advice are clinically safe.

Research design and methods
The study population consisted of patients with T2DM (aged 18-80) who used a once daily basal insulin as well as DNs from a university hospital. The usability evaluation consisted of think-aloud sessions with four patients and three DNs. Video data, audio data and verbal utterances were analyzed for usability problems encountered during PANDIT interactions. Usability problems were rated by a physician and a usability expert according to their potential impact on patient safety. The usability evaluation was followed by an implementation with a duration of four weeks. This implementation took place at the patients’ homes with ten patients to evaluate clinical safety of PANDIT advice. PANDIT advice were systematically compared with DN advice. Deviating advice were evaluated with respect to patient safety by a panel of experienced physicians, who specialized in diabetes care.

Results
We detected seventeen unique usability problems, none of which was judged to have a high risk of compromising patient safety. Most usability problems concerned the lay-out of the diary, which did not clearly indicate which data entry fields had to be entered in order to obtain an advice. 27 out of 74 (36.5%) PANDIT advice differed from those provided by DNs. However, only one of these (1.4%) was considered unsafe by the panel.
Conclusion

T2DM patients with no prior experience with the web-based self-management system were capable of consulting the system without encountering significant usability problems. Furthermore, the large majority of PANDIT advice were considered clinically safe according to the expert panel. One advice was considered unsafe. This could however easily be corrected by implementing a small modification to the system’s knowledge base.
**Introduction**

The clinical management of diabetes is already a challenge to health care professionals today and the rising incidence of type 2 diabetes mellitus (T2DM) will even put a further strain on the health care system in the future. Diabetes mellitus is associated with serious long-term complications such as blindness, renal failure and cardiovascular disease. The occurrence of these long-term complications can be significantly reduced by adequate treatment, which aims for a reduction of blood glucose levels to near normal values by diet adjustments, oral glucose-lowering medication and insulin therapy [1;2]. Treating patients with T2DM with basal insulin requires frequent evaluation of blood glucose levels and adjustment of the insulin dose (usually referred to as titration). Clinical practice and insulin trials often use a so-called “treat-to-target” design to achieve and maintain good glucose control in patients. A meta-analysis revealed that the frequency of glucose evaluations followed by possible dose adjustment was positively correlated with a reduction achieved in glycosylated haemoglobin (HbA1c), which is a widely accepted clinical marker in diabetes [3]. Therefore, we postulate that optimizing insulin use by providing sufficient titration opportunities to patients would improve glucose control in clinical practice.

A cost-effective solution to allow for more intensive treatment of patients and likewise improve clinical outcomes would be to shift the focus of diabetes treatment to self-management [4]. As internet access in patients’ homes will continue to increase over the coming years, decision support for the self-management of chronic diseases may potentially reach a large number of patients at low cost. The implementation of paper-based treat-to-target algorithms for insulin self-management has already been shown to lead to a decrease in HbA1c level of at least 1% [5]. A similar or greater decrease in HbA1c level can be expected when using an individualized algorithm delivered by a web-based clinical decision support (CDS) system, which saves the patients the effort of calculating a new insulin dose themselves. Therefore, we developed a web-based insulin self-titration system, named PANDIT (Patient Assisting Net-based Diabetes Insulin Titration), which incorporates a computerized treat-to-target algorithm to guide patients with T2DM in adjusting their basal insulin. Patients can use PANDIT at their home to supplement the regular visits with the care provider at the clinic and thereby increase insulin titration opportunities.
Related work

Drug monitoring and dosing is a well-defined area of CDS systems [6]. Most existing CDS systems in this area are developed to support physicians in their decision making to adjust drug therapy [6]. The majority of these systems focus on vitamin K antagonist dosing to be used in hospitals and community clinics in order to establish and maintain International Normalized Ratio (INR) in the therapeutic range for the prevention of thromboprophylaxis [7-9]. Also, many CDS systems for insulin therapy exist that give insulin dosing advice or support time of dosing to achieve glucose control in patients in intensive care units [10-12]. Although most of these systems have been shown to improve glycaemic control, in general the effects of CDS systems in drug monitoring and dosing on patient outcomes are heterogeneous and therefore inconclusive [6].

A number of CDS and telemedicine systems have been developed for diabetes care. Some systems assist patients and physicians in calculating the optimal pre-meal short-acting insulin dose for patients with type 1 diabetes [13-16]. Often, they were only tested on simulated data [13;16], although a few trials have shown that they are equally effective as traditional carbohydrate counting education [14;15]. Patients with T2DM initiating a basal insulin form a specific and relatively large patient group that also requires intensive titration of the insulin dose. However, T2DM patients cannot use systems targeted at patients with type 1 diabetes as adjustment of basal insulin requires a different strategy than adjustment of short-acting insulin. Most self-management CDS systems that support patients with T2DM focus on lifestyle interventions instead of the monitoring and dosing of insulin therapy [17].

State of the art

Web-based systems that assist T2DM patients in the titration of their insulin typically encompass only asynchronous telemedicine functionalities: patients enter their blood glucose readings on a daily or weekly basis and subsequently receive recommendations provided by medical professionals after the glucose readings are assessed [18]. These interventions have generally increased the time spent by professionals [19]. The addition of decision support algorithms that provide automated advice to patients may help alleviate care provider burden. Moreover, during the design of PANDIT, attention has been paid to aspects that facilitate the embedding of PANDIT in routine care. PANDIT incorporates an algorithm that decides when it is necessary for the care provider to review the patient data, allowing efficient utilization of professional care. Furthermore, care providers have their own Graphical User Interface (GUI) allowing
them to access patients records, and they also have the possibility to take over the
 provision of insulin dosing advice after reviewing the data. CDS interventions that
 fit into the workflow of the clinicians are associated with successful deployment of
decision support systems [20].

**Patient safety issues**
The use of CDS systems can improve the quality of care, but sometimes leads to
unexpected, adverse consequences for patients either because an unsafe advice is
given or because users experience problems in the interaction with the system [21].
As mentioned earlier, research has focused primarily on CDS systems that support
clinician decision making [22;23]. When CDS systems such as PANDIT are being used
by patients, safety issues become even more important because patients are less
likely to recognize incorrect system advice, compared to clinicians. With respect to
PANDIT, incorrect insulin dosing advice might induce hypoglycaemic events with
potential risks of coma or convulsions resulting in injury to the patient or others [24].
Therefore, evaluation of the safety of advice given by PANDIT is needed to prevent
adverse consequences and increase trust by stakeholders. But even when the advice is
considered safe, patients may enter incorrect data, misinterpret advice by the system,
or experience other usability problems that lead to safety issues. For this reason,
the interface of PANDIT has been designed by a multidisciplinary team consisting of
usability experts, medical informaticians, and diabetes physicians. Furthermore, we
performed an expert evaluation of PANDIT by cognitive walkthrough sessions [25].
However, typical target users of PANDIT are older people that often have minimal
experience with computers. Most of such chronic care patients are experiencing a
decline in their physical and cognitive abilities [26]; as a result, they may have more
difficulties when interacting with technology-driven solutions. This underlines the
importance of a thorough usability evaluation of PANDIT involving end-users before
conducting a pilot implementation in clinical practice.

**Objectives**
The objective of this study was to establish whether T2DM patients can safely use
PANDIT at home. To this end, we assessed whether patients experience usability
problems with a high risk of compromising patient safety when they interact with the
system, and whether PANDIT advice are considered clinically safe according to a panel
of experienced diabetes physicians. Patients and diabetes nurses (DNs) performed
a usability evaluation. This was followed by a four week pilot implementation at
patients’ homes under controlled conditions, with the goal of testing the safety of PANDIT advice. Preliminary results of the safety evaluation of PANDIT have been published previously [27].

**System description**

PANDIT consists of three different components: a clinical decision support system, a GUI and a relational database. The system architecture and patient GUI are shown in Figures 1 and 2. Because generation of insulin dosing advice is the main feature of the system, it was decided to use the Gaston framework to develop PANDIT [28]. Gaston is a state-of-the-art framework for building decision support systems, and consists of (i) an ontology-based knowledge representation language, (ii) a graphical modeling tool for encoding clinical decision rules and decision support algorithms, and (iii) an execution engine for reasoning and generation of advice. The GUI was developed using Microsoft Silverlight. Patient data are stored in a Microsoft SQL database. Upon each consultation of the system, the patient updates the diary. Subsequently, the GUI performs a store-operation on the database and sends a request to Gaston. Gaston then retrieves all relevant patient data from the database and executes the PANDIT algorithm. Finally, an advice is transmitted to the GUI, which stores the advice in the database and displays it to the user. PANDIT runs on a secured server and is password protected, with a personal account for each patient. The architecture of PANDIT meets the requirements of NEN7510, the Dutch standard for information security in health care. Secure Sockets Layer (SSL) is used to provide encrypted data exchange over the Internet.

![System architecture of PANDIT](image)

**Figure 1** System architecture of PANDIT
The patient GUI of PANDIT resembles a glucose diary in order to facilitate the collection of fasting plasma glucose (FPG) values. After a patient has logged in and opened the diary, a one-page screen is displayed containing five columns that show date, FPG values, insulin dosing advice as provided by the system, current dose of insulin used and remarks. Patients need to access PANDIT at least once every three days, in which they enter recently measured FPG values and their current insulin dose. Furthermore, they have to indicate whether they have experienced symptoms of hypoglycaemic episodes.

Figure 3 provides a flowchart model of the decision support algorithm that was implemented in Gaston. Table 1 presents the computerized treat-to-target algorithm as discussed in the introduction. In brief, the algorithm first verifies whether the patient has been using variable insulin doses, in which case (s)he is advised to contact the care provider. Otherwise, the algorithm checks if hypoglycaemic episodes have occurred, which is the case when blood glucose values referring to a hypoglycaemic episode have been entered or when the patient has explicitly indicated having experienced symptoms of hypoglycaemia. In both cases, a questionnaire is presented.
Figure 3 Flowchart model of the decision support algorithm of PANDIT

*: Table 1 describes the dose adjustments according to the lowest FPG value of the preceding three days. FPG, Fasting plasma glucose
to the patient in order to grade the severity of each episode. When a severe hypoglycaemic event has occurred, or when two non-severe hypoglycaemic events have occurred within a month, patients are advised to contact their care provider. PANDIT will then be automatically blocked from generating new insulin dosing advice and the care provider receives a message from PANDIT. In the absence of hypoglycaemia, the lowest value from three most recent glucose measurements is used to determine whether the patient’s glucose level is within the target range. If this is the case, the patient is advised to continue the current insulin dose; otherwise the system will advise to increase the insulin dose. The exact increase is based on the patient’s age and body weight.

In addition to the decision support algorithm, PANDIT also incorporates asynchronous telemedicine functionalities. These telemedicine functionalities are automatically triggered when the system is blocked from generating advice as described above, but can also be evoked by care providers when they think this is necessary. The telemedicine functionalities allow care providers to provide insulin dosing advice through the PANDIT interface to their patients. As soon as the patient is sufficiently stabilized, the care provider can decide the “unblock” the algorithm, and let PANDIT generate new insulin dosing advice again.

Table 1  Insulin titration algorithm

<table>
<thead>
<tr>
<th>FPG value</th>
<th>Insulin dose adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2.5 mmol/L</td>
<td>-0.04 IU/kg (min -4 IU)</td>
</tr>
<tr>
<td>&lt;4 mmol/L</td>
<td>-0.02 IU/kg (min -2 IU)</td>
</tr>
<tr>
<td>Target range: 4.0 – 5.5 mmol/L</td>
<td>Stable dose</td>
</tr>
<tr>
<td>5.6 – 9.9 mmol/L</td>
<td>+0.02 IU/kg (min +2 IU)</td>
</tr>
<tr>
<td>&gt;10 mmol/L</td>
<td>+0.04 IU/kg (min +4 IU)</td>
</tr>
</tbody>
</table>

* The lowest FPG of the preceding three days will be used as input.
+ The upper limit of the target FPG range can be tailored to the individual patient.
+ A patient above the age of 70 years will be up titrated with only 2 IU if FPG >10 mmol/L.

All dosing adjustments are rounded to full insulin units.
FPG, Fasting plasma glucose; IU, international units; 0.0X IU/kg (min X IU) = an adjustment of the insulin dose will be (0.0X x body weight (kg)) units with a minimum of X units.
Methods

Study population
The study population consisted of ten adult patients with T2DM who used a once daily basal insulin. The study population had a broad spectrum of HbA1c levels, ranging from normal to poor glucose control. Patients were recruited from the outpatient clinics of the Academic Medical Center in The Netherlands and several general practices in the Amsterdam region. Participation was on a voluntary basis. Informed consent was obtained prior to participation. Three DNs from the hospital provided reference insulin dosing advice with respect to the safety evaluation.

Usability evaluation of PANDIT
The usability evaluation of the PANDIT user interface was performed prior to the safety evaluation. Figure 4 provides a flowchart model of the inclusion of patients in the usability and safety study. Four patients and three DNs were asked to perform typical use scenarios while they interacted with PANDIT and verbalized their thoughts. By letting participants verbalize their thoughts, their cognitive processes and the ways they deal with problems encountered during system use are revealed. This type of “Think Aloud” (TA) testing is a direct way to gain insight into a system’s usability as perceived by its end users. We used the modified TA method of Boren and Ramey, as this method is particularly suited for usability testing of interactive computer applications [29].

Figure 4 Flowchart model of the inclusion of patients in the usability and safety study

To reduce environmental bias, the usability tests were performed at patients’ homes or at the clinic. We used a mobile testing environment for the TA sessions. The mobile environment consisted of a laptop running Windows XP with Morae (TechSmith
Corporation, Okemos, MI, USA) usability software and the web browser Internet Explorer 8. The laptop was equipped with a 17” widescreen, full-size keyboard, webcam and microphone. The usability software was used to collect video and audio recordings of the user test sessions. To test the PANDIT user interface on its usability by TA sessions, eight typical usage scenarios were developed, five for patients and three for DNs, to be completed in one TA session. TA sessions lasted approximately one hour. The supplemental file contains the description of the scenarios that were used for the TA sessions of both patients and DNs. The usage scenarios covered the subtasks in order to achieve the main goal for patients (obtain insulin dosing advice) as well as for DNs (provide insulin dosing advice). Table 2 describes the task and subtasks used for the TA usability evaluation sessions. The scenarios were presented to all participants in the same order, as each subsequent task required skills that were learned in previous tasks. The first scenario required patients to update their diary with one FPG value. In the second scenario the patient had to enter three FPG values and the occurrence of a hypoglycaemic event. In the 5th scenario the patient was asked to indicate that they had injected another insulin dose than was advised. Certain subtasks had to be repeated in all scenarios such as log in, open diary, enter FPG values and interpret advice. Patients’ performance on these repeated subtasks were used to assess persistency of usability problems. Video data, audio data, and computer screen recordings were analyzed for determining usability problems.

Each unique usability problem revealed during the TA sessions was (i) classified according to Nielsen’s severity rating and (ii) evaluated with respect to potential impact on patient safety. The severity rating of a usability problem according to Nielsen’s severity rating is a based on three factors: (i) its frequency of occurrence, (ii) its impact on task completion and (iii) its persistence when participants perform similar tasks in subsequent scenarios [30]. The potential impact on patient safety of each usability problem was evaluated by a physician and a usability expert with respect to the risk of compromising patient safety, which was classified as either “none”, “low”, or “high”. The risk was estimated by assessing whether the usability problem lead to an erroneous insulin dosing advice or a wrong interpretation of the presented advice, and whether this might cause a clinically unsafe situation. All data analyses were performed by two reviewers (ACRS and WTG).
Table 2  Task and subtasks used for the Think Aloud usability evaluation sessions/ Number of unique usability problems found per task and subtask

<table>
<thead>
<tr>
<th>Patient environment</th>
<th>Task and subtasks</th>
<th>No.</th>
<th>DN environment</th>
<th>Task and subtasks</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task and subtasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Obtain insulin dosing advice from the system</td>
<td>17</td>
<td>1. Provide a patient with insulin dosing advice</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.1) Log into PANDIT</td>
<td>2</td>
<td>(1.1) Log into PANDIT</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.2) Open your diary</td>
<td>0</td>
<td>(1.2) Open received message</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.3) Enter your FPG values</td>
<td>6</td>
<td>(1.3) Interpret message</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.4) Check the prefilled-in used insulin dose and adjust if necessary</td>
<td>1</td>
<td>(1.4) Open the patient’s diary</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.5) Indicate whether you have recently experienced a hypoglycaemic event</td>
<td>1</td>
<td>(1.5) Provide patient with insulin dosing advice</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.6) Save your entry</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.7) Answer questions concerning a recent hypoglycaemic event</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.8) Interpret (insulin) advice</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FPG, Fasting plasma glucose  

a: Indicates the existence of usability problems with potential impact on patient safety in a task or subtask;  
b: Indicates the existence of a problem with major severity according to Nielsen’s severity rating.

Safety evaluation of PANDIT

Safety of PANDIT’s insulin dosing advice was assessed in a pilot implementation with ten T2DM patients using PANDIT at their homes for four weeks. Four of the ten participating patients also participated in the usability evaluation prior to this safety evaluation. For the safety evaluation we used an adapted version of PANDIT. Normally, PANDIT is designed to provide insulin dosing advice to patients at home without interference of a DN. In this study, however, patients were blinded from the advice given by PANDIT and instead received – through the system’s user interface – insulin dosing advice given by DNs provided through the system’s asynchronous telemedicine functionalities. Figure 5 presents a schematic overview of the provision of insulin dosing advice by PANDIT and DNs. Patients were asked to measure FPG values every morning and to note whenever they experienced symptoms of hypoglycaemia. Once a patient had entered FPG measurements of three subsequent days and requested an insulin dosing advice, the DN received a notification from the system and could log in to provide this advice. As this happened asynchronously, there was usually a small time lag of a few days involved. Once DNs had entered an advice, patients received an SMS to notify them to read the advice in their PANDIT diary. DNs were instructed to
provide insulin dosing advice in concordance with daily practice. No adaptations were made to PANDIT between the usability and the safety evaluation.

After the pilot implementation, all advice provided by PANDIT were systematically compared with advice provided by DNs. In case of agreement, the advice given by PANDIT was considered to be safe. In case of disagreement, both advice were evaluated on safety by an expert panel consisting of six physicians who specialized in diabetes care. They were then presented as a clinical scenario containing the FPG values of the last seven days and the proposed insulin dosing advice from PANDIT and DNs. Two experts independently evaluated both advice on safety, taking into account the risk of hypoglycaemic events. The experts were blinded from the origin of either advice (PANDIT or DNs). Disagreements on the safety of an advice were resolved by a third expert. If both advice were safe, experts were also asked to indicate a preference for one advice, or to propose another advice (e.g., another insulin dosing advice or other clinical strategy) that would increase the clinical effectiveness of insulin therapy.

**Figure 5** Schematic overview of the provision of insulin dosing advice by PANDIT and DNs

CDSS, Clinical Decision Support system; DB, database; DN, diabetes nurse
Results

Usability evaluation of PANDIT

Patient user interface of PANDIT
Two female and two male patients with T2DM participated in the usability evaluation. Patients had an age varying from 52 to 73 years and a diabetes duration varying from two to 23 years. All patients had prior experience with the use of computers and the Internet. The TA sessions of three patients were performed at the outpatient clinic, whereas the TA session of one patient was performed at home. The TA sessions had a duration that varied from 27 to 86 min. TA usability testing with the four T2DM patients revealed a total number of 35 usability problems. The highest number of usability problems was encountered during the first scenario. Figure 6 presents the occurrence of all usability problems among the four patients for each subsequent scenario. Seventeen unique usability problems were detected. Table 2 describes the number of unique usability problems found per task and subtask. In four out of five scenarios all four participants successfully completed the main task (obtaining insulin dosing advice) without assistance. In the second scenario one participant needed assistance to obtain an insulin dosing advice.

None of the seventeen unique usability problems was rated as “catastrophic” in Nielsen’s severity rating. Two problems were rated as a “major” usability problem. These usability problems occurred while participants entered their FPG values in their diary. Most participants did not know they had to use a comma instead of a point as a decimal mark. Also, default values of 0.0 preset in the diary could not be overwritten unless patients would double-click on the value. This is an odd peculiarity in the Microsoft Silverlight GridView component upon which the diary is based. As a result all four participants experienced persistent difficulties in figuring out the right way to correctly enter the data.

Also none of the seventeen usability problems was rated as having a high risk of compromising patient safety. Five usability problems with a low safety risk were identified. Two of these problems concerned the lay-out of the diary that did not clearly indicate which data entry fields had to be filled in to retrieve an insulin dosing advice. As a result, one participant did not know that the system needed FPG values as input. Furthermore, neither participant realized that they could enter information on the amount of insulin actually used (by default, the system assumes that the previous insulin dosing advice was followed). However, this did not hamper participant in
The occurrence of all usability problems among four patients for each subsequent scenario

The persistency of usability problems in the main task (obtaining insulin dosing advice through PANDIT) was assessed using the total number of usability problems encountered by four patients for each subsequent scenario. Usability problems that were repeatedly encountered by a patient in one scenario were counted as a single occurrence. Coincidental findings refer to usability problems that patients only can encounter after coincidentally having performed certain actions. An example of a coincidental finding is that the login screen did not indicate when Caps Lock was coincidentally turned on in the login screen. If a usability problem was classified as having a risk on patient safety as well as a major usability problem, this problem was classified as having a risk on patient safety.

Figure 6 The occurrence of all usability problems among four patients for each subsequent scenario

updati ng their diary and retrieving a correct insulin dosing advice. The third and fourth problem concerned the browser pop-up messages. Browser pop-up messages are used to indicate that patients have entered an erroneous FPG value, and to present the insulin dosing advice. The small pop-up window and small font size of the pop-up text led to one participant to disregard the pop-up window altogether. Furthermore, two participants expressed the wish to retrieve the text after closing the pop-up window, which was not possible. The fifth usability problem concerned patients’ misinterpretation of a question in the hypoglycaemia questionnaire, reading “Did you measure a blood glucose value before treating the hypo?”. Again, neither of these problems stopped participants from completing the usage scenario in each question correctly.

DN user interface of PANDIT

Three female DNs with an age varying from 45 to 57 years participated in the usability evaluation. All DNs had more than five year of work experience as a DN. All DNs were familiar with the use of computers to record patient data and in their communication
with patients. The TA sessions were performed at the outpatient clinic and had a duration that varied from 29 to 34 min.

TA usability testing with the three DNs revealed a total number of thirteen usability problems. Ten unique usability problems were detected. None of these was rated as "catastrophical" according to Nielsen's severity rating or as having a high risk of compromising patient safety. Two problems were evaluated as having a low safety risk. The first one concerned the notifications provided by PANDIT to DNs in a pop-up window when patients had experienced hypoglycaemic events. The small font size of the pop-up text led two DNs to disregard the text altogether. The second problem concerned a marker about a hypoglycaemic event that was automatically placed in the patient diary, but was not noticed by one DN. However, the DN did become aware of this event when seeing the low blood glucose value in the patient’s diary.

**Safety evaluation of PANDIT**

Table 3 lists demographic and clinical characteristics of the ten patients that participated in the pilot implementation of PANDIT to assess safety of its insulin dosing recommendations. Each of them used PANDIT for four weeks. In total, 383 FPG values were entered into the PANDIT database and 74 insulin dosing advice (median 8 advice per patient) were provided by DNs. In addition, for each insulin dosing advice, a corresponding advice from PANDIT was assessed using the decision support algorithm and stored in the database. The median number of days between the previous advice and a patient’s demand for a next advice was three. The time lag between the demand for a next advice and the actual provision of advice by DNs varied from one to five days with a median of one day. Three patients experienced a total of three mild symptomatic hypoglycaemic events (all with glucose value >3.5 mmol/L) during the pilot implementation.

Twenty-seven out of 74 (36.5%) PANDIT advice deviated from the corresponding DN advice. One insulin dosing advice provided by PANDIT was considered unsafe by the expert panel. In this case the patient experienced a symptomatic hypoglycaemic event with a glucose value of 3.4 mmol/L followed by an insulin dosing advice on the same day. However, during the following four days FPG values were in the range of 6.4-8.7 mmol/L. Hereafter, the DN decided to continue the current dose, while PANDIT would have advised a dose increment of 2 insulin units based on the new FPG values of the last four days. The expert panel considered this increment to be unsafe because of the risk of recurrent hypoglycaemic events.
For most of these discrepancies the expert panel indicated to have no predominant preference for either the advice from PANDIT or the advice from the DN. For five discrepancies the advice by PANDIT was preferred. For two discrepancies the expert panel preferred the advice by the DN. The expert panel mostly favoured higher intensive insulin doses.

Although without implications for patient safety, four patients presented four unique clinical scenarios for which the expert panel proposed an alternative advice. In three unique clinical scenarios the expert panel proposed a strategy to improve the effect of basal insulin therapy on FPG values, e.g. to administer the insulin in the morning instead of the evening. For one unique scenario the expert panel recommended the performing of blood glucose measurements during the day in view of a more intensive insulin regimen, i.e. the addition of short-acting insulin. This patient reached near normal FPG values while having a disproportionally high HbA1c level, suggesting that basal insulin only would not meet the clinical needs of the patient.

Table 3  Demographic and clinical characteristics of patients participating in the safety evaluation (n=10)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female sex</td>
<td>4 (40)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>56.9 ± 8.9*</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>92.5 ± 18.7</td>
</tr>
<tr>
<td>BMI (kg/m2)</td>
<td>29.5 (26.8 - 41.3)</td>
</tr>
<tr>
<td>Diabetes duration since diagnosis (years)</td>
<td>7.0 (1 - 23)</td>
</tr>
<tr>
<td>FPG by central laboratory analysis (mmol/L)</td>
<td>8.2 (4.2-13.7)</td>
</tr>
<tr>
<td>HbA1c by central laboratory analysis (mmol/mol)</td>
<td>64.5 (37-102)</td>
</tr>
<tr>
<td>Insulin dose (IU)</td>
<td>23 (10-80)</td>
</tr>
<tr>
<td>Insulin use ≤1 year</td>
<td>6 (60)</td>
</tr>
<tr>
<td>Human insulin</td>
<td>5 (50)</td>
</tr>
</tbody>
</table>

* Age is mean ± SD  
Data are n (%) or median (min - max), unless otherwise indicated. BMI, Body Mass Index; FPG, Fasting plasma glucose; HbA1c, haemoglobin A1c; IU, international units

Discussion

To establish whether T2DM patients can safely use a web-based insulin self-titration system at home, we assessed whether patients experienced usability problems with a high risk of compromising patient safety when interacting with the system, and whether system advice were considered as clinically safe by a panel of experienced diabetes physicians.
Usability evaluation of PANDIT

Four T2DM patients and three DNs performed predefined use scenarios while they interacted with PANDIT using the TA method. The majority of unique usability problems determined during the TA usability testing were found in the patient interface. None of the usability problems in both interfaces were rated as “catastrophical” according to Nielsen’s severity rating or as having a high risk of compromising patient safety. The total number of usability problems in the patient interface decreased after participants had performed subsequent scenarios. This is most probably attributable to a learning effect.

Previous studies in the field of health information technology have shown that flaws in human-computer interaction can lead to clinical errors [21-23]. However, most previous usability research has focused on systems used by health care providers [21-23;31-34], while our study focuses on the usability of a self-management system used by patients. Only few studies have assessed the usability of patient-centered systems, mainly covering websites or telemedicine applications [35-40]. However, these studies did not assess usability problems with respect to safety risks.

We previously published the results of an evaluation of PANDIT by cognitive walkthrough sessions performed with usability experts [25]. Due to limited resources we did not redesign the system’s interface after this evaluation, but decided to take the results of both the expert evaluation and the TA sessions reported in this paper as the basis for a single redesign of PANDIT’s user interface. The cognitive walkthrough sessions detected 50% more usability problems in the patient interface when compared to the TA sessions. The problems revealed during the expert evaluation that were not detected in the TA sessions were minor usability problems or cosmetic problems that did not interfere with end-users’ work during the TA sessions. All usability problems with a potential impact on patient safety revealed in the expert evaluation sessions were also revealed in the current TA sessions. However, the TA sessions also allowed us to evaluate if a usability problem obstructed the patient from retrieving the correct insulin dosing advice, and therefore enabled us to better assess the risk of compromising patient safety. The TA sessions thus provided us with significantly more information to assess the risk of the usability problem on compromising patient safety.

Important strengths of the usability study reported in this paper are that we recruited real T2DM patients that use a once-daily basal insulin and are potential users of PANDIT, and that we performed the TA sessions with state of the art usability testing software. Many previous studies that assessed the usability of patient-centered
telemedicine systems applied only expert-based usability evaluations [37;38]. Another strength of this study was that it was designed to prospectively assess the potential impact of a usability problem on patient safety. Both a physician and a usability expert reviewed video and audio data that were recorded during the sessions to assess whether the usability problem led to erroneous insulin dosing advice or a wrong interpretation of the presented advice. Previous studies that evaluated the use of existing Computerized Physician Order Entry (CPOE) systems leading to medication errors had retrospective design and did not assess the impact of the usability flaw on patient safety [21-23].

Our study also has several limitations. First of all, we predefined the number of participants for the usability testing, while other studies have enrolled participants until saturation was attained, i.e. until no new information was uncovered [31]. Also, the assessment of persistency of usability problems in subsequent scenarios in the patients’ TA sessions was limited by design because some subtasks were only performed in a single scenario. Furthermore, scenarios that were presented at the end of the session were subjected to a detected learning effect. Although less feasible in a real world setting, from a research perspective it would have been better to minimalize the learning effect by providing training prior to the TA sessions. This might reduce the number of detected usability problems. Another limitation was that one patient needed assistance for the completion of a task, which may have biased the evaluation of persistency as in a real life setting this patient would not have been given any assistance. The number of scenarios in the evaluation of the DN user interface was too small to draw any conclusions with respect to persistency. Finally, while studies have shown that critical factors for effective design of CDS systems include integration with care provider workflow [41], we limited ourselves to the controlled conditions of the usability testing for the performance of TA sessions with DNs.

This study indicates that T2DM patients without prior experience with using PANDIT were capable of consulting the system without encountering significant usability problems. Nevertheless, we identified several usability problems with a low risk of compromising patient safety. Most of these usability problems concerned the lay-out of the diary, which did not clearly indicate which data entry fields had to be filled in order to obtain an advice. Patients using PANDIT were generally ignorant with respect to the information needed by the decision support algorithm to calculate an insulin dosing advice. This points us at one important recommendation to take into consideration when designing user interfaces for CDS systems directed at chronic
care patients. Patients, other than care providers, generally have a poor conceptual understanding of the information that is necessary to perform a specific clinical decision-making task. Therefore, transparency should be created in the patient interface with respect to the way a CDS system performs a specific decision support task, e.g. by using affordances to support users in building productive mental models.

For future research we recommend the identification of subpopulations of patients that need less or more guidance with regard to the use of web-based self-management interventions.

**Safety evaluation of PANDIT**

Ten T2DM patients used PANDIT at their home for four weeks under controlled conditions.

Out of 74 PANDIT advice, one insulin dosing advice provided by PANDIT was considered unsafe by the expert panel: in case of a recent hypoglycaemic event, rapid increments of the insulin dose after a few days might lead to recurrent hypoglycaemic events. Normally PANDIT would take into account only a small interval of days to decide on the next insulin dosing advice. However, we found that the occurrence of a hypoglycaemic event requires the evaluation of a longer number of days to decide on the optimal insulin dosing advice. Furthermore, for some clinical scenarios the expert panel preferred an alternative clinical advice, which encompassed a clinical strategy to improve the effect of the basal insulin or a recommendation to evaluate whether the patient should switch to another insulin regimen. Although having no impact on patient safety, disregard of these alternative clinical options might impede optimal care.

Two randomized controlled trials have investigated the safety of web-based systems in calculating the optimal pre-meal short-acting insulin dose for patients with type 1 diabetes [14;15]. They showed that the use of these web-based systems did not increase the risk of developing hypoglycaemic events. However, these studies did not provide insight into the causes of hypoglycaemic events. To our knowledge, no previous study has assessed safety of CDS systems targeted at patients by performing a process analysis that systematically evaluates system advice.

The main strength of the safety evaluation of PANDIT was the creation of a real world setting including ten patients with T2DM under the supervision of three DNs. The study population consisted of participants with a broad spectrum of age, sex and HbA1c levels, which were recruited from both general practices and academic hospitals, thus creating many possible scenarios to be tested on its safety. Furthermore,
we had access to the expertise of an extensive expert panel allowing an evaluation of all detected discrepancies between PANDIT and DNs advice and allowing us to obtain additional information on preferences and recommendations. Another strength of the safety evaluation was the controlled setting; arguments for each insulin dosing advice were recorded by DNs, which allowed for an easy recapitulation of the clinical scenarios to the expert panel afterwards. A limitation of the study was that actual use of PANDIT with unblinded advice might generate different clinical results.

The safety evaluation indicates that the majority of PANDIT advice was considered clinically safe by experienced diabetes physicians. However, the system could not be considered completely safe due to one flaw in the knowledge base. Similar advice can however easily be avoided in the future by implementing a small modification to the system’s knowledge base; if the patient experiences blood glucose values in or below target, the interval of days needed to decide on the next dosing advice should be increased. The clinical scenarios for which the expert panel preferred an alternative clinical advice occurred at the boundary of the system’s domain of expertise and exceeded the scope of the current knowledge base. This problem can be resolved by adding additional decision rules in the knowledge base that determine when the patient should be redirected to the professional.

The result of the safety evaluation implies that CDS systems can safely support patients in self-adjusting their drug therapy according to the treat-to-target strategy. This is valuable information as the treat-to-target of drug therapy implemented in the current knowledge base represents an indispensable approach to the management of some highly prevalent diseases [42]. However, to allow for graceful degradation outside the domain of expertise, attention should be given to these scenarios that require professional care.

**Conclusion**

PANDIT is a web-based self-management system using state of the art clinical decision support and telemedicine technology to provide T2DM patients with insulin dosing advice. T2DM patients without prior experience with this web-based system were capable of consulting the system without encountering significant usability problems. Furthermore, the large majority of system advice was considered clinically safe by experienced physicians specialising in diabetes care. One advice was considered unsafe. This would however easily be remedied by implementing a small modification to the system’s knowledge base. The results of the study imply that patients with
T2DM can safely interact with computer-based self-care systems, and CDS systems can safely support patients in self-adjusting their drug therapy. However, as patients, other than care providers, cannot critically reflect on all system’s advice, specific attention should be given to scenarios that occur at the boundary of a system’s domain of expertise in order to ensure that care responsibilities are timely handed back to professionals.
Supplemental file

Think Aloud – Scenarios for patients

Scenario 1:
Context:
• You have been using PANDIT for one week.
Information:
• Yesterday (March 28th), you have received your last insulin advice from PANDIT: 18 IU.
• Accordingly you administered 18 IU last evening (March 28th).
• Today, you have measured a fasting blood glucose value:
  • March 29th: 7.8 mmol/L

1. Update your diary.
2. What would you do in response to the text in the pop-up screen?

Scenario 2:
Context:
• You have been using PANDIT for a month.
Information:
• You have received your last insulin advice from PANDIT four days ago (March 26th): 24 IU.
• Accordingly you administered 24 IU each night.
• You have measured the following fasting blood glucose values:
  • March 27th: 6.7 mmol/L
  • March 28th: 7.1 mmol/L
  • March 29th: 6.9 mmol/L
• You have recently moved house to Hoefsmidhof 28, 1446 AV, Purmerend.

1. Update your diary.
2. What would you do in response to the text in the pop-up screen?
3. Change your address information.

Scenario 3:
Context:
• You have been using PANDIT for two months.
Information:

- You have received your last insulin advice from PANDIT the day before yesterday (March 27th): 32 IU.
- Accordingly you administered 32 IU each night.
- You have measured the following fasting blood glucose values:
  - March 28th: 4,5 mmol/L
  - March 29th: 3,0 mmol/L
- Despite the low fasting blood glucose value on March 29th, you did not experience symptoms of a hypoglycemia. You were alone at home and took two dextrose tablets.

1. Update your diary.
2. What would you do in response to the text in the pop-up screen?

Scenario 4:

Context:

- You have been using PANDIT for three months.

Information:

- You have received your last insulin advice from PANDIT three days ago (March 26th): 46 IU.
- Accordingly you administered 46 IU each night.
- You have measured the following fasting blood glucose values:
  - March 27th: 5,7 mmol/L
  - March 28th: 5,1 mmol/L
  - March 29th: 5,9 mmol/L
- On March 28th you experienced symptoms of hypoglycemia (sweating, shaking) before having dinner. Your measured blood glucose value at that moment was 2,3 mmol/L.
- You were alone at home and drank lemonade. The symptoms quickly disappeared.
- From now on you wish to receive PANDIT reminders by text-message every night at 9 PM.

1. Update your diary.
2. What would you do in response to the text in the pop-up screen?
3. Update your preferences for receiving reminders.
Scenario 5:
Context:
• You have been using PANDIT for one month.
Information:
• You have received your last insulin advice from PANDIT four days ago (March 26th): 24 IU.
• Accordingly you administered 24 IU each night, but not on March 28th: on that day you accidentally administered 18 IU.
• You have measured the following fasting blood glucose values:
  • March 27th: 6,7 mmol/L
  • March 28th: 7,1 mmol/L
  • March 29th: 6,9 mmol/L
• You would like to contact your diabetes nurse but you can’t reach her by phone.

1. Update your diary.
2. What would you do in response to the text in the pop-up screen?
3. Send your diabetes nurse an e-mail in which you tell her that you weren’t able to reach her.

Think Aloud – Scenarios for diabetes nurses

Scenario 1:
Information:
• You have received a new message stating one of your patients needs a new insulin advice.

1. Open your new message.
2. Provide the patient with a new (random) insulin advice.

Scenario 2:
Information:
• You have received a new message stating one of your patients needs a new insulin advice.
• You decide that the upper limit of the target value for this patient should be changed to 6.0 mmol/L.

1. Open your new message.
2. Provide the patient with a new (random) insulin advice.
3. Change the target value for this patient.
Scenario 3:
Information:
- You have received a new message stating one of your patients needs new insulin advice.
- You want to motivate the patient to enter his fasting blood glucose values into his diary every five days.
- Also, you decide to take over the provision of insulin advice and block PANDIT in providing advice.

1. Open your new message.
2. Provide the patient with a new (random) insulin advice.
3. Adjust the visiting profile of this patient.
4. Block PANDIT in providing insulin advice.
Reference List


40. Ostergren MJ, Karras BT. ActiveOptions: leveraging existing knowledge and usability testing to develop a physical activity program website for older adults. AMIA Annu Symp Proc 2007;578-582.
