Automated stage discrimination of Parkinson's disease

Vered Aharonson, Nabeel Seedat, Simon Israeli-Korn, Sharon Hassin-Baer, Michiel Postema, Gilad Yahalom

HAL Id: hal-03192434
https://hal.archives-ouvertes.fr/hal-03192434v2

Submitted on 16 Oct 2021

Automated Stage Discrimination of Parkinson's Disease

Vered Aharonson[1,2,*], Nabeel Seedat[1], Simon Israeli-Korn[3], Sharon Hassin-Baer[3,4], Michiel Postema[1] and Gilad Yahalom[3,4]

Abstract

Background: Treatment plans for Parkinson's disease (PD) are based on a disease stage scale, which is generally determined using a manual, observational procedure. Automated, sensor-based discrimination saves labor and costs in clinical settings and may offer augmented stage determination accuracy. Previous automated devices were either cumbersome or costly and were not suitable for individuals who cannot walk without support.

Methods: Since 2017, a device has been available that successfully detects PD and operates for people who cannot walk without support. In the present study, the suitability of this device for automated discrimination of PD stages was tested. The device consists of a walking frame fitted with sensors to simultaneously support walking and monitor patient gait. Sixty-five PD patients in Hoehn and Yahr (HY) stages 1 to 4 and 24 healthy controls were subjected to supported Timed Up and Go (TUG) tests, while using the walking frame. The walking trajectory, velocity, acceleration and force were recorded by the device throughout the tests. These physical parameters were converted into symptomatic spatiotemporal quantities that are conventionally used in PD gait assessment.

Results: An analysis of variance (ANOVA) test extended by a confidence interval (CI) analysis indicated statistically significant separability between HY stages for the following spatiotemporal quantities: TUG time ($p < 0.001$), straight line walking time ($p < 0.001$), turning time ($p < 0.001$), and step count ($p < 0.001$). A negative correlation was obtained for mean step velocity ($p < 0.001$) and mean step length ($p < 0.001$). Moreover, correlations were established between these, as well as additional spatiotemporal quantities, and disease duration, L-dihydroxyphenylalanine-(L-DOPA) dose, motor fluctuation, dyskinesia and the mobile part of the Unified Parkinson Disease Rating Scale (UPDRS).

Conclusions: We have proven that stage discrimination of PD can be automated, even to patients who cannot support themselves. A similar method might be successfully applied to other gait disorders.

Keywords

5-class discrimination, confidence interval analysis, Hoehn and Yahr stages, Parkinson's disease Gait characteristics, walker-mounted sensors.

[1]School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

[2]Department of Medical Engineering, Afeka, Tel-Aviv Academic College of Engineering, Israel

[3]The Movement Disorders Institute, Department of Neurology and Sagol Neuroscience Center, Chaim Sheba Medical Center, Tel-Hashomer, Israel

[4]Sackler Faculty of Medicine, Tel Aviv University, Israel

*Correspondence to:
Prof. Vered Aharonson, School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa.
E-mail: vered.aharonson@wits.ac.za

Introduction

The most common rating scales in Parkinson's disease (PD) are the Unified Parkinson Disease Rating Scale (UPDRS) and Hoehn and Yahr (HY) staging [1]. The HY 5-stages scale is the shorter of the two and primarily describes the progression of motor symptoms of PD [2]. This scale is based on the scenario that the motor symptoms of PD begin on one side of the body and then become bilateral, where compromise of balance/gait comes last. The HY scale thus grades PD progression, starting with a unilateral dysfunction (stage 1), following bilateral involvement, initially without postural instability (stage 2), then postural instability develops (stage 3) until physical independence is lost (stage 4), and at the terminal stage (stage 5) the patients become wheelchair bound or bed-ridden. The HY scale is weighted heavily toward postural instability, and does not sufficiently capture impairments or disability from other motor features of PD, such as manual dysfunction or tremor [3]. However, where gait disorders are examined, this scale can provide a disease stage description.

The staging of the HY scale involves subjective assessment of the examining physician. It may lead to inter-rater, and even to intra-rater variability [4]. Particularly, bias has been observed in the discrimination between stage 2 and 3 due to different skills and interpretation between different physicians. The inherent characteristics of the scale as categorical instead of numerical: The scores are not interval scales, hence distances between values on

Original Article

these scales are not quantified. The scale is non-linear in its description of progression between stages; i.e., a stage 1 PD subject who develops postural instability before developing bilateral signs must be rated as stage 3, having never been stage 2, additionally limits its capability on providing quantitative information. Last, but not least, the examination process involved in HY stage determination takes a considerable amount of clinicians' and other healthcare professionals' time and hence is extensive and expensive. This reduces the accessibility and affordability of this assessment to many patients.

The motor part of the UPDRS (mUPDRS) is a continuous numerical scale. This scale offers a more elaborate range of symptoms compared to the HY scale, and can complement the assessment of gait disorders [5]. However, mUPDRS still shares the limitations of the HY scale in its non-linear unquantified intervals between scores, as well as its length, labor and cost, and probable bias [1].

The Timed Up and Go (TUG) test, is an assessment tool, often coupled with the clinical HY scale to quantify the gait disorder in a shorter and simpler process. Initially introduced to assess functional mobility in the elderly [6] and in subjects during rehabilitation [7], this test has been proven instrumental for PD stage evaluation. In the standard TUG test, the subject is instructed to stand up from a chair, walk 3 m, turn back, walk 3 m, and sit down on the chair. Test completion time is measured with a stopwatch. This test has a potential to provide an objective measure of disease severity. However, the procedure still requires, the attention of a supervisor and relies on manipulation of a manual stopwatch. Moreover, as the TUG test measures only completion time, it does not quantify its different segments, like walking in a straight line and turning time which may provide a more complete gait characterization [8].

In view of the aforementioned limitations in prevalent PD severity scales, automated assessment tools were proposed. Automated assessments are inherently more objective and quantitative have the potential to aid in quantifying Parkinson's stage diagnosis and add to both the accuracy and efficiency of the assessment process.

Quantitative sensor-based methods were suggested in former studies to quantitatively asses gait disorders in PD. Many of these methods use the TUG test protocol, capture the subject's motion and provide quantitative models that discriminate PD subjects' gait from healthy control (HC) subjects' gait [9, 10]. The sensors used by these methods are either strapped on the patient's body [11, 12] or implemented as wearable sensors [13], or are fitted in walkway systems which measure the pressure exerted by the patient's foot as they walk [14, 15]. A drawback of the first two methods is their complexity, expense, and time demands. Additionally, these devices are often cumbersome and uncomfortable to wear, thereby negatively affecting the user's experience, especially for motor impaired persons [16, 17]. The walkway systems offer high accuracy and lower costs but require large physical space and a dedicated environment. All three methods are inappropriate for an assessment of severe cases of PD, when the patient requires a walking aid [16].

Previous sensor-based gait data acquisition methods to extract complex and abstract mathematical features from the sensors' outputs used machine learning tools for feature selection and discrimination. These computational analysis studies often used a combination of features, which could not be readily separated (i.e., using the principal component analysis) and interpreted [18–20]. This limits the usage of these methods for clinical use and for providing clinical insight into gait disorders in PD.

Previous sensor-based measurements of gait were used to discriminate PD patients from controls or to detect a specific symptom in PD gait, i.e., dyskinesia [21, 22]. Their analysis, however, aimed to distinguish between normal and impaired gait [23, 24] and did not attempt to assess disease severity or stage. One of the challenges involved in disease stage assessment is that statistical analysis methods can provide significance difference in terms of p-value, but this value is *not* indicative of the *magnitude* of differences between the different groups, nor does it quantify the amount of overlap between the groups.

The current study addresses all the aforementioned limitations. The data was acquired by an exo-body walking frame, fitted with sensors to monitor patient gait and support walking concurrently. This device offers a solution to the disadvantages of both strapped-on and walkway methods. Particularly, being a walking aid makes this device suitable for assessments of the severe stage of the disease, i.e., HY4. Preliminary results have shown that this device can provide accurate discrimination of PD patients and control subjects [16]. The measurements analysis in the current study considered only features which could be observed and related to the physical properties of the movement, and thus may provide an insight into the condition studied. Extended statistical analyses were employed to quantify these measurements' capability to discriminate the five HY stages of PD. Due to the inherent limitations of the HY scale, the automated analysis results were also tested for correlations with the mUPDRS and with complementary clinical data on the patients and their treatment.

Methods

Population

Sixty-six consecutive patients diagnosed with idiopathic PD according to the UK Bank criteria, attending a movement disorders institute at a tertiary medical center were recruited for the study. This patient cohort included stages 1 to 4 of the HY scale. Twenty-four healthy age-matched control subjects (HC, also designated as stage 0 of the HY scale) were recruited from the pool of hospital staff, patients' (unrelated) family members, caregivers or accompanying friends that arrived at the clinic. The exclusion criteria were: PD patients with additional neurologic disorders or any other disorder potentially affecting gait, patients who had had neurosurgical intervention for PD (such as deep brain stimulation and thalamotomy), patients with balance or gait disorder not related to PD, and patients with musculoskeletal problems causing gait impairment. The study was approved by the local institutional review board (IRB) of the Sheba Medical Center (Ethics number: 3036-16-SMC). All subjects signed an informed consent

form. Ethical approval for re-use of data was obtained from the University of the Witwatersrand, Human Research Ethics Committee, clearance number is M180202.

Instrumentation

The instrumented walker is an off-the shelve aluminum walker frame fitted with an instrumentation kit. The kit includes two force sensors underneath the hand grips that measure the grip force, two digital encoders on the walker's front wheels that measure the position and velocity of the walker and a tri-axial accelerometer in the control box of the walker (illustrated in **Figure 1** and described in details in [16]). An embedded microcontroller (Arduino Nano V3) in the control box executes the commands and control functionality and acquires the data at a sampling rate of 21.5 Hz. The data is written to a secured digital (SD) card in the form of a CSV file [16]. This data consists of the trajectory, velocity, acceleration, and force signals, which were recorded by the sensors throughout the subjects' walking experiment. The parameters computed from this data include the following spatiotemporal parameters: step count, mean step time, mean step length, mean step velocity, mean acceleration, standard deviation (STD) of step time, STD of step length, STD of step velocity, STD of acceleration, total TUG time, total walk time, total turn time, and cadence. The force sensors provided force, force difference between right and left force sensors and the correlation between right and left force sensors.

Protocol

The study was approved by the local IRB and all subjects signed an informed consent form. Each patient underwent a full neurological examination and was rated using part III (motor examination) of the UPDRS, yielding an mUPDRS score and the HY stage was determined. The presence of motor fluctuations and dyskinesia were specifically assessed and noted.
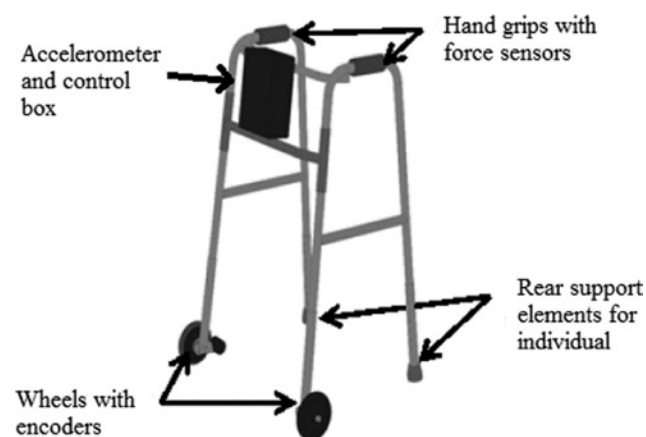

Figure 1 An illustration of the exo-body instrumented walker. The location of the encoders in the wheels, the pressure sensors in the hand grips and the accelerometer in the control box are marked on the figure [16].

All subjects underwent a TUG test while holding the instrumented walker: Subjects sat comfortably on a chair with no armrest and then spontaneously held on to the instrumented walker and stood up. The subjects then (holding the walker) walked at their natural speed straight ahead towards a cone positioned on the floor (3 m away from the start line), turned around the cone, walked back and then sat back down on the chair (still holding the walker). If a subject failed to perform the procedure correctly (e.g., due to poor understanding of the task or distraction), that trial was discarded and immediately repeated.

Data collection

Clinical data

The HY stage, mUPDRS score and the presence of motor fluctuations or dyskinesia were assessed and logged. Complementing clinical data including age at PD onset, disease duration and use of L-dihydroxyphenylalanine-(L-DOPA) in the medication regimen. Age and gender data were logged for all subjects.

Extracting features from the signals

Data analysis was performed on all the signals captured by the walkers' sensors. The preprocessing of the signals included noise and artifact removal, segmentation of the walking into strides in straight-line walking and turning, and footfall detection [25]. The signals were compressed into a set of mathematical variables, which represent the spatiotemporal parameters of gait, i.e., mean step time, mean step velocity. All these variables have been used in previous sensor-based studies on gait, and are easily interpreted into clinician observation of gait.

Statistical analysis

The study population represented five groups: PD patients according to HY stages 1–4 and HCs, which may be referred to as HY 0, respecting the hierarchical order between groups. The analysis aimed to determine the importance of each feature extracted from the signals, in terms of its discrimination performance of the five groups.

The first task in this analysis was to find the features which provide the highest differences between the groups: HY stages 0–4. The Kruskal–Wallis [one-way analysis of variance (ANOVA) on ranks] test was used to check for significant difference (p-value ≤ 0.05) between the five groups, for each variable, where the variables included both the demographic and clinical variables and the instrumented walker features.

A flaw in the ANOVA analysis methods is that their p-value is *not* indicative of the magnitude of the differences between the groups, nor does it indicate an overlap between the groups. The analysis was refined, using confidence

intervals (CIs), to estimate the probability that the *range* of values for a specific feature in one group does not overlap with the ranges of that feature in the other groups [26]. Plotting the CIs can provide a clear visual display of the overlaps and hence of the differences between multiple groups.

The difference between each pair of groups is based on the overlap of the two CIs belonging to two groups and is calculated as follows: an overlap greater than 50% corresponds to no statistical significance of difference; less than 50% corresponds to a 95% statistical significance of difference and no overlap corresponds to a 99% statistical significance of the difference [27].

Graphs of the CIs were plotted to illustrate the statistical differences between feature value ranges in the different HY groups, and CI overlaps computed.

Both p-values and CI-overlap values were used as metrics to determine the importance of each feature extracted from the walker signals. A ranking of the features according to these two measure was performed. The ranking indicates which features are the most informative in providing HY group separability.

Lastly, Pearson's correlation was employed to map the correlations between all pairs of instrumented-walker features and demographic/clinical variables. This correlation data and the corresponding p-values and significance (p < 0.05) were listed in a concluding table.

Results

The demographic and clinical characteristics of the study population are provided in **Table 1**. The demographic and/or clinical variables of the HCs (HY stage 0) and of the patients – HY stages 1–4 – are presented in the first five columns, respectively. The last column provides the p-values of the Kruskal–Wallis test for each characteristic, for the five groups (HY stages 0–4). The table conveys that age, disease duration, and the prevalence of L-DOPA treatment and motor fluctuations significantly increase with HY stage.

Table 2 lists the features extracted from the instrumented walker signals in the 3 m TUG test, in a format similar to **Table 1**: mean and STD of the extracted features are presented, for the five HY groups. The sixth displays the p-values of the Kruskal–Wallis test for each feature, for the five groups (HY stages 0–4).

Table 3 presents the pairwise CI overlap percentages for the first six gait features in **Table 2**. Zero overlaps, corresponding to a 99% statistical significance of the difference are marked by two asterisks. Overlaps of less than 50%, corresponding to a 95% statistical significance of difference, are marked by one asterisk. All other entries have overlaps larger than 50%, corresponding to statistically insignificant difference.

Table 1 Demographic and Clinical Characteristics of the Five HY Groups and p-Values Representing the Kruskal–Wallis Analyses of the Five Groups

	HY0 (control)	HY1	HY2	HY3	HY4	p-Value
Number of subjects	24	7	23	29	6	
Males (%)	10 (42)	5 (71)	17 (74)	19 (66)	3 (50)	0.270
Age in years	62.2 ± 13.3	59.4 ± 13.8	65.3 ± 10.3	70.0 ± 8.1	76.3 ± 6.5	**0.001**
Disease duration (years)	NA	3.1 ± 2.1	8.2 ± 4.4	8.8 ± 4.4	11.8 ± 5.0	**0.003**
L-DOPA treated (%)	NA	2 (29)	14 (61)	26 (90)	5 (83)	**0.003**

HY: Hoehn and Yahr scale; L-DOPA: L-dihydroxyphenylalanine. Bold value denote statistical significance of 95% or more.

Table 2 Mean and STD of the Features Extracted in the Instrumented Walker 3 m TUG Test. The First Five Columns Provide Mean ± STD Values for the Five HY Groups and the Sixth Displays the Computed ANOVA p-Value of the Difference Between the Groups' Means

	HY0 (Control)	HY1	HY2	HY3	HY4	p-Value
Total TUG time (s)	12.7 ± 4.2	11.7 ± 4.5	12.6 ± 4.0	19.9 ± 7.0	26.1 ± 12.8	**<0.001**
Straight-line walking time (s)	8.7 ± 3.7	8.6 ± 4.2	11.0 ± 3.9	13.3 ± 5.6	17.9 ± 10.9	**<0.001**
Turning time (s)	4.9 ± 2.4	3.2 ± 0.7	3.9 ± 1.3	7.2 ± 2.0	8.9 ± 3.7	**<0.001**
Mean step length (m)	0.5 ± 0.1	0.5 ± 0.2	0.4 ± 0.1	0.3 ± 0.1	0.2 ±0.06	**<0.001**
Mean step count (n)	14.0 ± 5.8	15.0 ± 6.3	17.0 ± 6.6	22.0 ± 10.0	47.0 ± 18.2	**<0.001**
Mean step velocity (m/s)	0.6 ± 0.2	0.7 ± 0.2	0.5 ± 0.15	0.4 ± 0.1	0.3 ± 0.09	**<0.001**
Step length variability (m)	0.7 ± 0.3	0.6 ± 0.2	0.6 ± 0.2	0.8 ± 0.3	0.6 ± 0.2	**0.03**
Mean acceleration (m/s^2)	0.03 ± 0.05	0.03 ± 0.04	0.03 ± 0.03	0.01 ± 0.03	0.01 ± 0.02	0.73
Step velocity variability (m/s)	0.3 ± 0.09	0.2 ± 0.06	0.3 ± 0.08	0.3 ± 0.09	0.3 ± 2.0	0.32
Step time (s)	0.7 ± 0.1	0.6 ± 0.02	0.7 ± 0.3	0.7 ± 0.07	0.7 ± 0.02	0.46
Step time variability (s)	0.3 ± 0.2	0.2 ± 0.03	0.4 ± 0.3	0.3 ± 0.2	0.4 ± 0.05	0.51
Cadence (step/s)	100.7 ± 9.6	107.1 ± 9.9	100.0 ± 8.8	98.2 ± 7.2	18.5 ± 3.1	0.25
Force sensor asymmetry (N)	30.1 ± 16.8	23.8 ± 6.0	32.8 ± 16.1	28.0 ± 17.8	23.8 ± 20.6	0.58
Force sensor asymmetry variability (N)	0.9 ± 0.2	0.9 ± 0.2	0.8 ± 0.2	0.9 ± 0.20	1.0 ± 0.1	0.305

HY: Hoehn and Yahr scale; N: Newtons; STD: standard deviation; TUG: Timed Up and Go test. Bold value denote statistical significance of 95% or more.

Table 3 The Overlap of CI, in Percentage, for the HY and HC Group Pairs. Overlaps of Less Than 50% are Marked by One Asterisk and Reflect 95% Significance of Difference Between the Groups. No Overlap (0%) Reflects 99% Significance of Difference Between the Groups

Group Pair	Feature					
	Turning Time (s)	Step Velocity (m/s)	Step Length (m)	Straight Line Walk Time (s)	Step Count	TUG Time (s)
HC–HY1	0**	19*	73	96	84	46
HC–HY2	21*	28*	36*	83	77	92
HY1–HY2	0**	0**	8*	75	91	71
HC–HY3	0**	0**	0**	0**	0**	0**
HC–HY4	0**	0**	0**	0**	0**	0**
HY1–HY3	0**	0**	0**	0**	0**	0**
HY1–HY4	0**	0**	0**	0**	0**	0**
HY2-HY3	0**	0**	0**	31*	13*	0**
HY2–HY4	0**	0**	0**	0**	0**	0**
HY3–HY4	33*	0**	0**	0**	0**	0**

CI: confidence interval; HC: healthy controls; HY: Hoehn and Yahr stage; TUG: Timed Up and go test. *Overlap corresponds to a 95% statistical significance of difference. **Overlap corresponds to a 99% statistical significance of difference.

The first three rows of **Table 3** indicate no statistical difference between the controls (HY stage 0) and HY stages 1 and 2 groups. The subsequent rows in the table demonstrate a 95% to 99% statistically significant difference between all other pairs of HY groups, for all features.

A visual display of the CI analysis is presented in **Figure 2**. The six graphs in the figure correspond to the six features in **Table 3**. The overlap or lack of overlap between groups can be observed in this graphical representation, as well as the range of values for each feature and for each group. These figures convey that the controls – HY stage 0 – and HY stage 1 patients have a higher step velocity and a greater step length, lower step count, slower straight-line walking, shorter turning time and shorter total TUG time. HY stages 2, 3 and 4 demonstrate increasingly lower step velocity, smaller step length, higher step count, and higher mean straight-line walking, turning, and total TUG time in all six graphs of

Figure 2. Patients with HY stage 4 differ significantly in all gait features from all other groups.

Table 4 presents the correlations between the pairs of demographic/clinical variables, including the HY stage and the mUPDRS, and the gait features. The rows in the table are the demographic and clinical variables and the columns are the instrumented-walker features. Each cell contains the Pearson's correlation coefficient and p-value for a pair of demographic/clinical and instrumented-walker features. Statistical significance is marked in bold numbers.

PD duration is inversely correlated with mean step length ($p = 0.04$) and mean step velocity ($p = 0.02$) and is correlated with mean step count ($p = 0.03$), TUG time ($p = 0.03$) and TUG turning time ($p = 0.03$). Mean TUG straight line walking time is not correlated to PD duration. TUG time, mean step count, mean step length and mean step velocity, show correlations to L-DOPA dose, which are similar to the correlations of these features to PD duration: L-DOPA dose
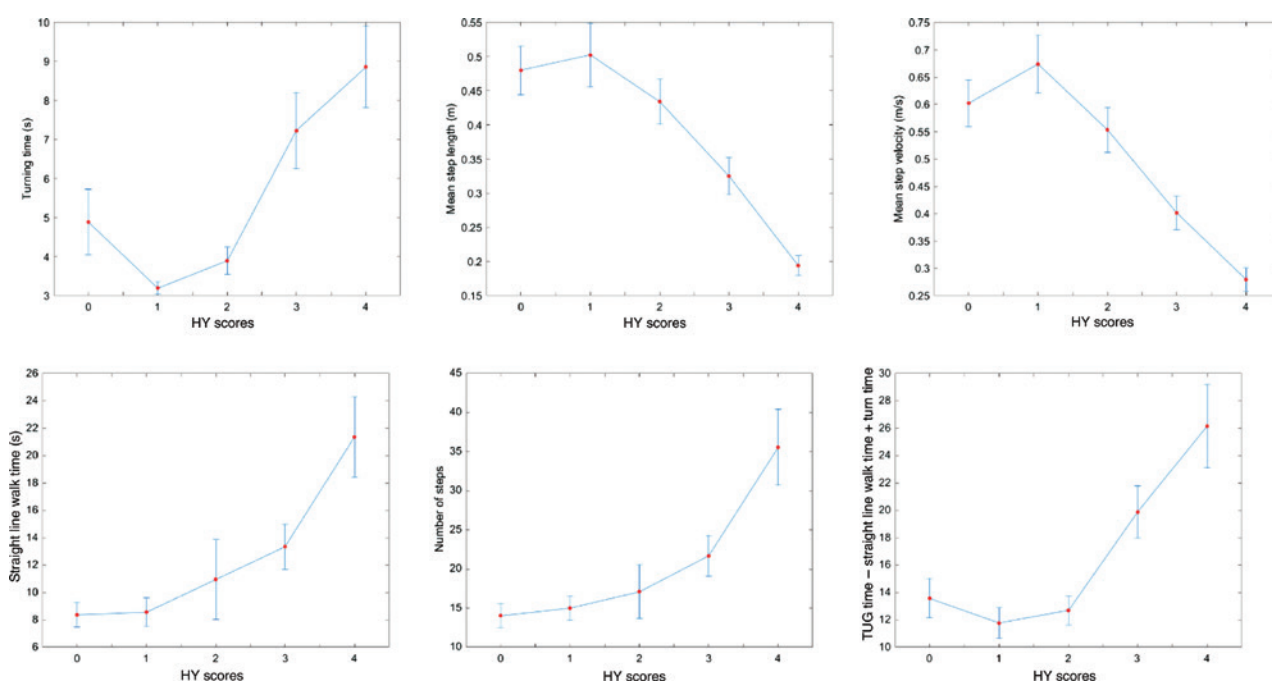

Figure 2 CI graphs of six instrumented gait features, for the five different HY groups. CI: confidence interval; HY: Hoehn and Yahr scale.

Table 4 Pearson's Correlation R-Values and p-Values (in Parenthesis) Between Sensor-acquired Gait Features Extracted from the 3 m TUG test, and Clinical Data and Observations

	Total TUG Time (s)	Straight Line Walking Time (s)†	Turning Time (s)	Mean Step Length (m)	Mean Step Count (n)	Mean Step Velocity (m/s)	Cadence (step/s)	Mean Acceleration (m/s²)	Force Sensor Asymmetry (N)	Step Length Variability (m/s)	Step Velocity Variability (m/s)	Force Sensor Asymmetry Variability (N)	Step Time Variability (s)	Mean Step Time (s)
Age	**0.349** **(0.006)**	**0.378** **(0.003)**	0.214 (0.10)	-0.155 (0.23)	0.201 (0.12)	-0.126 (0.33)	-0.03 (0.82)	0.001 (0.99)	0.03 (0.79)	-0.082 (0.53)	0.056 (0.67)	0.067 (0.54)	0.078 (0.55)	-0.002 (0.99)
PD Duration	**0.274** **(0.03)**	0.196 (0.13)	**0.275** **(0.03)**	**-0.264** **(0.04)**	**0.376** **(0.003)**	**-0.3** **(0.02)**	-0.113 (0.39)	0.153 (0.24)	0.15 (0.24)	-0.027 (0.84)	0.121 (0.35)	-0.03 (0.76)	0.12 (0.36)	0.092 (0.48)
L-DOPA dose	**0.275** **(0.03)**	**0.28** **(0.03)**	0.187 (0.15)	**-0.337** **(0.008)**	**0.274** **(0.03)**	**-0.354** **(0.005)**	-0.16 (0.22)	0.193 (0.14)	**0.34** **(0.006)**	0.055 (0.67)	0.026 (0.84)	-0.15 (0.22)	-0.021 (0.87)	-0.062 (0.64)
m-UPDRS	0.291 (0.09)	**0.381** **(0.02)**	0.246 (0.15)	-0.101 (0.56)	**0.37** **(0.03)**	-0.114 (0.51)	**-0.468** **(0.005)**	0.231 (0.18)	0.24 (0.17)	0.13 (0.46)	0.156 (0.37)	-0.14 (0.44)	0.169 (0.33)	0.226 (0.19)
HY stage	**0.597** **(<0.001)**	**0.525** **(<0.001)**	**0.495** **(<0.001)**	**-0.575** **(<0.001)**	**0.504** **(<0.001)**	**-0.632** **(<0.001)**	-0.291 (0.20)	0.187 (0.15)	0.04 (0.71)	0.245 (0.06)	0.238 (0.06)	0.11 (0.29)	0.173 (0.18)	0.068 (0.60)

TUG, Timed Up and Go test; PD, Parkinson's disease; m-UPDRS, motor Unified Parkinson's Disease Rating Scale. Statistically significant values (95% confidence) are marked in bold font. †The straight line walking time is the time taken to complete the tug excluding turning time.

is correlated with step count (p = 0.03), TUG time (p = 0.03) and inversely correlated with mean step velocity (p = 0.005) and mean step length (p = 0.008). Mean TUG straight line walking time is correlated to L-DOPA dose (p = 0.03), while mean TUG turning time was not. Dyskinesias are correlated with mean step length only (p = 0.05). The mean of force difference between right and left force sensors is correlated with L-DOPA dose (p = 0.006) and with motor fluctuations (p = 0.03).

Age was correlated with TUG time (0.006) and mean TUG straight line walking time (p = 0.003).

From the two PD rating scales, mUPDRS is correlated with TUG straight line walking time (p = 0.02) and with step count (p = 0.03) and was inversely correlated with cadence (p = 0.005), while HY stage is correlated with turning time (p < 0.001), step count (p < 0.001), TUG time (p < 0.001) and walking time (p < 0.001) and is inversely correlated with mean step velocity (p < 0.001), and mean step length (p < 0.001).

Discussion

The method proposed in this study provided automated discrimination of the five HY stages in PD, where previous studies aimed to distinguish between normal and impaired gait or to detect a specific symptom in PD gait, i.e., dyskinesia [21–24] and were not applied for disease severity or stage assessment. Importantly, the gait characteristics which were used in these five classes of discrimination provide an easily-interpretable, quantitative insight into gait change with disease progression. A modified HY scale was introduced by Hoehn and Yahr, which included additional "mid-scale" values of 1.5 and 2.5 [28]. The experiment in this study was conducted in a hospital where a five-stage HY scale is employed. This scale was therefore used as the ground truth in our analysis. The discrimination between the HCs and the four HY stages, as provided by the gait features, was quantified using two statistical methods: The traditional correlation and p-value and an augmented CI analysis.

The correlations and p-values analysis indicated seven gait features which were significantly correlated to the HY stage **(Table 2)**. However, this analysis could provide only mean and STD values of these gait features for the different HY groups. Moreover, the p-values are not indicative of the magnitude of the differences between the different HY groups, nor does it quantify the amount of overlap between the groups.

The CI analysis method manifested both the range of gait feature values in each HY group and the amount of overlap between these groups **(Figure 2** and **Table 3)**. Additionally, the CIs can be portrayed in a graph **(Figure 2)** and provide an intuitive, "at a glance" illustration of the differences between the groups, and of the discrimination power of each gait feature. **Table 3** highlights the features that provided no overlap, reflecting 99% significance of difference, between the HY groups, and features that provided overlaps of less than 50%, reflecting 95% significance of difference

between the groups. This analysis conveys that HY stages 3 and 4 were significantly different from each other and from the earlier stages (HY1 and 2) and from the HCs, according to all six gait features. Not all the features were able to significantly discriminate the earlier stages, HY1, 2, and the HCs. This information was not apparent in **Table 2**, where only p-values were computed, provided only average discrimination statistics. This highlights the importance of looking beyond p-values in discrimination problems of multiple classes.

As its widespread usage and the prolific literature convey, the HY scale effectively identifies and discriminates the stages of PD. The PD patients' cohort in this study were labeled and grouped by their HY stage. The four groups (stages 1–4) were significantly different in terms of age, disease duration and L-DOPA treatment prevalence **(Table 1)**. Similar to previous studies, however, the patients' data in this study does not convey significant correlation of the HY grouping to clinically observed symptoms such as motor fluctuations and dyskinesia [29, 30].

The gait data acquired by the walker-mounted sensors provided insight into the features that could characterize observed symptoms and clinical information of the patients. **Table 4** portrays correlations between the sensor-acquired features and the clinical data, also incorporating the motor part of the UPDRS, L-DOPA usage and dose, motor fluctuations, and dyskinesia. Each one of the six features that were indicated in **Tables 2** and **3** were correlated with some, but not all, of the clinical data: The total TUG time was correlated to subject's age, PD duration, L-DOPA dose, in addition to the strong correlation to the HY stage that was indicated in the previous analysis. These findings corroborate the efficacy of the TUG time in PD progression, as well as imply the ability of the walker-mounted sensors to accurately capture it automatically. However, the TUG walking trajectory incorporates both straight line and turning walking phases. The separation of these two parts of the TUG and their quantification by the straight line walking time and turning time features yields additional information: the straight line walking time is significantly correlated with age and L-DOPA dose whereas turning time is significantly correlated with PD duration. This finding may imply that the straight line walking time within the TUG is affected more by age and medication whereas turning time is more indicative of disease progression, as manifested in a patient's gait.

The sensor-acquired data enables a measurement and computation of step length, step count, and step velocity statistics. PD duration is correlated with all three features.

The mUPDRS was correlated only with the straight line walking and the step count features. It was noted in previous studies [1] that these two scales, the HY and mUPDRS are not well correlated. The results corroborate this findings.

Unlike previous studies [18–20], the current analysis considered only features which could be observed and related to the physical properties of the movement, and thus may provide an insight into the condition studied. Additionally, this study focused an analysis of each gait feature separately. Each one of the features was considered separately and its relevance and contribution to the discrimination of

the five HY stages was quantified. This analysis therefore simplified the interpretation and makes it more useful clinically. **Table 4** shows, however, that additional features, such as cadence and imbalance in pressure on the walker's handle sensors may be also manifested in PD progression. These features, however, do not show significant correlation to the HY stage. Previous studies have indicated an enhanced performance of sensor-based gait features, when all features were jointly analyzed using machine learning. An analysis of the full feature-set will be performed when a larger sample is collected.

All these observations and interpretation are limited by the relatively small number of subjects, and need to be validated by a larger cohort. Particularly, more patients from HY stages 1 and 4 are needed to provide a balanced dataset. The study is a preliminary one and the first to attempt to perform a 5-stage discrimination of PD using the automated frame method. Two statistical analysis methods were implemented in order to show which of the features measured by the device are adequate to imply a significant discrimination between the stages. The results in this preliminary study indicate a potential to provide insights into the manifestation of gait features in PD progression. In addition, the analysis method of CIs overlaps employed in this study is indicated as a reliable and useful metric to convey discrimination between multiple groups.

The present analysis was undertaken to provide insight into the manifestations of gait characteristics in PD stages. In this analysis we deviate from the use of complex features and machine learning [7, 9–15, 18–20, 22–24] and, with an integrated team of expert neurologists and engineers, investigated the task of discriminating the five stages of PD severity, using only clinically-interpretable features [1–5] and explanatory statistical analysis [26, 27].

This approach differs from earlier ones by the following traits: 1) Only features that could be clinically interpreted and used to characterize the symptoms observed by clinical experts were included in the analysis. 2) Among these features were parameters from the turning segment of the walking test, whereas most previous sensor-based methods focused on straight line walking. 3) Each of these features was individually examined using statistical analysis, and 4) this statistical analysis employed in addition to the traditionally-used one-way ANOVA, an extended statistical analysis of CIs to better suit the more complex problem of multiple groups – the 5 HY stages.

The results indicate that six of the gait parameters were able to provide a statistically significant discrimination of HY stages, as well as value ranges for each of the stages, as conveyed by **Table 3** and **Figure 2**. Two of the parameters had 95%–99% in their discrimination between all the pairs of the PD stages, one yielded 95%–99% discrimination significance for all pairs except the pair of HC and HY1, and the other three yielded 95%–99% discrimination significance for all pairs except HC–HY1, HC–HY2 and HY1–HY2. All six parameters were affirmed by the healthcare professionals as relevant to their clinical observations of gait and to their manifestation in the observation-based determination of the HY stage. Moreover, the discrimination of HC from the first stages of the disease, based on gait only, was

Original Article

clinically observed as challenging using traditional observation methods [1–4].

Our method thus quantifies gait parameters which are routinely observed by healthcare professionals in clinical procedures, and the relationship between disease progression and the degradation in this gait parameters values and thus provides an improved insight into PD manifestation in gait compared to the aforementioned studies.

The prediction power of the six gait features that provided the strongest discrimination between all PD stage groups will be further explored and generalized in future studies employing larger cohorts of patients.

Therefore, although the data size is small, the results convey high statistical significance, implying that the recorded gait parameters and their correlation with HY stages can be used for the discrimination/prediction of HY stages in PD. These results should be validated on a larger cohort of patients. Our findings, however, provide a clinical value of these gait parameters, which is superior to other previously reported methods by introducing only clinically-interpretable parameters, explainable statistical analysis and a capability to discriminate five stages of disease severity.

Conclusion

The feasibility of PD stage discrimination by a simple and low-cost walker-mounted sensors method may offer a potential in rapid, labor-saving screening in the follow-up of PD patients, both in clinics and remotely. This automated method can update the patient status on a regular basis and provide warning of deterioration.

Acknowledgment

References

[1] Perlmutter JS. Assessment of Parkinson disease manifestations. Curr Protoc Neurosci 2009;49:10.1.1–10.1.14. [PMID: 19802812 DOI: 10.1002/0471142301.ns1001s49]

[2] Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. Neurology 1998;50:427–42. [PMID: 9484345 DOI: 10.1212/wnl.50.2.318]

[3] Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations the Movement Disorder Society Task Force on rating scales for Parkinson's disease. Mov Disord 2004;19:1020–8. [PMID: 15372591 DOI: 10.1002/mds.20213]

[4] Jacobs JV, Horak F. Abnormal proprioceptive-motor integration contributes to hypometric postural responses of subjects with Parkinson's disease. Neuroscience 2006;141:999–1009. [PMID: 16713110 DOI: 10.1016/j.neuroscience.2006.04.014]

[5] Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): process, format, and clinimetric testing plan. Mov Disord 2007;22:41–7. [PMID: 17115387 DOI: 10.1002/mds.21198]

[6] Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatrics Soc 1991;39:142–8. [PMID: 1991946 DOI: 10.1111/j.1532-5415.1991.tb01616.x]

[7] Galán-Mercant A, Barón-López FJ, Labajos-Manzanares MT, Cuesta-Vargas AI. Reliability and criterion-related validity with a smartphone used in timed-up-and-go test. Biomed Eng Online 2014;13:156. [PMID: 25440533 DOI: 10.1186/1475-925X-13-156]

[8] Greene BR, O'Donovan A, Romero-Ortuno R, Cogan L, Scanaill CN, et al. Quantitative falls risk assessment using the timed up and go test. IEEE Trans Biomed Eng 2010;57:2918–26. [PMID: 20923729 DOI: 10.1109/TBME.2010.2083659]

[9] Bonora G, Carpinella I, Cattaneo D, Chiari L, Ferrarin M. A new instrumented method for the evaluation of gait initiation and step climbing based on inertial sensors: a pilot application in Parkinson's disease. J Neuroeng Rehabil 2015;12:45. [PMID: 25940457 DOI: 10.1186/s12984-015-0038-0]

[10] Washabaugh EP, Kalyanaraman T, Adamczyk PG, Claflin ES, Krishnan C. Validity and repeatability of inertial measurement units for measuring gait parameters. Gait Posture 2017;55:87–93. PMID: 28433867 DOI: 10.1016/j.gaitpost.2017.04.013]

[11] Salarian A, Russmann H, Vingerhoets FJ, Dehollain C, Blanc Y, et al. Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. IEEE Trans Biomed Eng 2004;51:1434–43. [PMID: 15311830 DOI: 10.1109/TBME.2004.827933]

[12] Dewey DC, Miocinovic S, Bernstein I, Khemani P, Dewey 3rd RB, et al. Automated gait and balance parameters diagnose and correlate with severity in Parkinson disease. J Neurol Sci 2014;345:131–8. [PMID: 25082782 DOI: 10.1016/j.jns.2014.07.026]

[13] Mariani B, Hoskovec C, Rochat S, Büla C, Penders J, et al. 3D gait assessment in young and elderly subjects using foot-worn inertial sensors. J Biomechanics 2010;43:2999–3006. [PMID: 20656291 DOI: 10.1016/j.jbiomech.2010.07.003]

[14] Godinho C, Domingos J, Cunha G, Santos AT, Fernandes RM, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. J Neuroeng Rehabil 2016;13:24. [PMID: 26969628 DOI: 10.1186/s12984-016-0136-7]

[15] Lynall RC, Zukowski LA, Plummer P, Mihalik JP. Reliability and validity of the protokinetics movement analysis software in measuring center of pressure during walking. Gait Posture 2017;52:308–11. [PMID: 28033577 DOI: 10.1016/j.gaitpost.2016.12.023]

[16] Aharonson V, Seedat N, Schlesinger I, McDonald A, Dubowsky S, et al., editors. Feasibility of an instrumented walker to quantify treatment effects on Parkinson's patient gait. 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT). IEEE; 2018. [DOI: 10.1109/EBBT.2018.8391457]

[17] Alwan M, Felder R. Eldercare technology for clinical practitioners. Berlin, Germany: Springer Science & Business Media; 2007.

[18] Mazilu S, Hardegger M, Zhu Z, Roggen D, Tröster G, et al., editors. Online detection of freezing of gait with smartphones and machine learning techniques. 2012 6th International Conference on Pervasive Computing Technologies for Healthcare

(PervasiveHealth) and Workshops. IEEE; 2012. [DOI: 10.4108/icst.pervasivehealth.2012.248680]

[19] Tucker C, Han Y, Black Nembhard H, Lewis M, Lee W-C, et al. A data mining methodology for predicting early stage Parkinson's disease using non-invasive, high-dimensional gait sensor data. IIE Trans Healthcare Syst Eng 2015;5:238–54. [PMID: 29541376 DOI: 10.1080/19488300.2015.1095256]

[20] Cole BT, Roy SH, De Luca CJ, Nawab SH. Dynamical learning and tracking of tremor and dyskinesia from wearable sensors. IEEE Trans Neural Syst Rehabil Eng 2014;22:982–91. [PMID: 24760943 DOI: 10.1109/TNSRE.2014.2310904]

[21] Pulliam CL, Burack MA, Heldman DA, Giuffrida JP, Mera TO. Motion sensor dyskinesia assessment during activities of daily living. J Parkinson's Dis 2014;4:609–15. [PMID: 25208729 DOI: 10.3233/JPD-140348]

[22] Ossig C, Antonini A, Buhmann C, Classen J, Csoti I, et al. Wearable sensor-based objective assessment of motor symptoms in Parkinson's disease. J Neural Trans 2016;123:57–64. [PMID: 26253901 DOI: 10.1007/s00702-015-1439-8]

[23] Jarchi D, Pope J, Lee TK, Tamjidi L, Mirzaei A, et al. A review on accelerometry-based gait analysis and emerging clinical applications. IEEE Rev Biomed Eng 2018;11:177–94. [PMID: 29994786 DOI: 10.1109/RBME.2018.2807182]

[24] Cancela J, Pastorino M, Arredondo MT, Nikita KS, Villagra F, et al. Feasibility study of a wearable system based on a wireless body area network for gait assessment in Parkinson's disease patients. Sensors 2014;14:4618–33. [PMID: 24608005 DOI: 10.3390/s140304618]

[25] Seedat N, Beder D, Aharonson V, Dubowsky S, editors. A comparison of footfall detection algorithms from walker mounted sensors data. 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT). IEEE; 2018. [DOI: 10.1109/EBBT.2018.8391456]

[26] Baguley T. Serious stats: a guide to advanced statistics for the behavioral sciences. London, UK: Macmillan International Higher Education; 2012.

[27] Cumming G. Inference by eye: reading the overlap of independent confidence intervals. Stat Med 2009;28:205–20. [PMID: 18991332 DOI: 10.1002/sim.3471]

[28] Clarke CE, Patel S, Ives N, Rick CE, Woolley R, et al. Clinical effectiveness and cost-effectiveness of physiotherapy and occupational therapy versus no therapy in mild to moderate Parkinson's disease: a large pragmatic randomised controlled trial (PD REHAB). Health Technol Assess 2016;20:1–96. [PMID: 27580669 DOI: 10.3310/hta20630]

[29] Denny AP, Behari M. Motor fluctuations in Parkinson's disease. J Neurol Sci 1999;165:18–23. [PMID: 10426141 DOI: 10.1016/s0022-510x(99)00052-0]

[30] Schrag A, Quinn N. Dyskinesias and motor fluctuations in Parkinson's disease: a community-based study. Brain 2000;123:2297–305. [PMID: 11050029 DOI: 10.1093/brain/123.11.2297]