Large-Scale Continuous Mobility Monitoring of Parkinson's Disease Patients Using Smartphones

Wei-Yi Cheng[1], Florian Lipsmeier[2], Andrew Creigh[2,3], Alf Scotland[2], Timothy Kilchenmann[2], Liping Jin[1], Jens Schjodt-Eriksen[2], Detlef Wolf[2], Yan-Ping Zhang-Schaerer[2], Ignacio Fernandez Garcia[2], Juliane Siebourg-Polster[2], Jay Soto[4], Lynne Verselis[1], Meret Martin Facklam[2], Frank Boess[2], Martin Koller[4], Machael Grundman[4,5,6], Andreas U. Monsch[7,8], Ron Postuma[9], Anirvan Ghosh[2], Thomas Kremer[2], Kirsten I. Taylor[2], Christian Czech[2], Christian Gossens[2(✉)], and Michael Lindemann[2,10]

[1] Roche Innovation Center New York, Roche TCRC Inc., New York, USA
[2] Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Basel, Switzerland
christian.gossens@roche.com
[3] University of Oxford, England, UK
[4] Prothena Biosciences Inc., South San Francisco, USA
[5] Global R&D Partners, LLC, San Diego, USA
[6] University of California, San Diego, USA
[7] University Center for Medicine of Aging, Memory Clinic, Felix Platter Hospital, Basel, Switzerland
[8] University of Basel, Basel, Switzerland
[9] Department of Neurology, McGill University, Montreal General Hospital, Montreal, Canada
[10] Baden-Wuerttemberg Cooperative State University, Loerrach, Germany

Abstract. Smartphone-based assessments have been considered a potential solution for continuously monitoring gait and mobility in mild to moderate Parkinson's disease (PD) patients. Forty-four PD patients from cohorts 4 to 6 of the Multiple Ascending Dose (MAD) study of PRX002/RG7935 and thirty-five age- and gender-matched healthy individuals (i.e. healthy controls - HC) in a separate study performed smartphone-based assessments for up to 24 weeks and up to 6 weeks, respectively. The assessments included "active gait tests", where all participants were asked to walk for 30 s with at least one 180° turn, and "passive monitoring", in which subjects carried the smartphone in a pocket or fanny pack as part of their daily routine. In total, over 6,600 active gait tests and over 30,000 h of passive monitoring data were collected. A mobility analysis indicates that patients with PD are less mobile than HCs, as manifested in time spent in gait-related activities, number of turns and sit-to-stand transitions, and power per step. It supports the potential use of smartphones for continuous mobility monitoring in future clinical practice and drug development.

P. Perego et al. (Eds.): MobiHealth 2017, LNICST 247, pp. 12–19, 2018.
https://doi.org/10.1007/978-3-319-98551-0_2

Keywords: Sensors · Activity recognition · Smartphone
Accelerometer · Machine learning · Deep learning
Parkinson's disease · Clinical trial

1 Introduction

Mobility-related symptoms are among the earliest symptoms of Parkinson's disease (PD) and are part of the clinical diagnosis [1]. Functional impact of PD on mobility-related activities such as walking, turning or rising from chair has an impact on patients' quality of life and is also used clinically as an indication for disease progression. In order to provide an objective and quantitative assessment of gait and mobility, many studies have implemented wearable systems or body-fixed sensors in both controlled [2,3] and free-living settings [4,5]. While the on-body sensors are light-weighted and inexpensive, the difficulties of integrating them into daily living pose an extra burden to the subjects and usually limit the length of the study to less than one month [5].

Most smartphones have built-in sensors to provide more environment-aware services and applications utilize these sensors, such as accelerometers, gyroscope and magnetometers. As smartphones have become standard equipment in daily life, they provide a more natural way of performing long-term mobility assessment in a clinical setting. In this paper, we present the results of a large-scale, longitudinal gait and mobility assessment of PD patients in a clinical trial setting using smartphones and compare them with age and gender matched controls.

2 Methods

2.1 Data Collection

Information on the MAD clinical trial of PRX002/RG7935 study can be found online [6]. Our analysis focused on exploring the between-group differences between HC and PD, and did not assess PRX002/RG7935-related effects. PD subjects from cohorts 4 to 6 of the MAD study and all participants from the HC study were provided with a locked-down smartphone only running a dedicated app, which allowed for the execution of active tests and continuous sensor recordings.

Active Gait Tests. Each participant was asked to perform at least one active test in the morning. Participants were requested to walk in a straight line with minimal turns for 30 s. In the trial, 5,107 active gait tests were completed by the PD cohort and 1,589 in the HCs. The sensor readouts of the accelerometer and gyroscopes were measured at 66 Hz. All of the gait and mobility features in the active tests were extracted while excluding the first and last 5 s of the test, as during these time spans often the subjects were putting the smartphones into or removing it from their pockets or fanny packs.

Passive Monitoring. Every day the participants were requested to carry the smartphone in their pocket or in a fanny pack for as long as the battery life allowed (approximately 6 h), while the sensors recorded their movement. In total, 24,104 h of passive monitoring data were recorded for the PD cohort, and 8,614 h for the HCs. In line with a previously published approach [7], we filtered out accelerometer data where the standard deviation of Euclidean norm was less than 0.03 m/s^2 for more than 30 min, as during these spans smartphones were likely not carried by the subjects. This step removed 24% of the passive monitoring data.

2.2 Human Activity Recognition (HAR)

A diagram of the 9-layer neural network model structure and an example data flow is shown in Fig. 1. Similar structures have been used previously for HAR and have been shown to out-perform the traditional machine learning methods [8]. Our HAR model was trained on two public data sets [9,10] to classify six activities: walking, stairs, jogging, sitting, standing, and lying down. The continuous accelerometer data were down-sampled into 20 Hz and segmented into 4-s windows with 75% overlapping with adjacent ones.

When making predictions, the model classifies the overlapping windows into the six activities. As each second in the trial data was covered by four windows, we determined the final predictions on each second by performing a majority vote using the predicted activities of the four windows. In case of a tie, the predicted activity was determined by the one with highest summed predicted probability.

For sit-to-stand and stand-to-sit (STS) detection, we counted the number of occurrence of sit or lying down spans preceded by or followed by stand, walk, stairs, or jog spans. The STS extraction algorithm would only accept the event as a true STS event if and only if during the span transition the phone orientation has changed, or step detection algorithm has detected steps in the gait spans.

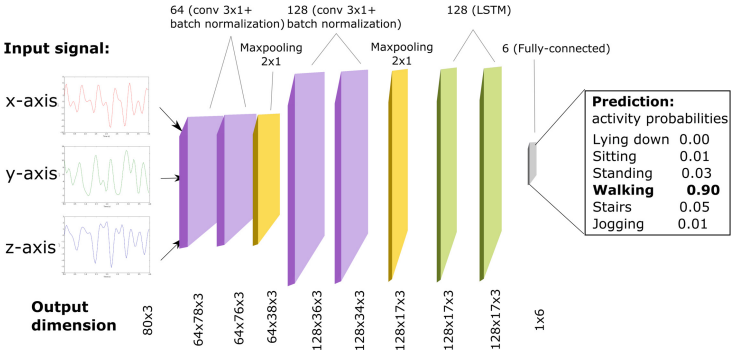


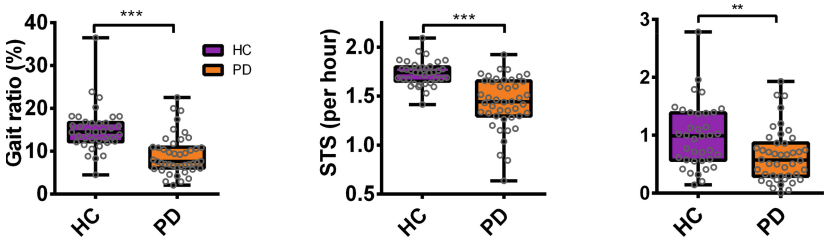
Fig. 1. Example dataflow from raw signal to activity classification

2.3 Gait and Mobility Feature Extraction

The mobility features we extracted from gaits and turns are all based on previously investigated features [5,11–13].

We applied the adaptive step detection algorithm proposed by Lee *et al.* [14] to determine the time points of step initiation. After steps were identified in the gait spans, we calculated average step frequencies (number of steps divided by time) per subject. For the passive monitoring data we only performed the step detection and feature extraction algorithm in gait-related activities (walking, stairs, jogging) that were longer than five seconds. To infer the power invested during walk, we calculated the integral of the variance of the Euclidean norm of the acceleration signal, divided by the number of steps as the per-step power coefficient. This metric is a surrogate coefficient for power given that we do not have mass of the subject.

For turn detection, we followed a three-step process. First, we used a minimization method to identify the optimal rotation matrix $R^* = R_x(pitch)R_y(roll)$ such that the average of the acceleration signal on the z-axis follows gravity. That is, $\text{argmin}_{pitch,roll}(\int_t R_x(pitch)R_y(roll)\text{a}_z g\,\text{d}t)$, where *pitch* and *roll* are the pitch angle and roll angel, g is the gravity, and a_z is the z-axis component of the acceleration signal. In the second step we applied the rotation matrix R^* on the gyroscope signal. To detect turns around the z-axis, we integrated the angular velocity on a 1.5 s rolling window to obtain the yaw angles: $\int_t (R^*\text{gyro})_z\,\text{d}t$ where $()_z^{\cdot}$ takes the z-axis component of the rotated signal. We then detect peaks in yaw angles that are higher than 1.5 rad (∼86°) as one turn event. Our method is comparable to similar methods described in [12,15], but uses a more stringent turn criterion in terms of degrees and speed of turns. For computational reasons we only estimated turns for gait spans longer than 18 s. This leaves about 57,000 walking spans collected during passive monitoring.



(a) Gait ratio

(b) Average STS events per hour

(c) Difference of average number of turns per walking minute

Fig. 2. Associations between HAR-profiled mobility measurements and clinical features. *Statistical significance:* $^{**}P$ value < 0.01; $^{***}P$ value < 0.001

3 Results

3.1 Human Activity Recognition Performance Validation

Before applied on the passive monitoring data, the HAR model was validated on two data sets. The first was the held-out test set from the same sources of training data. The HAR model was able to correctly distinguish gait activities (walking, stairs, jogging) from stationary activities (sitting, standing, lying down) with more than 98% of accuracy. In addition, we also performed validation using the labeled daily active test. The active gait test sensor data was used as the positive control for gait detection, while the balance test, during which the subjects stand still for 30 s, was used as negative control. We defined a correct prediction as more than 50% of the span was labeled with the correct activity. The HAR model was also able to successfully profile the gait segments with 96.9% of accuracy, and balance segments with 99.5% accuracy.

3.2 Activity Profiles and Mobility Features Comparison in Passive Monitoring

The mobility of each subject was quantified by gait ratio, defined as the proportion of time the subject engaged in gait activities over the total passive monitoring coverage. Figure 2a shows the between-group difference in gait ratios. In the PD cohort we detected a median of 9.7% of gait spans over all coverage spans as supposed to HC cohort's 15.1%, indicating that the HC cohort had a significantly higher per-subject gait activity level than PD cohort ($P < 0.001$, Mann-Whitney test).

From the activity profile, we calculated the coverage-normalized STS events for each subject. Concordant with the results previously presented [11], the median number of STS per hour in PD patients was 17% lower than in HC subjects, as shown in Fig. 2b ($P < 0.001$, Mann-Whitney test).

Difficulties while turning is a regular symptom of PD [16] and may influence the way PD patients walk during their regular daily life. From our turn detection we calculated the average number of turns per hour per subject. We observed a 38% median reduction in the PD cohort versus HCs ($P < 0.01$, Mann-Whitney test), as shown in Fig. 2c. This result differs from a similar albeit smaller and shorter study by others [13], where no significant difference could be observed. One possible reason for this could be our focus on turns with higher degree change that pose a much higher difficulty for gait impaired PD patients.

3.3 Comparison of Gait Features Between Active Tests and Passive Monitoring

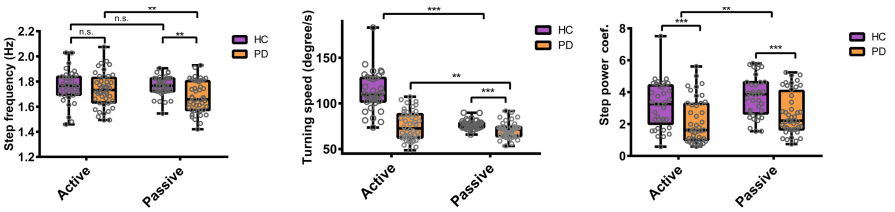
Active tests, such as our Gait Test, constitute an artificial situation and one open question is how much they reflect the subjects' behavior during their regular daily life. We compared three important and typical gait features: step frequency, step power and turning speed, between active test and passive monitoring for HC and

PD patients. Additionally, we also compared group differences between HC and PD cohorts. For the comparison we only used the subset of passive monitoring spans, which the HAR labeled as walking, and span length between 20 to 40 s.

Figure 3a shows that there is a significant mean reduction of 0.05 steps per second between active test and passive monitoring in the PD patient data ($P < 0.01$, paired t-test), whereas the difference in HCs is not significant ($P > 0.1$, paired t-test). However, we observed that during passive monitoring the HCs has much less variability in step frequency than that of PD. This low variability in HCs and the decrease in step frequency for PD patients may help explain the significant difference between HC and PD in passive monitoring ($P < 0.01$, Mann-Whitney test).

For turning speed, as depicted in Fig. 3b, we observed statistically significant difference between active and passive in both groups ($P < 0.001$, paired t-test). While the reduction in turning speed in PD is much more pronounced in the active tests, the differences in both active test and passive monitoring are statistically significant ($P < 0.001$, Mann-Whitney test). Similar to the case of step frequency, the HC cohort harbors much less variability than the PD cohort in passive monitoring.

Finally, Fig. 3c shows the power invested while walking in the active tests versus during passive monitoring. For both active and passive monitoring we observed significant lower power in PD versus HCs ($P < 0.001$, Mann-Whitney test). Both the PD cohort and HCs show significantly higher per-step power in passive monitoring then in the active gait tests ($P < 0.01$, paired t-test). The significance levels of comparisons between various groups are summarized in Table 1.



(a) Average step frequency (b) Average turning speed (c) Average step power coefficient

Fig. 3. Comparison of active vs. passive in HC and PD as well as between group comparison of HC vs. PD. *Statistical significance:* $^{**}P$ value < 0.01; $^{***}P$ value < 0.001

Table 1. Significance level for active test and passive monitoring comparison.

Test difference in % (P value)	Step frequency	Tuning speed	Step power coefficient
PD vs HC in active test	−2.0% (n.s.)	−33.6% (<0.001)	−50.0% (<0.001)
PD vs HC in passive monitoring	−6.1% (<0.01)	−6.4% (<0.001)	−40.0% (<0.001)
Active vs passive in PD	−4.3% (n.s.)	−2.1% (<0.001)	70.7% (<0.01)
Active vs passive in HC	−0.0% (<0.01)	−30.2% (<0.01)	42.4% (<0.01)

4 Discussions

As mobility reflects a very important aspect in the quality of life while being a diagnostic indication for PD, continuous long-term monitoring can provide valuable insight for treatment strategy development. Our study shows that such monitoring on mild to moderate PD patients is feasible using smartphones. Sensor data collected from both active and passive monitoring provide previously inaccessible information regarding patients' daily behavior and functioning. Specifically we demonstrated that patients with PD in our study were less mobile than healthy controls, as manifested in time spent in gait-related activities, number of turns and sit-to-stand transitions, and power per step. For further understanding on whether these features reflect disease severity as measured by physician rated assessment, we are performing ongoing analyses which include correlating the Movement Disorder Society unified Parkinson's disease rating scale (MDS-UPDRS), that is used in clinic to evaluate PD severity, with the smartphone data. Another important issue is how robust these measurements are across different mobile devices. In this study, all the subjects were using the same smartphone model. In the future, if similar study is to be performed using subjects' own devices, as different commercial smartphone models have different built-in inertial measurement unit (IMU), the acceleration and angular speed readings will require more comprehensive benchmarking so the derived features are comparable across devices.

The comparison of step features between active and passive monitoring provides valuable insight for future practice of remote monitoring programs. While for most step features we observed no significant difference in active test and passive monitoring, larger spread of step frequencies and turning speed in HC during active tests were clearly observed. We also observed that the step power is positively correlated with span length: both PD patients and HCs tended to invest more power in longer walking spans (data not shown). This agrees with the previously published result [5]. These pose interesting questions on human behavior and trial design for future studies.

Finally, the information from this study may be applicable to other motor function-related diseases to further understand disease progression or treatment effects, and eventually provides new perspectives in healthcare practices.

Acknowledgements. We thank Max A. Little for his technical input at the early stages of this research project.

References

1. Ramaker, C., Marinus, J., Stiggelbout, A.M., van Hilten, B.J.: Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. Mov. Disord. **17**(5), 867–876 (2002)
2. González, R.C., López, A.M., Rodriguez-Uría, J., Álvarez, D., Alvarez, J.C.: Real-time gait event detection for normal subjects from lower trunk accelerations. Gait Posture **31**, 322–325 (2010)

3. Zijlstra, A., Zijlstra, W.: Trunk-acceleration based assessment of gait parameters in older persons: a comparison of reliability and validity of four inverted pendulum based estimations. Gait Posture **38**(4), 940–944 (2013)
4. Godfrey, A., Lord, S., Galna, B., Mathers, J.C., Burn, D.J., Rochester, L.: The association between retirement and age on physical activity in older adults. Age Ageing **43**, 386–393 (2014)
5. Del Din, S., Godfrey, A., Galna, B., Lord, S., Rochester, L.: Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. J. NeuroEng. Rehabil. **13**, 46 (2016)
6. ClinicalTrials.gov. https://clinicaltrials.gov
7. Rai, A., Chintalapudi, K.K., Padmanabhan, V.N., Sen R.: Zee: zero-effort crowdsourcing for indoor localization. In: MobiCom (2012)
8. Ordóñez, F.J., Roggen, D.: Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. Sensors **16**, 115 (2016)
9. Weiss, G.M., Lockhart J.W.: The impact of personalization on smartphone-based activity recognition. In: Proceedings of the AAAI-12 Workshop on Activity Context Representation: Techniques and Languages, Toronto, Canada (2012)
10. Stisen, A., et al.: Smart devices are different: assessing and mitigating mobile sensing heterogeneities for activity recognition. In: 13th ACM Conference on Embedded Networked Sensor Systems, Seoul, Korea (2015)
11. Zijlstra, A., Mancini, M., Lindermann, U., Chiari, L., Zijlstra, W.: Sit-stand and stand-sit transitions in older adults and patients with Parkinson's disease: event detection based on motion sensors versus force plates. J. NeuroEng. Rehabil. **9**, 75 (2012)
12. El-Gohary, M., et al.: Continuous monitoring of turning in patients with movement disability. Sensors **14**, 356–369 (2014)
13. Mancini, M., et al.: Continuous monitoring of turning in Parkinson's disease: rehabilitation potential. NeuroRehabilitation **37**(1), 3–10 (2015)
14. Lee, H.H., Choi, S., Lee, M.J.: Step detection robust against the dynamics of smartphones. Sensors **15**(10), 27230–27250 (2015)
15. Pham, M.H.: Algorithm for turning detection and analysis validated under home-like conditions in patients with Parkinson's disease and older adults using a 6 degree-of-freedom inertial measurement unit at the lower back. Front. Neurol. **8**, 135 (2017)
16. Stack, E., Ashburn, A.: Dyscuntional turning in Parkinson's disease. Disabil. Rehabil. **30**(16), 1222–1229 (2008)