Using measurements from wearable sensors for automatic scoring of Parkinson's disease motor states: results from 7 patients

Thomas Ilias, Bergquist Filip, Constantinescu Radu, Nyholm Dag, Senek Marina, and Memedi Mevludin

Abstract— **The objective of this study was to investigate the validity of an objective gait measure for assessment of different motor states of advanced Parkinson's disease (PD) patients. Seven PD patients performed a gait task up to 15 times while wearing sensors on their upper and lower limbs. Each task was performed at specific points during a test day, following a single dose of levodopa-carbidopa. At the time of the tasks the patients were video recorded and three movement disorder experts rated their motor function on three clinical scales: a treatment response scale (TRS) that ranged from -3 (very bradykinetic) to 0 (ON) to +3 (very dyskinetic), a dyskinesia score that ranged from 0 (no dyskinesia) to 4 (extreme dyskinesia), and a bradykinesia score that ranged from 0 (no bradykinesia) to 4 (extreme bradykinesia). Raw accelerometer and gyroscope data of the sensors were processed and analyzed with time series analysis methods to extract features. The utilized features quantified separate limb movements as well as movement symmetries between the limbs. The features were processed with principal component analysis and the components were used as predictors for separate support vector machine (SVM) models for each of the three scales. The performance of each model was evaluated in a leave-one-patient out setting where the observations of a single patient were used as the testing set and the observations of the other 6 patients as the training set. Root mean square error (RMSE) and correlation coefficients for the predictions showed a good ability of the models to map the sensor data into the rating scales. There were strong correlations between the SVM models and the mean ratings of TRS (0.79; RMSE=0.70), bradykinesia score (0.79; RMSE=0.47), and bradykinesia score (0.78; RMSE=0.46). The results presented in this paper indicate that the use of wearable sensors when performing gait tasks can generate measurements that have a good correlation to subjective expert assessments.**

*Research supported by the Swedish Knowledge Foundation (Sweden), Vinnova (Sweden), Acreo (Sweden), Cenvigo (Sweden), Sensidose (Sweden), Uppsala University (Sweden), and Dalarna University (Sweden)

Ilias Thomas is with Dalarna University, Falun, 79 131, Sweden (tel: 0046 23 77 82 82; e-mail: ith@du.se).

Filip Bergquist (e-mail: filip.bergquist@gu.se) and Radu Constantinescu (e-mail: radu.constantinescu@vgregion.se) are with Gothenburg University, Gothenburg, 405 30, Sweden.

Dag Nyholm (e-mail: dag.nyholm@neuro.uu.se) and Marina Senek (e-mail: marina.senek@neuro.uu.se) are with the Dept. of Neuroscience, Neurology, Uppsala University, Uppsala, 752 36, Sweden

Mevludin Memedi is with Dalarna University and Orebro University, Orebro, 702 81, Sweden (e-mail: mmi@du.se, mevludin.memedi@oru.se).

I. Introduction

Parkinson's disease (PD) is a neurological disorder associated with motor (e.g. bradykinesia, rigidity, and tremor) and non-motor symptoms. It is estimated that about 9 million people will suffer from the disease by 2030 in the world's ten most populous nations [1]. PD symptoms can be assessed by rating motor functions using standardized clinical scales. In advanced stages of the disease, motor complications start to appear and disease progression should be monitored closely. However, clinician-based assessments are not able to capture the day-to-day variation of the symptoms. That is because physicians can only get a snapshot of the disease during patients' visits. Capturing symptoms continuously gives the treating physician a more comprehensive view of a patient's condition [2].

Wearable technologies could assist with that task, as they have shown the ability to produce objective measurements of motor functions [3]. More specifically, wearable sensors are found to be useful in providing: remote and continuous monitoring of motor symptoms [4]; accurate and valid objective assessment of symptoms and the ability to discriminate between ON, OFF and dyskinesia [5, 6]. Bradykinesia (or OFF) is the state where the patients are under-medicated, ON is the state where they are properly medicated, and dyskinesia is the state where they experience motor complications due to over-medication.

The goal of this study was to investigate the validity of an objective gait measure for assessment of different motor states of advanced PD patients. Subjective expert assessments based on three rating scales for PD where used to train machine learning algorithms [7]. The performance of machine learning models was evaluated by comparing the predictions to the experts' ratings. The three scales were: the Treatment Response Scale (TRS) [8], a dyskinesia score and item 31 from unified Parkinson's disease rating scale (bradykinesia). The TRS scale is particularly interesting as it provides an overall picture of patients' motor function with respect to bradykinesia and dyskinesia, with gait weighted as more important in the case of mixed patterns. The dyskinesia score provides information about potential over-medication of patients and the bradykinesia score addresses under-medication. The combined use of all scales could provide the treating physicians with a better picture when evaluating medication efficacy.

978-1-5090-2809-2/17/$31.00 ©2017 IEEE

II. METHODS

A. Dataset

Nineteen participants with advanced PD, experiencing fluctuations, were recruited in a single center open label clinical trial at Uppsala University Hospital, Sweden in 2015. The study design was approved by the regional ethics review board in Uppsala and the participants had given written consent to the trial, in accordance with the Helsinki declaration.

B. Data collection

The trial consisted of single levodopa-carbidopa dose administration, followed by motor assessments. After at least 8h washout of the subjects' regular Parkinson medication, a levodopa dose equalling 150% of their normal morning dose was administered [9, 10]. The reason for administering a larger dose than normal was to increase the likelihood that the subjects would display all three treatment states (OFF, ON, dyskinesia) during the measurements. The subjects performed a 2.5-meter straight walk, repeating it three times by taking 2 U-turns in the process. The task was performed before the dose (T=-20 min) at the time of the dose (T=0) and then at specific time points after the dose was taken (20, 40, 60, 80, 110, 140, 170, 200, 230, 260, 290, 320 and 350 min). It was repeated 15 times, or until the subjects felt that the drug effect had worn off and they could not remain unmedicated any longer. Each gait test took 15 seconds to complete.

While performing the tests the patients were wearing one sensor on each limb. The patients were also video recorded and the videos were presented in a randomized order to three movement disorder experts. The motor function of the patients was rated on the TRS, ranging from -3 (very OFF) to +3 (very dyskinetic), dyskinesia score, ranging from 0 (no dyskinesia) to 4 (severe dyskinesia), and bradykinesia score that ranges from 0 (no bradykinesia) to 4 (extreme bradykinesia). The scores of the raters were averaged and the mean values were used in the analyses.

Sensor data were collected from all participants but most of them had partially missing data. Seven participants that had complete sensor datasets were included in this first analysis. The population characteristics can be seen in table 1.

TABLE I: Population characteristics of the 7 individuals participating in the study.

Sex	Mean (SD) age (years)	Mean (SD) height (m)	Mean (SD) weight (kg)	Mean (SD) years with the disease	Mean (SD) years on Levodopa
6M/1W	70 (5.9)	1.76 (0.11)	77.4 (9.9)	11 (3.9)	10 (3.2)

C. Sensor description and feature extraction

Four sets of Shimmer3 sensors were used to record movements during the gait tests. Each sensor consisted of a 3D accelerometer and gyroscope and recorded movements on three axes (x, y and z). The sensors had a sampling rate of 102.4Hz with wide accelerometer range +/-16g and gyroscope range +/- 2000 dps and were placed at the lower and upper limbs: left wrist (LW), right wrist (RW), left leg (LL), and right leg (RL).

Time series analysis in the time domain and the frequency domain was performed to extract features for the sensors. Specifically, discrete wavelet transforms features, means, standard deviations, skewness and approximate entropy features were extracted. In total 152 spatiotemporal features were calculated: 32 features from each sensor and 24 features that quantified the differences of symmetry between the sensors.

For each sensor, the following eight signals were extracted and used in subsequent analysis: Xacc, Yacc, Zacc, Xgyr, Ygyr, Zgyr, Macc, and Mgyr. Xacc, Yacc, Zacc represent the acceleration in the three axes and Xgyr, Ygyr, Zgyr the orientation in the three axes. These values were directly extracted from the sensor readings. Macc and Mgyr represent the magnitudes of acceleration and gyroscope signals and were calculated by taking the square root of the sum of the squares for the three individual axes as seen in equations 1 and 2.

$$Macc = \sqrt{Xacc^2 + Yacc^2 + Zacc^2} \quad (1)$$

$$Mgyr = \sqrt{Xgyr^2 + Ygyr^2 + Zgyr^2} \quad (2)$$

For each of the eight signals and for each sensor (LW, RW, LL, and RL) the following features were calculated:

- The first three statistical moments: mean, standard deviation and skewness. Those were calculated to quantify level, variation and symmetry in movements during the walking trial. A total of 96 features (3 moments X 8 signals X 4 sensors) were extracted.

- Irregularities in movements during walking were quantified by applying Approximate Entropy (ApEn) on the two magnitude signals. For each of the four sensors, two features (one for Macc and one for Mgyr) were calculated, resulting in 8 more features.

- A three level Discrete Wavelet Transform (DWT) using Daubechies (db10) wavelet family was applied on the two magnitude signals.

  - ❖ The first level resulted in low-frequency components between 0 and 25.6Hz and high-frequency components between 25.6 and 51.2Hz.
  - ❖ In the second level, the low frequency components of the first level were further decomposed into low (0-12.8Hz) and high (12.8-25.6Hz) frequencies.
  - ❖ Finally, the third level decomposed the low frequency components of the second level into low (0-6.4Hz) and high (6.4-12.8Hz) frequencies.

  The first two statistical moments (mean and standard deviation) of Macc were calculated for 3 DWT signals.

The first level high frequencies, the second level high frequencies, and a combined signal, which consisted of the following components in the following order: 1) third level low frequencies, 2) third level high frequencies, 3) second level high frequencies and, 4) first level high frequencies. This analysis resulted in 24 features (2 moments X 3 DWT signals X 4 sensors).

- Finally, to capture dissimilarities in movements between the four limbs during walking, Dynamic Time Warping (DTW) and Cross Approximate Entropy (Cross-Apen) were applied on the Macc and Mgyr signals. For each of these two signals, DTW and Cross-Apen were applied to calculate features for the 6 sensor combinations (LW vs. RW, LW vs. LL, LW vs. RL, RW vs. LL, RW vs. RL, and LL vs. RW). This analysis resulted in 24 additional features (2 signals X 6 sensor combinations X 2 methods).

In order to reduce the number of features, principal component analysis (PCA) was applied to the 32 features of each sensor. The first ten principal components (PC) were kept from each sensor that explained about 90% of the total variation. That analysis reduced the total number of features to 64 (10 PC from each sensor, plus 24 features that capture dissimilarities).

D. Predictive modelling

The PCs of the four sensors together with the 24 DTW and Cross-Apen features were used as predictors in three separate support vector machine (SVM) models. For each of the three scales, each individual model mapped the predictors into the mean ratings of the three experts. The SVMs used radial basis function with two parameters, gamma and cost parameter, and epsilon value fixed at 0.1. The parameters of the SVM models were optimized based on a grid search of the parameter space to optimize the performance of the model in a cross validation (CV) setting. Inclusion of the predictors in the final models was decided based on a bidirectional elimination method. The criterion was whether the performance of the models (in a CV setting) improved upon the inclusion or removal of a predictor. A leave-one-individual-out CV method was selected to evaluate the performance of the models, where data from six subjects were used during training and data from one subject were used during testing. Two statistics metrics were used for the evaluation of the performance of the SVM models: the correlation of the predictions to the mean ratings of the three experts and the root mean squared error (RMSE) of the predictions.

III. RESULTS

Table 2 shows an analytical description of the number of predictors used for each SVM model. The predictors have been organized into three groups: PCs from the upper limbs, PCs from the lower limbs and the symmetry features (DTW and Cross-Apen features). The results indicate that the predictors from the upper limbs were the most important when training and testing the machine learning models.

TABLE II: The number of the three predictor groups in the support vector machine models.

	TRS model	Bradykinesia model	Dyskinesia model
Upper limb PC (number)	3	4	7
Lower limb PC (number)	-	2	1
Symmetry features (number)	3	1	-

Preliminary results on data from 7 patients showed good predictive ability of the SVM models in a leave-one-individual out CV setting. The predictions of the models in all rating scales were comparable to the mean ratings of the three experts with correlation coefficients and RMSE values of 0.79 and 0.70 for TRS, 0.79 and 0.47 for dyskinesia, and 0.78 and 0.46 for bradykinesia. The results of the three models can be seen in figures 1, 2, and 3 respectively.

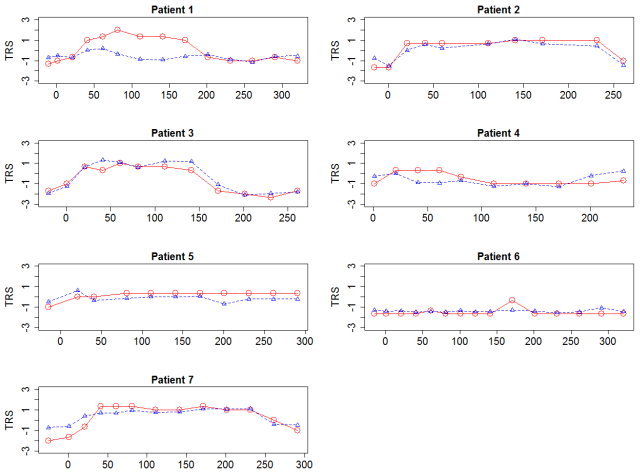

Fig. 1: Cross validation results from the support vector machine model on the TRS scale. The raters' values are colored in red (circles) and the model's predictions are colored in blue (triangles). On the x-axis is time (in minutes).

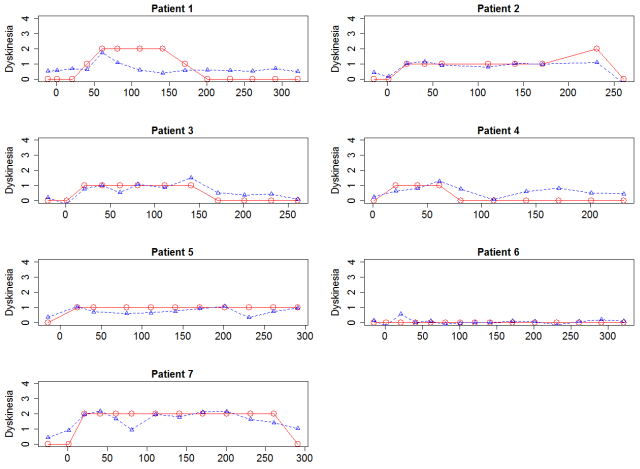

Fig. 2: Cross validation results from the support vector machine model on the dyskinesia score. The raters' values are colored in red (circles) and the model's predictions are colored in blue (triangles). On the x-axis is time (in minutes).

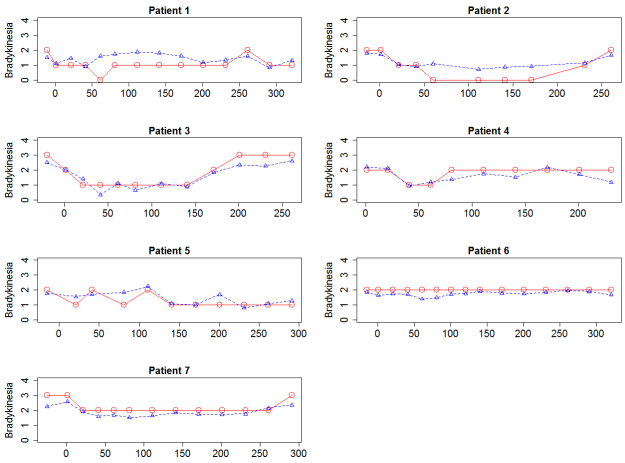

Fig. 3: Cross validation results from the support vector machine model on the bradykinesia score. The raters' values are colored in red (circles) and the model's predictions are colored in blue (triangles). On the x-axis is time (in minutes).

IV. Conclusion

In this paper we have assessed the feasibility of objective measures of gait in quantifying PD motor functions, using data from wearable sensors. The results from the 7 first analyzed patients suggest that the use of wearable sensors when performing a walking task can capture useful information for assessing treatment-related changes in advanced PD. High correlation was found for the TRS scale. Because TRS is a combination of the bradykinesia and dyskinesia scores, high correlation was also expected in those scores.

However, the models had a tendency to under-predict the high scores. This limitation can be explained by the low number of observations used for building and testing the SVM models. An interesting finding is the fact that in all models, the strongest predictors were the ones from the upper limb sensors. More specifically, there were no predictors from the lower limbs that were relevant to be included in the model for predicting the TRS. Similarly, in the dyskinesia model there were 7 predictors from the upper limbs and 1 from the lower limbs and in the bradykinesia model there were 4 features from the upper limbs, 2 from the lower limbs and 1 symmetry feature. These findings are consistent with previous research that suggests hand movements are one of the strongest predictors of PD severity [11]. Another explanation for these results would be the fact that the current models are data-driven and the features extracted from the sensor signals are not related to the kinematics and biomechanics of the movements during gait. Therefore, future work will investigate a knowledge-based approach to examine whether gait-related features would be more correlated to clinician-based ratings of TRS, dyskinesia, and bradykinesia scales [12].

Furthermore, the results presented here are preliminary, based on data from 7 patients, out of 19 that completed the study. Future steps will focus on analyzing sensor data from all subjects. The current models could require certain adjustments and optimization to generalize better to the general patient population and, in later stages, be applied to free walking. Moreover, the plans include evaluation of other clinimetric properties of the methods such as reliability and responsiveness to treatment. Additionally, the sensors could be applied to establish what normal behaviour is (from healthy controls) and the results could then be used to set a target of normal functions (when medicating).

In conclusion, the results presented in this paper indicate that the use of wearable sensors when performing walking tasks can generate measurements that have a good correlation to subjective expert assessments. This could be useful during individualized evaluation of symptoms and treatments.

References

[1] ER Dorsey, R Constantinescu, JP Thompson, KM Biglan, RG Holloway, K Kieburtz, FJ Marshall, BM Ravina, G Schifitto, A Siderowf, et al. "Projected number of people with Parkinson's disease in the most populous nations, 2005 through 2030." *Neurology,* 68(5):384–386, 2007.

[2] D Isacson, K Bingefors, Ivar Sønbø Kristiansen, and D Nyholm. "Fluctuating functions related to quality of life in advanced Parkinson disease: effects of duodenal levodopa infusion." *Acta Neurologica Scandinavica,* 118(6):379–386, 2008.

[3] Alberto J Espay, Paolo Bonato, Fatta B Nahab, Walter Maetzler, John M Dean, Jochen Klucken, Bjoern M Eskofier, Aristide Merola, Fay Horak, Anthony E Lang, et al. "Technology in Parkinson's disease: Challenges and opportunities." *Movement Disorders*, 2016.

[4] Walter Maetzler, Josefa Domingos, Karin Srulijes, Joaquim J Ferreira, and Bastiaan R Bloem. "Quantitative wearable sensors for objective assessment of Parkinson's disease." *Movement Disorders*, 28(12):1628–1637, 2013.

[5] Andong Zhan, Max A Little, Denzil A Harris, Solomon O Abiola, E Dorsey, Suchi Saria, and Andreas Terzis. "High frequency remote monitoring of Parkinson's disease via smartphone: Platform overview and medication response detection." *arXiv preprint* arXiv:1601.00960, 2016.

[6] Manuela Galli, Sara L Vimercati, Elena Manetti, Veronica Cimolin, Giorgio Albertini, and Maria F De Pandis. "Spiral analysis in subjects with Parkinson's disease before and after levodopa treatment: a new protocol with stereophotogrammetric systems." *Journal of applied biomaterials & functional materials*, 12(2), 2014.

[7] Ken J Kubota, Jason A Chen, and Max A Little. "Machine learning for wearable sensor data in parkinson disease: concepts, pitfalls and futures." *Movement Disorders*, 2016.

[8] Dag Nyholm, AIM Nilsson Remahl, N Dizdar, R Constantinescu, B Holmberg, R Jansson, S-Aquilonius, and H°akan Askmark. "Duodenal levodopa infusion monotherapy vs oral polypharmacy in advanced parkinson disease." *Neurology*, 64(2):216–223, 2005.

[9] Claire L Tomlinson, Rebecca Stowe, Smitaa Patel, Caroline Rick, Richard Gray, and Carl E Clarke. "Systematic review of levodopa dose equivalency reporting in Parkinson's disease." *Movement disorders*, 25(15):2649–2653, 2010.

[10] Marina Senek, Sten-Magnus Aquilonius, Håkan Askmark, Filip Bergquist, Radu Constantinescu, Anders Ericsson, Sara Lycke, Alexander Medvedev, Mevludin Memedi, Fredrik Ohlsson, et al. "Levodopa/carbidopa microtablets in Parkinson's disease: a study of pharmacokinetics and blinded motor assessment." *European Journal of Clinical Pharmacology*, pages 1–9, 2017.

[11] Anat Mirelman, Hagar Bernad-Elazari, Avner Thaler, Eytan Giladi-Yacobi, Tanya Gurevich, Mali Gana-Weisz, Rachel Saunders-Pullman, Deborah Raymond, Nancy Doan, Susan B Bressman, et al. "Arm swing as a potential new prodromal marker of Parkinson's disease." *Movement Disorders*, 31(10):1527–1534, 2016.

[12] Olumide Sofuwa, Alice Nieuwboer, Kaat Desloovere, Anne-Marie Willems, Fabienne Chavret, and Ilse Jonkers. "Quantitative gait analysis in Parkinson's disease: comparison with a healthy control group." *Archives of physical medicine and rehabilitation*, 86(5):1007–1013, 2005.