A web application for follow-up of results from a mobile device test battery for Parkinson’s disease patients

Mevludin Memedi\textsuperscript{a,b,*}, Jerker Westin\textsuperscript{a}, Dag Nyholm\textsuperscript{c}, Mark Dougherty\textsuperscript{a}, Torgny Groth\textsuperscript{d}

\textsuperscript{a} Department of Economy and Society, Computer Engineering, Dalarna University, Borlänge, Sweden
\textsuperscript{b} Academy of Science and Technology, Örebro University, Örebro, Sweden
\textsuperscript{c} Department of Neuroscience, Neurology, Uppsala University, Uppsala, Sweden
\textsuperscript{d} Department of Medical Sciences, Biomedical Informatics and Engineering, Uppsala University, Uppsala, Sweden

A R T I C L E   I N F O

Article history:
Received 13 April 2011
Received in revised form
26 July 2011
Accepted 27 July 2011

Keywords:
Parkinson’s disease
Test battery
Web application
Decision support
Remote patient monitoring
Telemedicine
Principal component analysis

A B S T R A C T

This paper describes a web-based system for enabling remote monitoring of patients with Parkinson’s disease (PD) and supporting clinicians in treating their patients. The system consists of a patient node for subjective and objective data collection based on a handheld computer, a service node for data storage and processing, and a web application for data presentation. Using statistical and machine learning methods, time series of raw data are summarized into scores for conceptual symptom dimensions and an “overall test score” providing a comprehensive profile of patient’s health during a test period of about one week. The handheld unit was used quarterly or biannually by 65 patients with advanced PD for up to four years at nine clinics in Sweden. The IBM Computer System Usability Questionnaire was administered to assess nurses’ satisfaction with the web application. Results showed that a majority of the nurses were quite satisfied with the usability although a sizeable minority were not. Our findings support that this system can become an efficient tool to easily access relevant symptom information from the home environment of PD patients.

\textcopyright 2011 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Parkinson’s disease (PD) is a progressive neurological disorder, which results from a degeneration of dopamine-producing nerve cells in the region of the brain known as substantia nigra that controls movements. The cardinal symptoms of PD are slowness of movements, stiffness, instability, and tremor. Additionally, non-motor symptoms are common. Standard treatment is levodopa which works as dopamine replacement and helps in relieving the symptoms and prolonging life expectancy. In the early stages of the disease, levodopa works well in relieving the symptoms and improving patient’s motor function. However, with the disease progression, patients start to experience motor fluctuations, their motor condition fluctuates between the “off” state (as a result of insufficient medication) and the “on” state (in which levodopa levels are enough for the patient to respond as a non-parkinsonian person). In addition to these two motor conditions, patients in the “on” state may develop involuntary movements, also known as drug-induced dyskinesias, in response to peak levels of medication. In the presence of motor and non-motor fluctuations, single or few observations are not enough to give full information on a patient’s condition and may also result in bias in measuring treatment effects [1].
Most of the existing approaches to remote monitoring of PD symptoms have been based on the use of wearable sensors [2–6]. Other approaches to test motor function via keyboards [7] and custom hardware testing devices [8] have also been tried. Typically, these approaches do not address the important issue of defining scores that combine different aspects of patient function in order to facilitate monitoring of the patients. To date, repeated patient assessments are usually summarized individually as percentages of responses above some levels (e.g. percentage self assessed good “on”-time, that is “on” time without troublesome dyskinesias) or as mean values of test results.

Our research team has developed a test battery implemented on a Personal Digital Assistant (PDA) with touch screen and built-in mobile communication to be used by patients at home as a telemedicine approach [9]. The tests consist of self-assessments of symptoms and fine motor function tests (tapping and spiral drawing). Assessments are performed several times per day in the patient’s home during test periods of approximately one week in duration.

In this paper, we present an integrated system for delivering assessment support information to the treating clinical staff for monitoring PD symptoms in their patients and assisting decision making concerning treatments. The system compiles the test battery raw data, concerning subjective and objective measurements, into scores for conceptual symptom dimensions and an overall test score reflecting the overall condition of the patient during a test period and presents them graphically in a web application.

2. Computational methods and theory

2.1. Datasets used in system development

The test battery handheld unit has been used by 65 patients with advanced PD (treated with intraduodenal levodopa/carbidopa gel infusion, Duodopa®, or candidates for this treatment) at nine clinics around Sweden in the clinical study DAPHNE (Duodopa in Advanced Parkinson’s: Health Outcomes & Net Economic Impact, Eudract No. 2005-002654-21). A detailed description of the test battery can be found elsewhere [9], however an outline of it is given here. The patients entered responses to seven self assessment questions (q1–q7) and performed fine motor tests four times per day during one to ten test periods of seven days’ length. Most of the diary questions (q1, “Ability to Walk”, q3, “Off at worst”, q4, “Dyskinetic at worst”, q5, “Painful cramps”, and q6, “Satisfied with functioning”) relate to the last 4h/this morning and are of verbal descriptive scale type with answer alternatives ranging from 1 (worst) to 5 (best). There is another question (q2) in which patients are asked to mark the proportion of time they spent in “off”, “on” and “dyskinetic” states during the last 4h/this morning. In addition, patients were asked (q7) to mark the right now condition in one of seven categories of motor conditions ranging from “very off” to “on” to “very dyskinetic”. In addition patients performed different tapping tests (with and without visual cueing) and traced a pre-drawn Archimedean spiral three times per test occasion. In total there were 379 test periods and 10,439 test occasions. For 223 of these test periods, Hoehn and Yahr [10] and Unified Parkinson’s Disease Rating Scale (UPDRS [11]) ratings were performed in afternoons at the start of the week. Baseline characteristics are shown in Table 1.

2.2. Calculation of test battery dimensions

Some of the test battery items are highly correlated indicating that they measure the same concept. Because of this redundancy, it is possible to reduce these items into a smaller number of principal components without much loss of information for representing conceptual symptom dimensions. This can be achieved with a linear dimension reduction technique, in terms of least mean squared error, known as Principal Component Analysis (PCA) where the first principal component is the linear combination of the original variables, giving the largest variance in the data [12]. According to the information content of the test battery, a test period can be described by six dimensions: Walking (based on q1), Satisfaction (q6), Dyskinesia (q2-dyskinetic and q4), Off (q2-off and q3), Tapping (all tapping tests) and Spiral (all spirals).

Initially, series times of test battery items were summarized by calculating statistical features, such as the level (MEAN), fluctuation (standard deviation, SD) and the mean squared deviation (MSD) from “the best” answer alternative for diary questions q1, q3, q4 and q6, on a test period level. For instance, MSD for q1 was calculated with the following formula:

\[
MSD = \frac{1}{n} \sum_{i=1}^{n} (5 - x_i)^2
\]

where \( n \) is the total number of test occasions in a test period, 5 is the best answer alternative meaning “Walking without a problem” and \( x_i \) is the patient’s test response. For each patient and test period we obtained a total of 28 features. Table 2 shows which features belong to which dimension.

The rationale for selection of these statistical features was to define scores taking into account the intensity, frequency and importance of occurring symptoms. Mean values are the obvious choice to represent levels and standard deviations are obvious for representing overall variation. In order to determine how much of the level and how much of the variation to include, PCA was applied. To have at least three features to base PCA on, MSD was also used. This measure combines both level and variation.

For each dimension, we performed PCA using the correlation matrix method applied to the statistical features for the diary questions or motor tests that the dimension was based on by retaining the first principal components. The motivation for using correlations instead of covariances was based on the

| Patients (n, gender) | 65 (43 m; 22 f) |
| Age (years) | 64.7 ± 7.4 |
| Years with levodopa | 13.5 ± 5.8 |
| Hoehn and Yahr | 2.8 ± 0.8 |
| UPDRS-Section 2 (activities of daily living) | 13.6 ± 6.2 |
| UPDRS-Section 3 (motor) | 22.3 ± 10 |
| Total UPDRS | 45.8 ± 16.7 |

Table 1 – Clinical features at baseline (mean ± standard deviation).
the corresponding dimensions and percent variation involved in each of them. Many criteria are used for determining the adequate number of principal components to be retained and used in model development. A popular approach is to select a cumulative percentage of total variation to which it is desired that the selected components should contribute more than 70% [12]. In this work, the justification of retaining only the first principal components for each dimension was based on the fact that we needed a single value representing them.

Linear transformations were used to scale dimensions to a scale from 0, worst, to 1, best score based on their minimum and maximum values. To evaluate whether dimensions measure the same construct of patient’s health status, Cronbach’s $\alpha$ for the six dimensions was calculated. The test showed that there was good internal consistency between them with Cronbach’s $\alpha$ of 0.81.

### 2.3. Calculation of overall test score

Using the six derived dimensions, the overall score can be calculated by using a regular hexagon where each dimension becomes a corner on it. According to domain experts the geometrical placement of dimensions in the hexagon should look like in Fig. 5. Using the side-angle-side method the areas of each triangle of the hexagon can be calculated separately and then added up to form an unweighted overall score.

A problem with this approach is that all dimensions, concerning self-assessments and motor tests, have the same weight in the assessment of the overall condition of the patient which is in contradiction with the common clinical rating scales, e.g. UPDRS, used in clinical practice where the highest weight is given to the evaluation of motor symptoms.

To overcome this problem, an alternative way to define the overall score was to weight symptom dimensions using the UPDRS, since it is the most widely used scale for assessing PD today [13]. The standard least-squares linear regression was used for examining the relationship between first components of dimensions and the patient’s UPDRS score.

---

**Table 2 – Percentage total variation accounted for by first principal components of the six dimensions and their feature weights. MSD is mean squared deviation from the best answer alternative. NA = not available, i.e. best score is undefined.**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Test battery-item (question/motor test)</th>
<th>Mean (%)</th>
<th>SD (%)</th>
<th>MSD (%)</th>
<th>Total variation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>Q1 (ability to walk)</td>
<td>39.4</td>
<td>21.6</td>
<td>39.1</td>
<td>71</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>Q6 (pleased with function)</td>
<td>35.7</td>
<td>27.3</td>
<td>36.9</td>
<td>77</td>
</tr>
<tr>
<td>Dyskinesic</td>
<td>Q2 (portion of time spent in dyskinesic)</td>
<td>20.3</td>
<td>18</td>
<td>NA</td>
<td>21.9</td>
</tr>
<tr>
<td>Q4 (dyskinesic at worst)</td>
<td>22.1</td>
<td>17.6</td>
<td>NA</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Spiral</td>
<td>All spirals (mean “spiral score”)</td>
<td>50</td>
<td>50</td>
<td>NA</td>
<td>51</td>
</tr>
<tr>
<td>Tapping</td>
<td>Tapping (alternately)-both hands (mean speed)</td>
<td>8.6</td>
<td>1.1</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tapping (alternately)-both hands (rate of correct taps)</td>
<td>9.4</td>
<td>7.5</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tapping (increasing speed)-dominant hand (rate of correct taps)</td>
<td>12.8</td>
<td>11.6</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tapping (random chasing)-dominant hand (mean speed)</td>
<td>11.7</td>
<td>11.1</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tapping (random chasing)-dominant hand (rate of correct taps)</td>
<td>12.7</td>
<td>13.4</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2 (portion of time spent in “off”)</td>
<td>21.4</td>
<td>17</td>
<td>NA</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Q3 (”off” at worst)</td>
<td>23.3</td>
<td>14.7</td>
<td>23.7</td>
<td></td>
</tr>
</tbody>
</table>

---

**Fig. 1 – The three principal component axes plotted with the original features (MEAN, SD and MSD) for Walking dimension. The blue line represents the first principal component and accounts for 71% of the data variability, the red represents the second and the green represents the third principal component. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)**
Fig. 2 – The overall system architecture.

Fig. 3 – Navigation diagram for the WA.

Fig. 4 – Patient status report in WA with graphical visualization of time series of test scores.
UPDRS weighted overall score (UPW) was defined as a linear combination of the first principal components, with numerical weights estimated by regression technique to fit simultaneous clinical ratings on the UPDRS scale. This was done on the basis of the following equation:

\[
\text{UPW} = 0.0265 \times \text{Walking} + 0.0933 \times \text{Satisfaction} + 0.0513 \\
\times \text{Dyskinetic} - 0.0007 \times \text{off} + 0.119 \\
\times \text{Tapping} + 0.203 \times \text{Spiral}
\]

(3)

The UPW was then scaled to a scale from 0, worst, to 1, best score using linear transformations based on its maximum and minimum values. The dataset used for optimizing this statistical procedure was collected from DAPHNE study. From Eq. (3), motor test results, tapping tests and spirals, were given high weight (65%), just as the motor section (III) has the highest weight in the UPDRS.

Correlations between UPW and other clinical rating scores were assessed by Spearman rank correlations. The correlations between UPW and total UPDRS (−0.6, \( p < 0.001 \)) were adequate whereas between UPW and Hoehn and Yahr (−0.44, \( p < 0.001 \)) were medium.

3. System description

Decision support systems comprise components for (i) database management capabilities with access to internal and external data, information, and knowledge, (ii) modelling functions accessed by a model management system, and (iii) user interface designs that enable interactive queries, reporting, and graphing functions (cf. e.g. [14]). In our case, these processes are handled separately where data collection, storage and processing takes place before presentation. The former two processes are handled by the so-called data processing sub system and the later one is handled by the web application.

3.1. Data processing sub system

For each test occasion, raw test data are sent from the PDA hand unit over the mobile net to the so-called remote device manager (RDM). The RDM is a commercial product responsible for collecting and storing the data from the hand unit. There is a communication protocol between these two systems which handles the transfer of the data. The system architecture is illustrated in Fig. 2.

The data processing sub system (DPSS) is a stand-alone application, which incorporates knowledge to analyze and interpret the raw test battery data. It parses, processes and stores the data into relational database tables and at the same time calculates and stores test scores on test period level. A connection with RDM is first established following by receiving and parsing XML data from files. Once files are received, they can be directly interpreted by the DPSS which runs during a specified time interval, e.g. every single hour. The data messages consist of patient identification, test period, hand unit identification, starting and ending time of the test occasion and responses to the test battery items. The main reason for processing data centrally instead of locally is the risk of losing raw data in the remote devices if it is not uploaded regularly. Access to raw data centrally also facilitates future research and method developments. The data collection in the hand units is designed to minimize upload bandwidth [9].
[15], using Discrete Wavelet Transform and PCA, a single piece of M-code was written in Matlab\textsuperscript{a} (MathWorks Inc.). An error handler was designed so the application can recover from possible run-time errors without terminating by rolling back all the information and saving the error information to a log file.

3.2. Web application

The web application (WA) is a feedback system comprising a secure web server and a database with web-based access for medical staff. The main role of the WA in the overall system is to present test results to the end-user clients. It is designed in line with the users’ requirements and needs for decision support by displaying the test results per patient and test period in graphical and tabular format. It consists of different web pages and the navigation between them, as illustrated in Fig. 3.

Some important user requirements specifications for the WA were: easy to understand and use, user-friendly design, easy navigation, fast response, rapid and convenient screening of patients, a comprehensive overview of a patient’s condition on a single page. They were drafted in collaboration with experts and different prototype versions were developed in several iterations.

Data security is assured at three levels: forms-authentication level; web-server level; and database-server level. Forms authentication is an authentication model in which users, in our case the treating clinical staff and administrators, access the WA by providing their unique usernames and passwords in a web page form. Depending on the user credentials and their access permissions, the application limits what resources or functionalities are accessible (e.g. it enables access only to data belonging to a specific clinic). On a web-server level, the Integrated Windows Authentication mechanism is used for connecting clients to the server. It uses a hash algorithm for sending credentials before sending them across the network. On a database-level, database administrators have their Active Directory accounts where they have full access to the data and can manage users, groups and passwords, whereas the users of WA that are not part of this domain can access the data by a SQL-specific account, also known as SQL login, incorporated in the configuration file of the WA.

4. Samples of typical system runs

To enable rapid patient status assessment, the information in the WA is displayed and ordered using a top-down approach where the general overview of the patient’s performance per test period is given. Raw data and other detailed information may be accessed in more “advanced” displays. After the user logs in and selects a relevant patient, the main page displays the graphs of the patient scores on a test period level, focusing on the overall test score, dimensions and daily summaries as shown in Fig. 4. The graphs can be updated interchangeably. The horizontal axis on the overall test score graph displays the test periods and the vertical axis displays the score of the patient on a scale ranging from 0 to 1, where zero is considered as the worst score and one is regarded as the best.

Along with overall test score the user has the option to add to the graph dimensions’ scores by using the respective checkboxes. In addition to providing the overall score graph, a table is shown with the completed test periods and their respective number of test occasions where test periods with low compliance, e.g. less than 15 test occasions, are highlighted with a special colour. In the lower left side of the web page, two more graphs are generated to represent dimensions of symptom severity arranged in regular hexagons for two test periods, i.e. for the first and the last one. Next, another graph shows the trend of statistical summaries, mean and standard deviation, for test responses over the time-of-day sessions (8–, 12–, 16– and 20–o’clock) for a selected test period.

In a separate web page, the user can observe the trend of patient’s performance on the six symptom dimensions over all test periods as shown in Fig. 5.

In another web page, a table of statistical summaries is presented for test battery items and their derived quantities, dimensions and overall score. The table also shows the summaries for the DAPHNE reference group from the existing database to allow users to evaluate and compare individual patient’s state with the reference group.

Test responses for different test battery items over a test period are shown in time series graphs where the user can subsequently alter the diary question, motor test, or test period by using dropdown menus. The user can click on any designated point over the time series plot to show a popup window displaying the three spiral drawings along with their respective drawing completion times in seconds and the patient’s self-assessment of the motor state at the time of the particular test occasion.

4.1. System’s usability evaluations

An advisory board consisting of 14 neurologists evaluated the system after presentation and demonstration of its functionality. The neurologists were based in the following countries: USA 3, Germany 3, Italy 2, Spain 1, Netherlands 1, UK 1, Sweden 1, Denmark 1 and Finland 1. They responded to questions by pressing keypads. To a question about their overall impression of the system, eleven had a positive impression, one had a neutral impression and two had negative impression. The most important benefits they could see were an increased ability to identify patients who are not doing well and facilitated follow-up optimization of an individual’s treatment. The system was seen as most important for complicated patients and for regional patients, that is patients living in regions far away from a clinic. The general conclusion of the board was that the system was recognized as a tool that will assist in management of patients.

The WA was demonstrated to fifteen nurses from the nine clinics in Sweden who had experience using the mobile test battery device in the DAPHNE study, but they had not previously seen the WA. The evaluation was performed as a presentation session for the nurses where they asked the presenter to display certain data from a particular patient in whom they had an interest. The presenter took notes about the reactions from the group. At least one patient per clinic was shown. The responses were summarized in a qualitative manner as follows: (i) the WA is very useful, (ii) the results
during test periods showed agreement with qualitative observations of the patient during that test period, for example, “one patient was in a bad condition in baseline, he improved after starting Duodopa, then he became worse again, 24-h infusion started and the patient became better again; we can clearly follow this in the WA”, (iii) comparisons between patients are possible (one patient is in better/worse condition than another).

One year later, the IBM Computer System Usability Questionnaire (CSUQ) was administered to evaluate the nurses’ satisfaction with the WA. They were asked to perform a series of tasks using the WA, such as (1) login, (2) select a patient, (3) check the patient’s performance by looking at the graphs of summarized scores and compare these results with own clinical observations, (4) select and check the other patients’ results repeating steps (2) and (3), (5) complete the survey, and (6) logout. The questionnaire was web-based and more than one person per clinic involved in the study could check and update the answers to the questionnaire giving a consensus response. CSUQ consists of 19 items on a seven-point Likert [16] rating scale ranging from 1, “strongly agree” to 7, “strongly disagree”. Four usability scores can be derived by averaging responses to the CSUQ items: Overall Satisfaction, System Usefulness, Information Quality and Interface Quality [17]. Data about gender, age and previous experience with computer applications were noted. Responses to the CSUQ were obtained from seven of the nine clinics and the results were mixed. A majority of the clinics were quite satisfied with the usability although a sizeable minority were not. All evaluators were female and ages ranged from 38 to 61 (mean value 49). Two out of seven asserted that they had much previous experience with computers, whereas four had some experience and one had little experience.

To analyze potential benefits of using the system in everyday clinical practice, a “market acceptance test” is planned. The goal of the test is to determine whether compliance and usability are similar as in clinical studies. A further question is if there are country-specific issues posing different requirements for introduction into clinical practice. Other important issues that will be addressed include whether the usage of the system will improve treatment of patients during clinical practice and if a cost benefit analysis is favourable.

5. Hardware and software specifications

DPSS was written in the VB.NET programming language to allow data parsing and processing. It is designed to run with a PC-compatible machine with a minimum requirement of a Pentium processor with 256 of RAM. The M-function to calculate the spiral score was encrypted and wrapped into a C# interface by using the Matlab Builder for .NET to be accessed by the DPSS. This pre-built interface requires the Matlab Compiler Runtime to be installed in the running machine.

The WA was developed in ASP.NET as a code-behind model and ADO.NET was used as a standard way to connect to a database. Structured Query Language (SQL) was employed as a query language to access the data stored in Microsoft SQL Server. Three-tier architectures ensure a high-level of availability where the different application components can be easily replicated to increase the overall performance [18]. WA supports this architecture in which a web browser sends HTML requests using HTTP to a web server, which in turn passes the request to a common gateway interface application program. This application server then sends requests to a database server, which generates the query result set and sends it back in order to be formatted and presented to the end-user clients as an HTML page.

Microsoft Web Application Stress Tool (WAST) was used to stress the web server by realistically simulating a large number of users accessing the WA at the same time. The main measurements include response time and throughput [19]. The tests were run from a separate machine, running at 100 Mbits/s with varying load levels (concurrent connections) of 1, 10, 50, and 100. The web server, IIS 6.0, runs on Windows Server 2003 with CPU speed of 2.93 GHz and 1 GB of RAM. A script was developed, capturing a test scenario with a typical web browser. The test variability of the metrics was measured by repeating each test 10 times per load level of 1 min run time. The results showed that as the load level increased from 50 to 100, the number of requests that can be successfully served by the application per unit time starts to saturate. The average response time varied from page type to page type. For example, the login page introduced more delay compared to other pages; this was because it used the POST method to send data to the server. All average request times were in the range of 0.5 s indicating the users will not wait for too long for each page. However, memory requirements on the web server side tend to increase over time as the number of users accessing the database table storing spiral drawings data (x, y and time) increases. This is because this table contains a large number of records (around 8 million in the current dataset).

6. Mode of availability of the system

Currently the system is in a prototype stage and is being used in studies. To date it has been used by 65 Swedish patients in the DAPHNE study and by 35 patients in a validation study in Italy [20]. The next stage of development will be to perform the market acceptance test mentioned in Section 4.1. Thus, the system is not yet publicly available.

7. Discussion

In summary, there are four steps in our research concerning the system; first selection of what and how to measure to capture the symptom state; second comes validation that the measurements are accurate and measure the right things; third comes providing access to methods useful for decision making and fourth comes analysis of the implications of use. This work focuses on step three.

The use of an overall score may facilitate screening patients and help avoiding sub-optimization of treatments. Since symptom profiles are so different in PD patients, an overall score can be beneficial for deciding if a treatment change leads to an improvement of a patient’s general condition or not. The summary scores have been validated in a separate study (in Italy, 35 patients), with objective to assess test-retest
reliability, correlations to other assessment methods and ability to detect differences between patient groups at different disease stages. Compliance and reliability of the test battery were good. Correlations to rating scales were adequate and difference in overall test score between stable and fluctuating patient groups was detected [20].

The system is still under development, but our experiences this far indicate that it can be used as a tool for frequent assessments of PD symptoms in home environments. It is able to summarize the various time series of self-assessments and motor test results during week-long test periods and present them in a useful manner. The convenient access to current symptom profile as well as symptom history provides a basis for individualized evaluation and adjustment of treatments.

One drawback of the usability evaluation is the absence of feedback from PD physicians. This limitation will be taken into account in the future usability tests for the final version of the system. Additional functionalities that are not yet implemented include functions for producing alerts/highlights if symptoms are worse compared to the previous test period, providing a distributed access control to connect the treating clinical staff with their patients’ hospital records and to decide whether the requested information is allowed to be viewed by this treating clinical staff.

Conflicts of interest

The authors are shareholders and J. Westin, D. Nyholm, M. Dougherty and T. Groth are board members in Jemndator AB, the provider of the test battery.

Acknowledgements

This work was performed in the framework of the E-MOTIONS project, funded by the Swedish Knowledge Foundation, Abbott Product Operations AG, Allschwil, Switzerland, Nordforce Technology, Stockholm, Sweden and Dalarna University, Borlänge, Sweden.

REFERENCES