FinJeHeW Finnish Journal of eHealth and eWelfare

SCIENTIFIC PAPERS
VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Walking detection for Parkinson's disease patients and healthy control subjects measured with a smartphone accelerometer using mean amplitude deviation algorithm

Milla Juutinen[1], Jari Ruokolainen[2,3,4], Juha Puustinen[3], Anu Holm[5], Mark van Gils[1], Antti Vehkaoja[1]

[1] Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland; [2] Faculty of Business and Management, Tampere University, Tampere, Finland; [3] Unit of Neurology, Satasairaala Central Hospital, Wellbeing Services County of Satakunta, Finland; [4] School of Technology and Innovations, University of Vaasa, Finland; [5] Satakunta University of Applied Sciences, Pori, Finland

Milla Juutinen, Faculty of Medicine and Health Technology, Tampere University, Korkeakoulunkatu 10, FI-33720 Tampere, FINLAND. Email: milla.juutinen@tuni.fi

Abstract

Parkinson's disease is a neurodegenerative disorder that affects mobility, leading to a decline in the patient's quality of life. Analyzing gait for these patients aims to improve the mobility of the Parkinson's disease patients. The goal of this study was to validate mean amplitude deviation for detecting gait in Parkinson's disease patients and healthy controls. This method is robust and orientation-independent and has accurate results on physical activity detection for different study populations. The novelty of this study is to use and evaluate a previously validated method with Parkinson's disease patients instead of healthy young subjects as in earlier studies. We utilized inertial measurement unit data measured using smartphones from pre-existing datasets with pre-defined and labeled activities and free-living data containing continuously collected over three consecutive days. One dataset included 30 healthy adults, and the other two included in total of 62 and 68 Parkinson's disease patients and 40 and 39 healthy controls, respectively. The sensitivity of the algorithm in a controlled measurement setting was 100% and 98.7% for healthy adults and a combined dataset of Parkinson's disease patients and control subjects, respectively. Correspondingly, the specificity was 74.9% and 81.6%. Visual inspection of the free-living data showed that the algorithm provided durations and timings of walking activities, and walking took place during the daytime as anticipated for subjects with a typical daily rhythm. Median walking times were under ten minutes per hour. The results reached the same performance range as earlier studies with an orientation-independent approach, justifying the feasibility of this method. Therefore, this study validated the use of mean amplitude deviation for walking detection in Parkinson's disease patients and healthy adults. Future research will utilize the detected walking segments in analyzing the motor symptoms of the disease aiming to improve patient well-being through identifying the needs for additional healthcare interventions.

Keywords: Parkinson's disease, gait, neurodegenerative diseases, smartphone, acceleration

Introduction

Walking detection and step counting could provide objective insights to different gait issues related to diseases or aging in addition to traditional visual gait evaluation. For example, gait analysis is an essential tool for monitoring the progression and treatment efficacy of Parkinson's disease (PD), as it provides insight into the motor symptoms of the disease: tremor, rigidity, and slowness of movement [1]. Traditionally, these symptoms are evaluated visually by healthcare professionals at specific appointments, e.g., by utilizing the Movement Disorder Society Unified Parkinson's Disease Rating Scale [2], but usually, the specific tests are not recorded systematically and saved to the electronic health records. However, there is a need to analyze walking during daily living, since the symptoms vary substantially e.g., between ON and OFF medication states [3]. To achieve this, walking segments or steps need to be identified from other activities or rest.

Based on existing literature, free-living studies are less frequent than studies in laboratory settings [4]. Activity recognition algorithms are often developed and tested in supervised laboratory conditions [5,6], where a fixed location of the sensors, a clear separation of different activities, and control over the walking speed are used. For example, walking speed greatly affects gait phases, as demonstrated in several earlier studies [7–9]. The ground truth for these measurements could be obtained by using a video recording, step counters, and manual data labeling. Some settings could be also semi-controlled, where the instructions for completing tasks were given, but the participants were not monitored or supervised during the conduction of the experiment [4,10].

However, studies for implementing walking detection in less controlled environments have also been carried out [11–13]. Brajdic and Harle [11] tested several algorithms for both walking detection and step counting and evaluated the performance of the algorithms in terms of detection accuracy and computational efficiency. They concluded that for walking detection a standard deviation-based thresholding, and for step detection a windowed peak detection algorithm were the most efficient methods. Iluz et al. [12] studied the detection of missteps on free-living settings and concluded that they can be classified in PD patients with high accuracy. Pham et al. [13] validated a step detection algorithm, first in a laboratory environment and then tested the system in home-like setting.

Walking and step detection algorithms may be implemented with direct methods based on setting individual [14] or universal [15,16] thresholds for resultant acceleration, specific pattern or area of the acceleration signal [10,17,18]. The advantage of direct methods is the computational efficiency and general suitability to a new dataset. Still, the challenge is the selection and validation of threshold values and minimize the interference from walking characteristics or deviations. Micó-Amigo et al. [19] tested a walking detection algorithm created by Iluz et al. [12] based on identifying segments within 0.5–3 Hz frequency band to detect gait segments with PD patients resulting in accuracy of 96%. Dijkstra et al. [14] discovered that accelerometers provide more accurate results in step detection for PD patients than pedometers and they are less dependent on walking speed.

In addition to threshold-based algorithms, machine learning classifiers have been developed for walking and step detection [4,6,20], but they require large training datasets and more computational power for operation. Classifiers and direct methods can be compared against a ground truth annotated from video recordings of the experiment, or by

visually observing and manually annotating and timing the performed activities.

Identification of walking segments can be performed on fixed windows by analyzing the changes in the signal within consecutive windows. One of these methods is mean amplitude deviation (MAD), first introduced for physical activity detection by Vähä-Ypyä et al. in 2015 [15,16,21]. MAD measures how much a sample in the magnitude of acceleration deviates from the mean of the neighboring samples. It originated from the calculation of mean absolute deviation but is adjusted for acceleration time series. MAD is calculated for the magnitude of the three-dimensional acceleration vector and thus, the orientation of the device does not affect the interpretation of MAD.

The relationship between MAD and physical activity levels has been studied extensively [15,16,21–25]. The studies have identified MAD limits for physical activities for children and adults using different commercial acceleration sensors. Also, Vähä-Ypyä et al. [16] have identified universal cut-off limits based on three different commercial acceleration sensors. Based on the highest observed threshold values, sedentary activities are below 44 mg0 [24], light activities below 338 mg0 [21], moderate activities below 603 mg0 [21], and vigorous activities over that. There are considerable differences between limits found in different studies. Haapala et al. [25] and Gao et al. [23] utilized the limits defined by Vähä-Ypyä et al. [15]. Based on the existing literature, MAD has not yet been used to separate walking from other types of physical activities in healthy subjects, elderly, or specific patient groups.

Since the walking of PD patients might have additional changes to their walking style compared to age-matched healthy subjects, as shown by Yoneyama et al. [26], we have previously demonstrated how PD patients and healthy controls can be detected using controlled gait data [27] and concluded that a machine learning algorithm can separate between these groups based on characteristics of individual steps. We have also demonstrated a walking detection algorithm for controlled walking tests using signal magnitude area, as presented by Mehrang et al. [10]. The characteristics of walking have been shown to change with the severity of the disease and thus, PD symptom severity can be monitored using foot worn sensors and a random forest classifier with appropriate features [28–30].

The objective of this study was to address the walking detection challenge by validating that MAD could be a robust and accurate algorithm for detecting gait from accelerometer data of PD patients. These patients may have walking related issues both due to their age and progressive disease. The qualitative hypothesis was that if the algorithm functions well with labeled reference datasets, then we can also assume that the walking detections from free-living dataset can be accepted. Since the evaluation of the free-living dataset was qualitative rather than quantitative, specific statistical hypotheses with p-values were not defined.

Materials and methods

Materials

The materials used in this study originate from two different studies: an open-access dataset provided by Reyes-Ortiz et al. [31] from 30 healthy adults, and two datasets from the Käveli project [32]. The complete measurement protocol of the Käveli project (NCT 03366558) conducted at Satakunta Wellbeing County in Pori, Finland, is described by Jauhiainen et al. [32]. This dataset consists subjects having PD and healthy control subjects. Out of 114 subjects 97 were included in both Dataset B and C, five subjects only in Dataset B, ten subjects only in

Dataset C, and two subjects in neither of the Datasets. These exclusions were caused by missing data. All patients were able to walk without another person aiding them, which was also an inclusion criteria for the study [32].

The aim of using different datasets was to first validate the walking detection algorithm with two types of annotated data that contain walking before experimenting with the performance on free-living and non-annotated data.

Dataset A: Human Activities and Postural Transitions

"Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set" (HAPT, referred as Dataset A from now on) [31] is an open-access dataset available at the University of California Irvine Machine Learning repository [33]. It contains smartphone (Samsung Galaxy S II) IMU recordings from 30 healthy adult volunteers aged 19–48 years [33]. The dataset has 61 raw triaxial acceleration signals recorded at 50 Hz, and the unit used for acceleration is the gravity of Earth, g [31].

As Reyes-Ortiz et al. [31] describe, the measurement protocol consists of six basic activities: standing, sitting, lying, walking, walking downstairs, and walking upstairs. They have also recorded transitions between the static postures as follows: stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand. Walking activities were recorded for 115.4 minutes, whereas the total length of the recorded data was 271.5 minutes.

Dataset B: Käveli Clinic Dataset

The Käveli clinic dataset (Dataset B from here on) consists of structured walking tests. The subjects walked 20 steps in a straight hallway as instructed by the study physiotherapist. The test was performed two times for each participant [32].

This study used a triaxial accelerometer from a Nokia 6 smartphone sampled at 50–100 Hz [32]. The same setup was used for Dataset C. This dataset consists of 188 successful walking tests, collected from 102 subjects. The demographics of this dataset are presented in Supplementary Material A.

Dataset C: Käveli Free-Living Dataset

Dataset C contains free-living data from each subject for 72 hours [32]. The sensors were not worn during sleep or during water activities, or when the phone battery was being charged. This dataset consists of 107 subjects. The demographics of this dataset are presented in Supplementary Material A.

Free-living data may contain any activities the subjects were performing during the recording period, and their behavior was not guided or recorded in any way to minimize burden especially for the PD patients and any effects to their activities. Free-living data may also contain noisy segments, e.g. when the smartphone pouch was put on or taken off. However, these segments should be very short interruptions in the data, and they should not interfere with the walking detection significantly.

Methods

All data processing and analysis steps have been done with Python 3.7 using standard libraries. In summary, the data processing consists of pre-processing the raw data and implementing the mean amplitude deviation algorithm for the acceleration data. Segments fulfilling the set thresholds are classified as walking segments.

All datasets were preprocessed with a Butterworth lowpass filter with a cut-off frequency of 20 Hz and they were expressed in unit g. The resultant acceleration vector was calculated from the triaxial signal and thus, the vector is independent of the

device orientation. The resultant vector r is defined as follows:

$$r_i = \sqrt{(x_i^2 + y_i^2 + z_i^2)}. \quad (1)$$

with x_i, y_i, and z_i being the acceleration at time *i* in the corresponding orientations of the acceleration sensor.

The resultant acceleration is further used in calculating MAD. MAD is the average of the absolute difference between samples and their neighboring mean within the segment

$$MAD = \frac{1}{n}\Sigma|r_i - \bar{r}|. \quad (2)$$

In Equation 2, r_i is the resultant vector from *i*th sample within the segment, $\bar{r}$ is the mean of the resultant acceleration vector over the fixed window including sample *i*, and *n* is the number of samples in the segment. Thus, the resulting value of MAD for one window is the mean of these remainders over the selected segment [16,34]. A fixed non-overlapping calculation window of 5 seconds was used in the analysis based on earlier studies [15,25] for all datasets.

We set the threshold values based on earlier studies, and additional discussions with other data science professionals in the field. We set the lower limit for walking to 50 mg_0 and the upper limit to 600 mg_0 to filter out sedentary activities such as standing and sitting, and vigorous activities such as running. We estimated that PD patients might move more slowly than healthy people, thus the lower limit was also decreased.

The results of Dataset A and B are analyzed using common performance descriptors: accuracy, balanced accuracy, precision, sensitivity, specificity, F1 score, and error rate. F1 score is the harmonic mean of precision and recall, and it represents the predictive performance of the model. The performance was evaluated against ground truth, for Dataset A it was available within the dataset. For Dataset B, the authors visually classified each segment by determining that at least 50% of the segment needs to be in the selected category. The results are compared with existing literature.

Dataset C cannot be explored through statistical measures, since no ground truth is available of how much walking the participants have conducted. Instead, we visualize the amount of detected walking and inspect examples of walking distribution on different measurement days.

Results

Results for Datasets A and B

When dividing data into 5-second non-overlapping segments, Dataset A had 3507 and Dataset B 858 5-second segments, respectively. For Dataset A the classification results were 1150 true positive, 1765 true negative, 0 false negative, and 592 false positive predictions. For Dataset B, the results were 546 true positive, 249 true negative, seven false negative, and 56 false positive predictions. Classification results are summarized in Table 1, and visual examples of the classification results are in Figure 1.

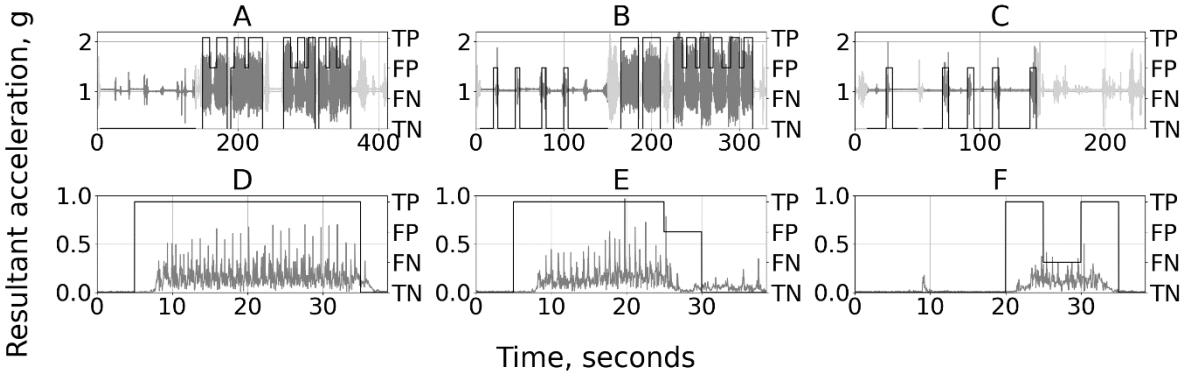

Figure 1. 6 Examples of classification results in Dataset A (A–C) and Dataset B (D–F). TN= True negative, FN=false negative, FP=false positive, and TP=true positive. For Dataset A the segments without a reference label are plotted in lighter gray since segments with missing label are not included in the analysis.

Table 1. Performance measures of walking detection results. "N = the number of five second segments and PD = Parkinson's disease"

Measure	Dataset A (n = 3507)	Dataset B (all n = 858, PD n = 535, Control n = 323)		
	All (%)	All (%)	PD (%)	Control (%)
Accuracy	83.1	92.7	92.5	92.9
Balanced accuracy	87.4	90.2	90.1	90.4
Sensitivity (Recall)	100.0	98.7	98.0	100.0
Specificity	74.9	81.6	82.2	80.8
Precision	66.0	90.7	91.2	89.8
F1 score	79.5	94.5	94.5	94.6

MAD is performing well for both datasets based on the performance descriptors. Accuracy and sensitivity are high in both datasets, but the higher number of false positives in Dataset A shows in specificity, precision and therefore, the F1 score. There is no large difference between PD patients and healthy controls (Dataset B) in terms of sensitivity and specificity. Sensitivity was higher for the control group but specificity for the PD group.

The percentage of false positives in Dataset A is 16.9% and in Dataset B 6.5%. In Dataset A (Figure 1A–C), the false positives come from the short segments between walking tests, where labels were not used. Another source for false positive detection is postural changes, which can be seen in Figures 1B and 1C as segments with lower resultant acceleration than walking. Dataset A does not have any false negatives; thus, every segment containing walking was detected.

Dataset B (Figure 1D–F) shows correctly classified walking tests, false positive segments in the beginning or end of the walking segment where walking is present less than 50% of the segment, and false negative segment in the middle of continuous walking. Dataset B had seven false negative segments, which occurred amongst two PD subjects in the

dataset. In both cases the resultant acceleration seems smaller than for other subjects, as can be seen in Figure 1F.

Results for dataset C

Figure 2 presents a summary of detected walking concerning the time of the day using a 24-hour clock. As expected for normal daily activities, most of the walking occurs during the day (7:00–23:00), and the median walking time is less than ten minutes per hour. However, only few subjects walked more than 30 minutes per hour and thus, those are marked as outliers in the figure. Outliers are defined as 1.5 times the interquartile distance. Several walking segments were detected during the night-time (Midnight to 7:00) for durations of less than 20 minutes. It should be noted that one outlier circle in Figure 2 denotes one person's walking sum during the whole three-day recording. Some individuals reported sleeping issues, and some control subjects may also work in shifts. Thus, they may be more active during unconventional hours. Similar figures were also drawn for each person individually showing that most of the subjects only walked during daytime, and the measurement phone was charged at night.

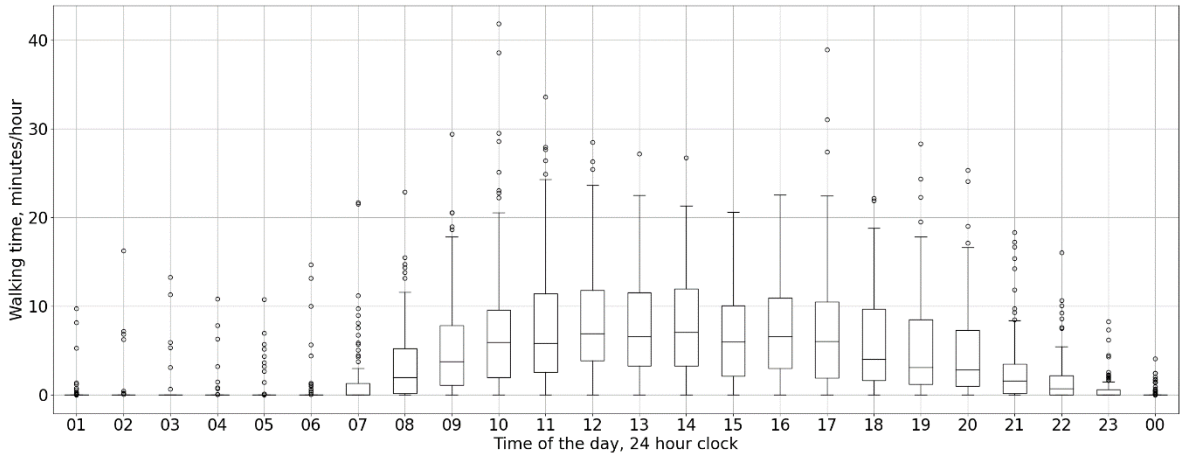

Figure 2. Detected walking distribution for all subjects in Dataset C with respect to a 24-hour clock. Boxes denote the middle quartiles (line is median), whiskers denote the first and fourth quartile, and dots denote outliers defined as 1.5 times the interquartile distance. Walking sums per hour were calculated for all subjects during all measurement days.

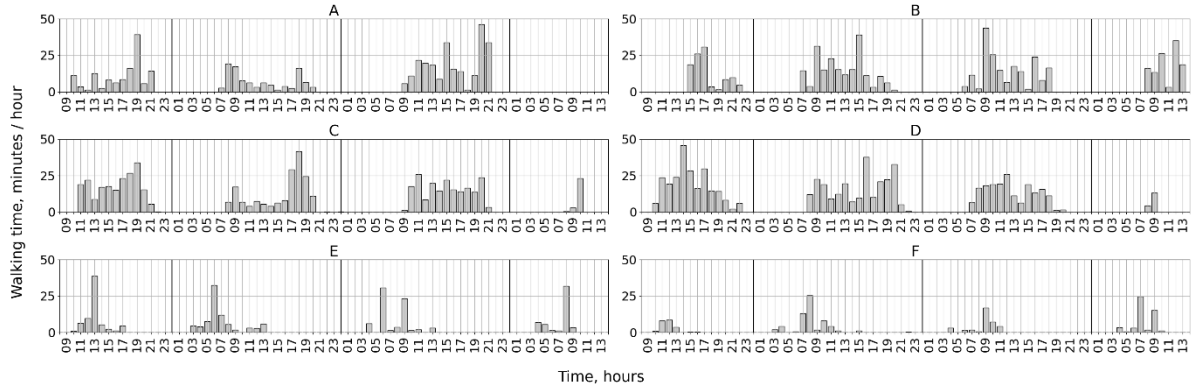

Figure 3. 72-hour free-living data from six example subjects. The subjects in the left column are Parkinson's disease patients, and the subjects in the right column are control subjects. Graphs A–D include four subjects that have walking during the day and not walking during the night. Graphs E–F include two examples present walking also during the night, approximately starting at 3:00. One of these two subjects had reported that they are usually awake during the night. Vertical lines denote the changing of the day at midnight.

Figure 3 shows the 72-hour walking statistics for six example subjects, three of which were PD patients (Figure 3 A, C, and E) and three healthy control subjects (Figure 3 B, D, and F). two subjects show walking at night, starting at 3:00–4:00 (Fig 3. E–F). The other four subjects have walking data collected during the daytime only. Walking activity on the first and last day may depend on the starting time of the measurement, which varied between the subjects.

Discussion

The objective of this study was to validate a walking detection algorithm for analyzing free-living gait of PD patients and healthy control subjects. Even though a step detection algorithm was implemented in our previous study [27], we concluded that an intensity-based walking detection algorithm is more suitable for long, free-living recordings. Also, as Brajdic and Harle [11] discovered that if the walking detection algorithm produces false positive classifications, the step recognition algorithms also tend to miscalculate the number of steps. Dijkstra et al. [14] also reported that slow walking speed and short trajectories cause an underestimation of steps and an overestimation of gait duration. The detection of walking segments instead of individual steps was justified with existing literature, especially when considering that the walking speed of PD patients might be reduced both due to progression of the disease as well as aging of the patients [1].

Algorithms for walking detection can be based on directly implementing thresholds on a straightforward calculation from the signal [11]. Our approach was to build a simple system with adjustable threshold values without needing extensive model training. Previous literature supported using MAD algorithm for physical activity recognition tasks in different study populations and regardless of the accelerometer brand [15,16,21]. The novelty and uniqueness of this study is to apply and evaluate MAD to PD patients and especially walking data instead of only healthy and young subject during vigorous physical activity, as in prior literature. Also, since MAD is based on calculating the resultant vector of the triaxial acceleration, the device

orientation does not affect the results, as was also noted by Dijkstra et al. [14].

According to existing literature, walking detection for healthy subjects resulted in varying accuracies, the best was the Hidden Markov model tested by Brajdic & Harle [11] with <2.0% error rate with threshold-based methods. Others also received relatively high results [4,20]. Wang et al. [35] used the same dataset (Dataset A) for a neural network and received a recognition rate of 95.9%. When considering the existing literature for PD patients, a high accuracy (96.0%) was obtained in the study by Micó-Amigo et al. [19].

Compared to earlier literature, our algorithm worked well on supervised datasets. The number of false negatives was very low in both datasets, and the number of false positives in Dataset B was also small. The accuracy of Dataset B outperformed few of the earlier studies [4,13,14,36]. Because the results are comparable to the existing literature and the algorithm does not overestimate the number or length of the walking segments, we can expect the algorithm also to detect walking segments from free-living data.

Both controlled datasets presented imbalance between walking and non-walking segments: 32.8% and 64.5% of the 5 second segments were walking segments for Datasets A and B, respectively. This was due to the nature of the study protocols Dataset A contained six basic activities and static posture changes and therefore, more than 50% of the signals is something other than walking. On the other hand, Dataset B contains mostly walking and there are only short segments of standing in the beginning and end of each file. In Dataset A, balanced accuracy was higher than accuracy and in Dataset B it was lower. This indicates that if the class is underrepresented in the dataset the detection accuracy is also underestimated, and if the class is overrepresented the detection accuracy is therefore overestimated. However, the differences between accuracy and balanced accuracy in this study are not substantial.

When analyzing the results of Dataset C, the visual inspections of free-living data performed by the authors showed credible results regarding walking amounts and times. For example, walking was mainly performed during the day and the smartphones were charged at night. Subjects were walking after they left the clinic appointment, and subjects indicated having routines e.g., waking up roughly at the same time each day. Few active subjects also made additional and voluntary journal notes matching the results.

Overall, while the study provides valuable insights into the detection of walking of PD patients using free-living acceleration data, its limitations must be acknowledged when interpreting the results. The study was conducted on a specific population group, i.e., PD patients and their healthy control group (Datasets B and C). Therefore, the results may not be generalizable to other populations, such as younger individuals, professional athletes, or persons with other diseases affecting mobility. Not including reference annotations was already decided when planning the study protocol and data collection [32] so it was not possible to add them afterward. Regarding the noisy segments in the signals, it was estimated that they would be short if they were originating from putting on and taking off the sensor and would not affect walking detection. Finally, the mean age of PD patients was ten years older than for the control subjects (in Dataset B and C), which may cause differences in walking styles and pace and therefore affect the walking detection results. The results of controlled walking tests did not reveal that either of the groups was under

or over detected, but the difference in age must be considered when conducting future studies.

Future research should address the limitations and implement the proposed method with broader datasets. The algorithm could be developed to improve computational efficiency related to reading long data segments. With a validated walking detection algorithm for controlled and uncontrolled walking of PD patients, future research will focus on observing changes in PD related motor symptoms that affect walking over the follow-up period of three days. This could provide more insight e.g. to the medication response or other factors affecting the well-being of the patient at home. Objective analysis of uncontrolled walking may also help patients with cognitive challenges related to PD, since they may have difficulty reporting detailed events from the past weeks or months to the clinician.

Overall, this method could provide a robust alternative for walking detection to be used to quantitatively assess walking to provide insights into intervention needs or physical condition. Since the Käveli project [32] objectives emphasized a mobile and/or remote solution for assessing PD gait, the objectives of this study were also aligned with the initial project.

Conclusions

This study validated the use of MAD in walking detection in labeled and free-living data for PD patients and healthy controls. Both classification performance metrics for labeled datasets and visual inspection of the free-living dataset performed by the authors suggested that the method is feasible for walking detection in PD patients and healthy subjects (Datasets B and C). The results also aligned with previous literature. The results of this experiment will be later used for analyzing PD symptoms in walking in more detail to support clinical decision-making regarding medication responses or need for health care interventions.

Conflict of interest statement

The authors declare no conflict of interest.

Acknowledgements

Jari Ruokolainen was the director of KÄVELI project. Physiotherapist Terhi Pakkanen from Satasairaala Central Hospital, Wellbeing Services County of Satakunta, Pori, Finland, performed the clinical study protocol. Anu Holm, Saeed Mehrang, and Hannu Nieminen participated in the research plan and ethical application in Käveli project.

Funding

This work was supported by Tampere University, Orion Pharma Oy, Suunto Oy, Forciot Oy, and Business Finland – The Finnish Funding Agency for Technology and Innovation (5137/31/2016), all from Finland, funded the KÄVELI project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, Obeso J, Marek K, Litvan I, Lang AE, Halliday G, Goetz CG, Gasser T, Dubois B, Chan P, Bloem BR, Adler CH, Deuschl G. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2015 Oct;30(12):1591-601. https://doi.org/10.1002/mds.26424

[2] Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Poewe W, Sampaio C,

Stern MB, Dodel R, Dubois B, Holloway R, Jankovic J, Kulisevsky J, Lang AE, Lees A, Leurgans S, LeWitt PA, Nyenhuis D, Olanow CW, Rascol O, Schrag A, Teresi JA, van Hilten JJ, LaPelle N; Movement Disorder Society UPDRS Revision Task Force. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord. 2008 Nov 15;23(15):2129-70. https://doi.org/10.1002/mds.22340

[3] Cunningham L, Mason S, Nugent C, Moore G, Finlay D, Craig D. Home-based monitoring and assessment of Parkinson's disease. IEEE Trans Inf Technol Biomed. 2011 Jan;15(1):47-53. https://doi.org/10.1109/TITB.2010.2091142

[4] Bao L, Intille SS. Activity recognition from user-annotated acceleration data. In: In: Ferscha A, Mattern F (eds). Pervasive Computing: Second International Conference, Pervasive 2004, Linz/Vienna, Austria, April 21-23, 2004. Lecture Notes in Computer Science, vol 3001. Berlin, Heidelberg: Springer; 2004. p. 1–17. https://doi.org/10.1007/978-3-540-24646-6_1

[5] Kang W, Han Y. SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization. IEEE Sens J. 2015;15(5):2906–16. https://doi.org/10.1109/JSEN.2014.2382568

[6] Zhao G, Wang X, Zhao H, Jiang Z. An improved pedestrian dead reckoning algorithm based on smartphone built-in MEMS sensors. AEU Inf J Electron C. 2023;168:154674. https://doi.org/10.1016/j.aeue.2023.154674

[7] Hebenstreit F, Leibold A, Krinner S, Welsch G, Lochmann M, Eskofier BM. Effect of walking speed on gait sub phase durations. Hum Mov Sci. 2015 Oct;43:118-24. https://doi.org/10.1016/j.humov.2015.07.009

[8] Quintero D, Lambert DJ, Villarreal DJ, Gregg RD. Real-Time Continuous Gait Phase and Speed Estimation from a Single Sensor. Control Technol Appl. 2017 Aug;2017:847-852. https://doi.org/10.1109/CCTA.2017.8062565

[9] Soltani A, Aminian K, Mazza C, Cereatti A, Palmerini L, Bonci T, Paraschiv-Ionescu A. Algorithms for Walking Speed Estimation Using a Lower-Back-Worn Inertial Sensor: A Cross-Validation on Speed Ranges. IEEE Trans Neural Syst Rehabil Eng. 2021;29:1955-1964. https://doi.org/10.1109/TNSRE.2021.3111681

[10] Mehrang S, Jauhiainen M, Pietil J, Puustinen J, Ruokolainen J, Nieminen H. Identification of Parkinson's Disease Utilizing a Single Self-recorded 20-step Walking Test Acquired by Smartphone's Inertial Measurement Unit. Annu Int Conf IEEE Eng Med Biol Soc. 2018 Jul;2018:2913-2916. https://doi.org/10.1109/EMBC.2018.8512921

[11] Brajdic A, Harle R. Walk detection and step counting on unconstrained smartphones. In: Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing (UbiComp '13). New York, NY, USA: Association for Computing Machinery; 2013. p. 225–34. https://doi.org/10.1145/2493432.2493449

[12] Iluz T, Gazit E, Herman T, Sprecher E, Brozgol M, Giladi N, Mirelman A, Hausdorff JM. Automated detection of missteps during community ambulation in patients with Parkinson's disease: a new approach for quantifying fall risk in the community setting. J Neuroeng Rehabil. 2014 Apr 3;11:48. https://doi.org/10.1186/1743-0003-11-48

[13] Pham MH, Elshehabi M, Haertner L, Del Din S, Srulijes K, Heger T, Synofzik M, Hobert MA, Faber GS, Hansen C, Salkovic D, Ferreira JJ, Berg D, Sanchez-Ferro Á, van Dieën JH, Becker C, Rochester L, Schmidt G, Maetzler W. Validation of a Step Detection Algorithm during Straight Walking and

Turning in Patients with Parkinson's Disease and Older Adults Using an Inertial Measurement Unit at the Lower Back. Front Neurol. 2017 Sep 4;8:457. https://doi.org/10.3389/fneur.2017.00457

[14] Dijkstra B, Zijlstra W, Scherder E, Kamsma Y. Detection of walking periods and number of steps in older adults and patients with Parkinson's disease: accuracy of a pedometer and an accelerometry-based method. Age Ageing. 2008 Jul;37(4):436-41. https://doi.org/10.1093/ageing/afn097

[15] Vähä-Ypyä H, Vasankari T, Husu P, Mänttäri A, Vuorimaa T, Suni J, Sievänen H. Validation of Cut-Points for Evaluating the Intensity of Physical Activity with Accelerometry-Based Mean Amplitude Deviation (MAD). PLoS One. 2015 Aug 20;10(8):e0134813. https://doi.org/10.1371/journal.pone.0134813

[16] Vähä-Ypyä H, Vasankari T, Husu P, Suni J, Sievänen H. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. Clin Physiol Funct Imaging. 2015 Jan;35(1):64-70. https://doi.org/10.1111/cpf.12127

[17] Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans Inf Technol Biomed. 2006 Jan;10(1):156-67. https://doi.org/10.1109/TITB.2005.856864

[18] Lee HH, Choi S, Lee MJ. Step Detection Robust against the Dynamics of Smartphones. Sensors (Basel). 2015 Oct 26;15(10):27230-50. https://doi.org/10.3390/s151027230

[19] Micó-Amigo ME, Bonci T, Paraschiv-Ionescu A, Ullrich M, Kirk C, Soltani A, Küderle A, Gazit E, Salis F, Alcock L, et al. Assessing real-world gait with digital technology? Validation, insights and recommendations from the Mobilise-D consortium. J Neuroeng Rehabil. 2023 Jun 14;20(1):78. https://doi.org/10.1186/s12984-023-01198-5

[20] Zhen T, Yan L, Yuan P. Walking gait phase detection based on acceleration signals using LSTM-DNN algorithm. Algorithms. 2019;12(12):253. https://doi.org/10.3390/a12120253

[21] Aittasalo M, Vähä-Ypyä H, Vasankari T, Husu P, Jussila AM, Sievänen H. Mean amplitude deviation calculated from raw acceleration data: a novel method for classifying the intensity of adolescents' physical activity irrespective of accelerometer brand. BMC Sports Sci Med Rehabil. 2015 Aug 7;7:18. https://doi.org/10.1186/s13102-015-0010-0

[22] Gao Y, Haapala EA, Vanhala A, Sääkslahti A, Rantakokko M, Laukkanen A, Pesola AJ, Rantalainen T, Finni T. Sedentary Thresholds for Accelerometry-Based Mean Amplitude Deviation and Electromyography Amplitude in 7-11 Years Old Children. Front Physiol. 2019 Aug 7;10:997. https://doi.org/10.3389/fphys.2019.00997

[23] Gao Y, Rantalainen T, Finni T, Portegijs E, Eronen J, Rantanen T, Rantakokko M. Daily Physical Activity and Sedentary Time Assessed by Acceleration Based on Mean Amplitude Deviation among Older People. Int J Environ Res Public Health. 2020 Sep 21;17(18):6887. https://doi.org/10.3390/ijerph17186887

[24] Bakrania K, Yates T, Rowlands AV, Esliger DW, Bunnewell S, Sanders J, Davies M, Khunti K, Edwardson CL. Intensity Thresholds on Raw Acceleration Data: Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) Approaches. PLoS One. 2016 Oct 5;11(10):e0164045. https://doi.org/10.1371/journal.pone.0164045

[25] Haapala EA, Rantalainen T, Hesketh KD, Rodda CP, Duckham RL. Accelerometer-based osteogenic indices, moderate-to-vigorous and vigorous

physical activity, and bone traits in adolescents. J Musculoskelet Neuronal Interact. 2022 Dec 1;22(4):514-523.

[26] Yoneyama M, Mitoma H, Okuma Y. Accelerometry-based long-term monitoring of movement disorders: from diurnal gait behavior to nocturnal bed mobility. J Mech Med Biol. 2013;13(2):1350041. https://doi.org/10.1142/S0219519413500413

[27] Juutinen M, Wang C, Zhu J, Haladjian J, Ruokolainen J, Puustinen J, Vehkaoja A. Parkinson's disease detection from 20-step walking tests using inertial sensors of a smartphone: Machine learning approach based on an observational case-control study. PLoS One. 2020 Jul 23;15(7):e0236258. https://doi.org/10.1371/journal.pone.0236258

[28] Açıcı K, Erdaş ÇB, Aşuroğlu T, Toprak MK, Erdem H, Oğul H. A random forest method to detect Parkinson's disease via gait analysis. In: Boracchi G, Iliadis L, Jayne C, Likas A (eds). Engineering Applications of Neural Networks. EANN 2017. Communications in Computer and Information Science, vol 744. Springer; 2017. p. 609–19. https://doi.org/10.1007/978-3-319-65172-9_51

[29] Aşuroğlu T, Açıcı K, Erdaş ÇB, Toprak MK, Erdem H, Oğul H. Parkinson's disease monitoring from gait analysis via foot-worn sensors. Biocybern Biomed Eng. 2018;38(3):760–72. https://doi.org/10.1016/j.bbe.2018.06.002

[30] Faiem N, Asuroglu T, Acici K, Kallonen A, Van Gils M. Assessment of Parkinson's Disease Severity Using Gait Data: A Deep Learning-Based Multimodal Approach. In: Särestöniemi M, et al. Digital Health and Wireless Solutions. Nordic Conference on Digital Health and Wireless Solutions 2024. Communications in Computer and Information Science, vol 2084. Springer; 2024. p. 29–48. https://doi.org/10.1007/978-3-031-59091-7_3

[31] Reyes-Ortiz JL, Oneto L, Samà A, Parra X, Anguita D. Transition-aware human activity recognition using smartphones. Neurocomputing. 2016;171:754–67. https://doi.org/10.1016/j.neucom.2015.07.085

[32] Jauhiainen M, Puustinen J, Mehrang S, Ruokolainen J, Holm A, Vehkaoja A, Nieminen H. Identification of Motor Symptoms Related to Parkinson Disease Using Motion-Tracking Sensors at Home (KÄVELI): Protocol for an Observational Case-Control Study. JMIR Res Protoc. 2019 Mar 27;8(3):e12808. https://doi.org/10.2196/12808

[33] UC Irvine. UCI Machine Learning Repository [Internet]. UC Irvine [cited 2022 Dec 7]. Available from: http://archive.ics.uci.edu/ml/index.php

[34] McClain JJ, Abraham TL, Brusseau TA Jr, Tudor-Locke C. Epoch length and accelerometer outputs in children: comparison to direct observation. Med Sci Sports Exerc. 2008 Dec;40(12):2080-7. https://doi.org/10.1249/MSS.0b013e3181824d98

[35] Wang H, Zhao J, Li J, Tian L, Tu P, Cao T, An Y, Wang K, Li S. Wearable Sensor-Based Human Activity Recognition Using Hybrid Deep Learning Techniques. Secur Commun Netw. 2020;2020(1):2132138. https://doi.org/10.1155/2020/2132138

[36] Marschollek M, Goevercin M, Wolf KH, Song B, Gietzelt M, Haux R, Steinhagen-Thiessen E. A performance comparison of accelerometry-based step detection algorithms on a large, non-laboratory sample of healthy and mobility-impaired persons. Annu Int Conf IEEE Eng Med Biol Soc. 2008;2008:1319-22. https://doi.org/10.1109/IEMBS.2008.4649407

Appendix. Supplementary material

Demographics of the Käveli dataset.

Table A1. Demographics of the Käveli dataset.

	Käveli clinic recordings (Dataset B)			Käveli free-living recordings (Dataset C)		
	PD	**Control**	**Sum**	**PD**	**Control**	**Sum**
N	62	40	102	68	39	107
UPDRS score (1–5)	1.7 ± 0.8			1.7 ± 0.8		
Age (years)	68.7 ± 8.5	58.7 ±14.1	64.8 ± 12.0	68.6 ± 8.4	60.3 ± 14.1	65.6 ±11.5
Gender (M | F)	34 | 28	7 | 33	41 | 61	38 | 30	8 | 31	46 | 61
Body mass index (kg/m^2)	26.4 ± 4.9	26.7 ± 4.3	26.5 ± 4.7	26.6 ± 4.8	26.8 ± 3.7	26.7 ± 4.4

Results expressed as mean ± standard deviation.
UPDRS = Universal Parkinson's Disease Rating Scale
PD = Parkinson's disease

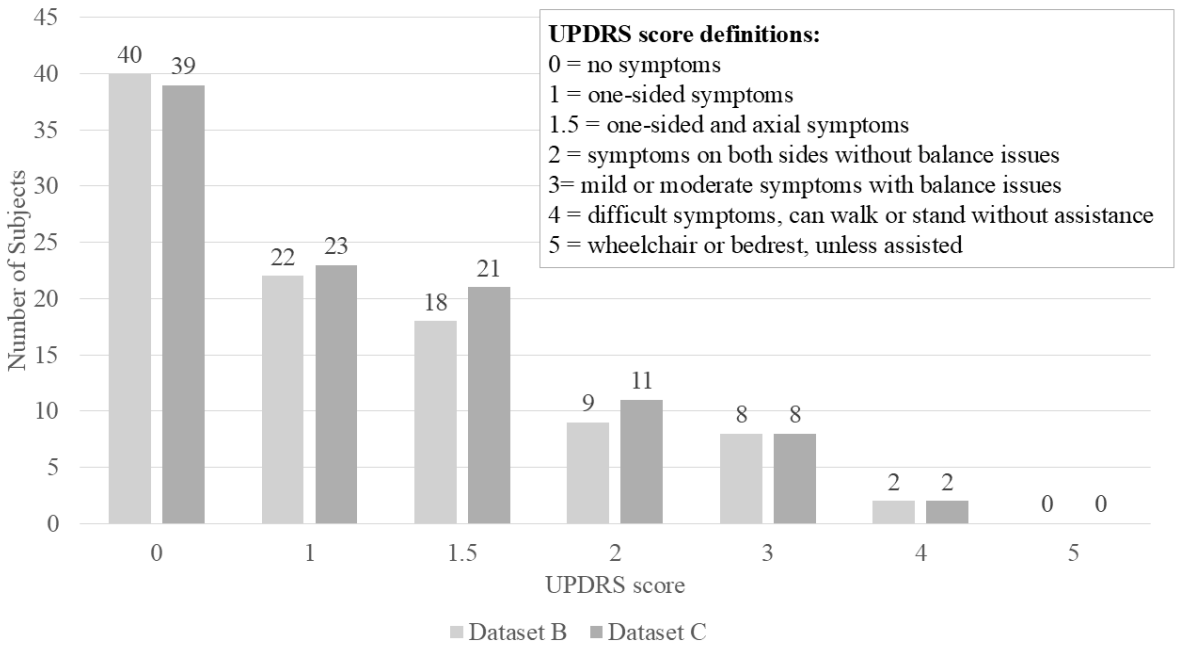

Figure A1. The distribution of Universal Parkinson's Disease Rating Scale (UPDRS) score in Dataset B and C. Score 0 (= no symptoms) was only achieved by control subjects. Patients with a score 5 (= not able to walk) were not included in the study.