



**HAL**  
open science

## Automated stage discrimination of Parkinson's disease

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

► **To cite this version:**

Vered Aharonson, Nabeel Seedat, Simon Israeli-Korn, Sharon Hassin-Baer, Michiel Postema, et al.. Automated stage discrimination of Parkinson's disease. *BIO Integration*, 2020, 1 (2), pp.55-63. 10.15212/bioi-2020-0006 . hal-03192434v1

**HAL Id: hal-03192434**

**<https://hal.science/hal-03192434v1>**

Submitted on 11 Apr 2021 (v1), last revised 16 Oct 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Automated stage discrimination of Parkinson's disease

Vered Aharonson<sup>1,2</sup>, Nabeel Seedat<sup>1</sup>, Simon Israeli-Korn<sup>3</sup>, Sharon Hassin-Baer<sup>3,4</sup>, Michiel Postema<sup>1</sup>, Gilad Yahalom<sup>3,4</sup>

<sup>1</sup>School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa. <sup>2</sup>Department of Medical Engineering, Afeka, Tel-Aviv Academic College of Engineering, Israel. <sup>3</sup>The Movement Disorders Institute, Department of Neurology and Sagol Neuroscience Center, Chaim Sheba Medical Center, Tel-Hashomer; <sup>4</sup>Sackler Faculty of Medicine, Tel Aviv University, Israel.

## **Abstract**

Treatment plans for Parkinson's disease are based on a disease stage scale, which is generally determined using a manual, observational procedure. Automated, sensor based discrimination saves labour and cost 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.

Since 2017, a device has been available that successfully detects Parkinson's disease and operates for people who cannot walk without support. In the present study, the suitability of this device for automated discrimination of Parkinson's disease stages is tested. The device consists of a walking frame fitted with sensors to simultaneously support walking and monitor patient gait. Sixty-five Parkinson's disease patients in HYstages 1 to 4 and twenty-four 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 spatio-temporal quantities that are conventionally used in Parkinson's disease gait assessment.

An ANOVA Test extended by a confidence interval analysis indicated statistically significant separability between HYstages for the following spatio-temporal quantities: TUG time ( $p<0.001$ ), straight like 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 spatio-temporal quantities, and disease duration, levodopa dose, motor fluctuation, dyskinesia and the mobile part of the unified Parkinson's disease rating scale.

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

## **Introduction**

The most common rating scales in Parkinson's disease are the Unified Parkinson Disease Rating Scale (UPDRS) and the Hoehn and Yahr (HY) staging (1). The HY 5-stages scale is the shorter of the two and primarily describes the progression of motor 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 HYscale 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 terminal stage (stage 5) the patients become wheelchair bound or bedridden. The HYscale 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 HYscale 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 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 HYstage determination takes a considerable time of clinicians and other healthcare professionals and hence is extensive and expensive. This reduces the accessibility and affordability of this assessment to many patients.

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

The timed up and go (TUG) test, is an assessment tool, often coupled with the clinical HY 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 Parkinson's disease stage evaluation. In the standard TUG, the subject is instructed to stand up from a chair, walk 3 meters, turn back, walk 3 meters 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. The procedure still requires, however, the attention of a supervisor and relies on manual stopwatch manipulation. Moreover, as TUG measures only completion time, it does not quantify its different segments, like 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 accuracy and efficiency of the assessment process.

Quantitative sensor-based methods were suggested in former studies to quantitatively assess gait disorders in Parkinson's disease. Many of these methods use the TUG protocol, capture the subject's motion and provide quantitative models that discriminate PD subjects' gait from healthy control 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 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 extracted complex and abstract mathematical features from the sensors' outputs used machine learning tools for feature selection and discrimination. These computational analysis studies often used combinations 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 is 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 severe stage of the disease, ie 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 patients 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 after 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 of the Sheba Medical Center (Ethics number: 3036-16-SMC). All subjects signed an informed consent form. Ethics 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 shelf 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. 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 spatio-temporal 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 timed up and go (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 institutional review board 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 Unified PD rating scale (UPDRS, yielding a m-UPDRS score) and the Hoehn and Yahr (HY) stage was determined. The presence of motor fluctuations and dyskinesia were specifically assessed and noted.

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 meters 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*

HYstage, motor UPDRS 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 the straight-line walking and turning, and footfall detection (25). The signals were compressed into a set of mathematical variables, which represent the spatio-temporal parameters of gait, ie 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 5 groups: PD patients according to HYstages 1-4 and healthy controls, which may be referred to as HY0, 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: HYstages 0 to 4. The Kruskal Wallis (“one-way ANOVA on ranks”) test was used to check for significant difference ( $p\text{-value} \leq 0.05$ ) 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 Confidence Intervals 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 Confidence Intervals 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 HYgroups, 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 measures was performed. The ranking indicates which features are most informative in providing HYgroup 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 healthy controls (HYstage 0) and of the patients - HYstages 1 to 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 (HYstages 0 to 4). The table conveys that age, disease duration, and the prevalence of L-dopa treatment and Motor fluctuations significantly increase with HYstage.

Table 2 lists the features extracted from the instrumented walker signals in the 3 meter TUG test, in a format similar to Table 1: Mean and Standard deviation of the extracted features are presented, for the five HYgroups. The sixth displays the p-values of the Kruskal Wallis test for each feature, for the five groups (HYstages 0 to 4).

Table 3 presents the pair-wise 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.

The first three rows of table 3 indicate no statistical difference between the controls (HY stage 0) and HYstages 1 and 2 groups. The subsequent rows in the table demonstrate a 95% to 99% statistically significant difference between all other pairs of HYgroups, 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 figure conveys that the controls -HYstage 0 - and HYstage 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 HYstage 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 HYstage 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 feature. 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 Levodopa dose, which are similar to the correlations of these features to PD duration: Levodopa dose 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 Levodopa 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 levodopa 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 HYstage 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 five HY stage in PD, where previous studies aimed to distinguish between normal and impaired gait or to detect a specific symptom in PD gait, ie. Dyskinesia (21-24) and were not applied for disease severity or stage assessment. Importantly, the gait characteristics which were used in this five-classes 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 contain additional “mid-scale” values of 1.5 and 2.5 (28). A five-stage scale is, however, still widely used in clinical practice. The experiment in the study included a ground truth of clinicians’ HY stage determination. This was conducted as a part of the regular clinical assessment during the patients’ visit at the hospital, as described in the Methods section. This shorter scale was thus the one used for the the study, aiming at a preliminary feasibility of multi-class discrimination.

The discrimination between the healthy controls 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 confidence intervals analysis.

The correlations and p-values analysis indicated seven gait features which were significantly correlated to the HY stage (Table 2). This analysis could provide, however, only mean and standard deviation 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 confidence intervals analysis method manifested both the range of gait feature values in each HY group and the amount of overlap between these groups (fig. 2 and Table 3). Additionally, the confidence intervals can be portrayed in a graph (fig. 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 (HY 1 and 2) and from the healthy controls, according to all six gait features. Not all the features were able to significantly discriminate the earlier stages, HY 1, 2, and the healthy controls. 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 prolific literature convey, the HY scale effectively identify and discriminate the stages of PD. The PD patients’ cohort in this study were labeled and grouped by their HY stage. The four groups (stages 1 to 4) were significantly different in terms of age, disease duration and L-dopa treatment prevalence (Table 1). Similar to previous literature, 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, incorporating also the motor part of the UPDRS, Levodopa 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 age, PD duration, Levodopa 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. The TUG walking trajectory incorporates, however, 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 Levodopa 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 on 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 m-UPDRS 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 m-UPDRS 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 separately considered and its relevance and contribution to the discrimination of the five HYstages 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 HY. 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 the above observations and interpretation are limited by the relatively small number of subjects, need to be validated by on a larger cohort. Particularly, more patients from HY stages 1 and 4 are needed to provide a balance 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 measure 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 confidence intervals 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, 9-16, 18-20, 22-24) and, with an integrated team of expert neurologists and engineers, investigate the task of discriminating 5-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 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 confidence-intervals to better suit the more complex problem of multiple groups - the 5 Hoehn and Yahr 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 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 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 5 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**

The authors would like to thank Vered Livneh, Tsvia Fay-Karmon, Sandra Elinx-Benizri, Yarin Rubel and Aviran Ohayon who assisted in the data collection.

## References

1. Perlmutter JS. Assessment of Parkinson disease manifestations. *Current protocols in neuroscience*. 2009;49(1):10.1. 1-1. 4.
2. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology*. 1998;50(2):318-.
3. Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, 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. *Movement disorders*. 2004;19(9):1020-8.
4. Jacobs JV, Horak F. Abnormal proprioceptive-motor integration contributes to hypometric postural responses of subjects with Parkinson's disease. *Neuroscience*. 2006;141(2):999-1009.
5. Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): process, format, and clinimetric testing plan. *Movement Disorders*. 2007;22(1):41-7.
6. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *Journal of the American geriatrics Society*. 1991;39(2):142-8.
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. *Biomedical engineering online*. 2014;13(1):156.
8. Greene BR, O'Donovan A, Romero-Ortuno R, Cogan L, Scanail CN, Kenny RA. Quantitative falls risk assessment using the timed up and go test. *IEEE Transactions on Biomedical Engineering*. 2010;57(12):2918-26.
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. *Journal of neuroengineering and rehabilitation*. 2015;12(1):45.
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.
11. Salarian A, Russmann H, Vingerhoets FJ, Dehollain C, Blanc Y, Burkhard PR, et al. Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. *IEEE transactions on biomedical engineering*. 2004;51(8):1434-43.
12. Dewey DC, Miocinovic S, Bernstein I, Khemani P, Dewey III RB, Query R, et al. Automated gait and balance parameters diagnose and correlate with severity in Parkinson disease. *Journal of the neurological sciences*. 2014;345(1-2):131-8.
13. Mariani B, Hoskovec C, Rochat S, Büla C, Penders J, Aminian K. 3D gait assessment in young and elderly subjects using foot-worn inertial sensors. *Journal of biomechanics*. 2010;43(15):2999-3006.
14. Godinho C, Domingos J, Cunha G, Santos AT, Fernandes RM, Abreu D, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. *Journal of neuroengineering and rehabilitation*. 2016;13(1):24.
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.

16. Aharonson V, Seedat N, Schlesinger I, McDonald A, Dubowsky S, Korczyn A, editors. Feasibility of an instrumented walker to quantify treatment effects on Parkinson's patient gait. 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT); 2018: IEEE.
17. Alwan M, Felder R. Eldercare technology for clinical practitioners: Springer Science & Business Media; 2007.
18. Mazilu S, Hardegger M, Zhu Z, Roggen D, Tröster G, Plotnik M, 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; 2012: IEEE.
19. Tucker C, Han Y, Black Nembhard H, Lee W-C, Lewis M, Sterling N, et al. A data mining methodology for predicting early stage Parkinson's disease using non-invasive, high-dimensional gait sensor data. IIE transactions on healthcare systems engineering. 2015;5(4):238-54.
20. Cole BT, Roy SH, De Luca CJ, Nawab SH. Dynamical learning and tracking of tremor and dyskinesia from wearable sensors. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2014;22(5):982-91.
21. Pulliam CL, Burack MA, Heldman DA, Giuffrida JP, Mera TO. Motion sensor dyskinesia assessment during activities of daily living. Journal of Parkinson's disease. 2014;4(4):609-15.
22. Ossig C, Antonini A, Buhmann C, Classen J, Csoti I, Falkenburger B, et al. Wearable sensor-based objective assessment of motor symptoms in Parkinson's disease. Journal of neural transmission. 2016;123(1):57-64.
23. Jarchi D, Pope J, Lee TK, Tamjidi L, Mirzaei A, Sanei S. A review on accelerometry-based gait analysis and emerging clinical applications. IEEE reviews in biomedical engineering. 2018;11:177-94.
24. Cancela J, Pastorino M, Arredondo MT, Nikita KS, Villagra F, Pastor MA. Feasibility study of a wearable system based on a wireless body area network for gait assessment in Parkinson's disease patients. Sensors. 2014;14(3):4618-33.
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); 2018: IEEE.
26. Baguley T. Serious stats: A guide to advanced statistics for the behavioral sciences: Macmillan International Higher Education; 2012.
27. Cumming G. Inference by eye: reading the overlap of independent confidence intervals. Statistics in medicine. 2009;28(2):205-20.
28. Clarke CE, Patel S, Ives N, Rick CE, Woolley R, Wheatley K, 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 Technology Assessment. 2016;20(63):1-96.
29. Denny AP, Behari M. Motor fluctuations in Parkinson's disease. Journal of the neurological sciences. 1999;165(1):18-23.
30. Schrag A, Quinn N. Dyskinesias and motor fluctuations in Parkinson's disease: A community-based study. Brain. 2000;123(11):2297-305.

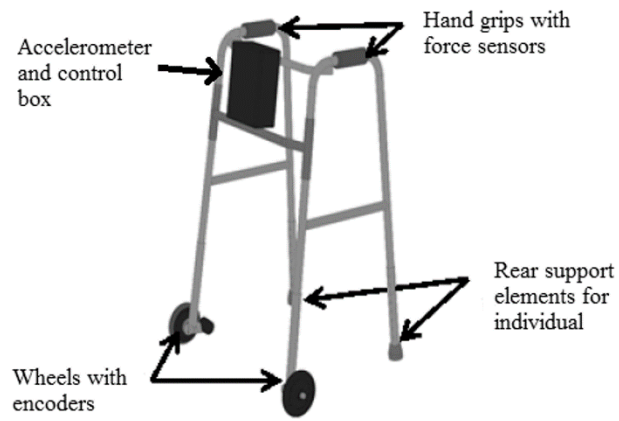


Figure 1: An illustration of the exo-body instrumented walker. The location of the encoders in the wheels, the pressure sensors in the hand grips and the accelerometer in the control box are marked on the figure.

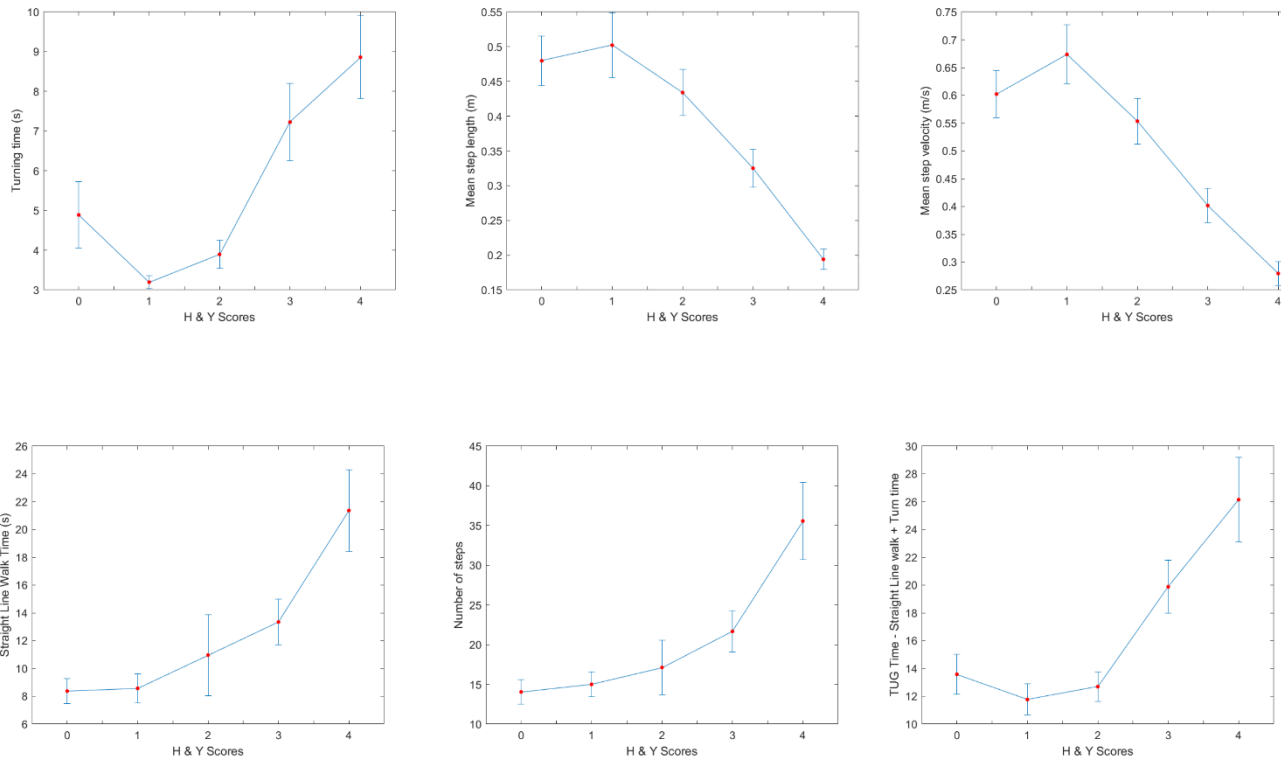


Figure 2: Confidence Interval graphs of six instrumented gait features, for the five different HYgroups.

Table 1: Demographic and clinical characteristics of the five HYGroups and p-values representing the Kruskal-Wallis analyses of the 5 groups.

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

Table 2: Mean and Standard deviation of the features extracted in the instrumented walker 3 meter TUG test. The first five columns provide Mean  $\pm$  Standard deviation values for the five HYgroups and the sixth displays the computed ANOVA p-value of the difference between the groups' means.

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

Abbr  
 eviat  
 ion: s  
 =  
 seco  
 nds;  
 std =  
 stand  
 ard  
 devia  
 tion;  
 m =  
 mete  
 r;  
 m/s  
 =  
 mete  
 r per  
 seco  
 nd;  
 N=  
 Newt  
 ons

Table 3: The overlap of Confidence Interval (CI), in percentage, for the Hoehn and Yahr (HY) and healthy controls (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%) reflect 99% significance of difference between the groups.

<b>Feature Group Pair</b>	<b>Turning Time (s)</b>	<b>Step Velocity (m/s)</b>	<b>Step Length (m)</b>	<b>Straight Line walk time (s)</b>	<b>Step count</b>	<b>TUG time (s)</b>
<b>HC - HY1</b>	0**	19*	73	96	84	46
<b>HC - HY2</b>	21*	28*	36*	83	77	92
<b>HY1 - HY2</b>	0**	0**	8*	75	91	71
<b>HC - HY3</b>	0**	0**	0**	0**	0**	0**
<b>HC - HY4</b>	0**	0**	0**	0**	0**	0**
<b>HY1 - HY3</b>	0**	0**	0**	0**	0**	0**
<b>HY 1- HY4</b>	0**	0**	0**	0**	0**	0**
<b>HY 2- HY3</b>	0**	0**	0**	31*	13*	0**
<b>HY 2- HY4</b>	0**	0**	0**	0**	0**	0**
<b>HY 3- HY4</b>	33*	0**	0**	0**	0**	0**



Abbreviations: TUG – timed up & go; HC – healthy controls; HY – Hoehn and Yahr stage

\* Overlap corresponds to a 95% statistical significance of difference

\*\* Overlap corresponds to a 99% statistical significance of difference

Table 4: Pearson’s correlation r-values and p-values (in parenthesis) between sensor-acquired gait features extracted from the 3m TUG, and clinical data and observations. 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.

	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/sec)	Mean acceleration (m/S <sup>2</sup> )	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	<b>0,349</b> (0,006)	<b>0,378</b> (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	<b>0,274</b> (0,03)	0,196 (0,13)	<b>0,275</b> (0,03)	<b>-0,264</b> (0,04)	<b>0,376</b> (0,003)	<b>-0,3</b> (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	<b>0,275</b> (0,03)	<b>0,28</b> (0,03)	0,187 (0,15)	<b>-0,337</b> (0,008)	<b>0,274</b> (0,03)	<b>-0,354</b> (0,005)	-0,16 (0,22)	0,193 (0,14)	<b>0,34</b> (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)	<b>0,381</b> (0,02)	0,246 (0,15)	-0,101 (0,56)	<b>0,37</b> (0,03)	-0,114 (0,51)	<b>-0,468</b> (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	<b>0,597</b> (<0,001)	<b>0,525</b> (<0,001)	<b>0,495</b> (<0,001)	<b>-0,575</b> (<0,001)	<b>0,504</b> (<0,001)	<b>-0,632</b> (<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)

Abbreviations: TUG = Timed Up and Go; PD = Parkinson’s disease; m-UPDRS = motor Unified Parkinson’s disease Rating Scale.