Reliability of average daily steps measured through a consumer smartwatch in Parkinson's disease phenotypes, stages and severities: a cross-sectional study.

Edoardo Bianchini, Domiziana Rinaldi, Lanfranco De Carolis, Silvia Galli, Marika Alborghetti, Clint Hansen, Antonio Suppa, Marco Salvetti, Francesco Ernesto Pontieri, Nicolas Vuillerme

Submitted to: Journal of Medical Internet Research
on: June 16, 2024

Table of Contents

Reliability of average daily steps measured through a consumer smartwatch in Parkinson's disease phenotypes, stages and severities: a cross-sectional study.

Edoardo Bianchini[1, 2] MD; Domiziana Rinaldi[1] MD; Lanfranco De Carolis[1] MD; Silvia Galli[1] MD; Marika Alborghetti[1] MD; Clint Hansen[2, 3]; Antonio Suppa[4, 5] MD; Marco Salvetti[1, 5] Prof Dr Med; Francesco Ernesto Pontieri[1] Prof Dr Med; Nicolas Vuillerme[2]

[1]Department of Neuroscience, Mental Health and Sensory Organs (NESMOS) Sapienza University of Rome Rome IT
[2]AGEIS, Université Grenoble Alpes Grenoble FR
[3]Department of Neurology Kiel University Kiel DE
[4]Department of Human Neurosciences Sapienza University of Rome Rome IT
[5]IRCCS Neuromed Institute Pozzilli IT

Corresponding Author:
Edoardo Bianchini MD
Department of Neuroscience, Mental Health and Sensory Organs (NESMOS)
Sapienza University of Rome
Via di Grottarossa 1035
Rome
IT

Abstract

Background: Average daily steps (avDS) could be a valuable indicator of real-world ambulation in people with Parkinson's disease (PwPD) and previous studies reported the validity and reliability of this measure. Nonetheless, no study to date has considered disease phenotype, stage and severity when assessing reliability of consumer wrist-worn devices to estimate daily step count in unsupervised, free-living conditions in PwPD.

Objective: To assess and compare the reliability of a consumer wrist-worn smartwatch (Garmin Vivosmart 4) in counting avDS in PwPD in unsupervised, free-living conditions among disease phenotypes, stages, and severity groups.

Methods: One-hundred-four PwPD were monitored through Garmin Vivosmart 4 for 5 consecutive days. Total daily steps for each day were recorded and avDS were calculated. PwPD were dichotomized into tremor dominant (TD) (N=39) or postural instability and gait disorder (PIGD) (N=65), presence (N=57) or absence (N=47) of tremor, and mild (N=65) or moderate (N=39) disease severity. Based on modified Hoeh and Yahr scale (mHY), PwPD were further dichotomized into earlier (mHY 1-2) (N=68) or intermediate (mHY 2.5-3) (N=36) disease stage. Intraclass correlation coefficient (ICC) (3, k), standard error of measurement (SEM) and minimum detectable change (MDC) were used to evaluate the reliability of avDS for each subgroup. The threshold for acceptability was set at an ICC ? 0.8 with a lower bound of 95% confidence interval (CI) ? 0.75. Student's t-tests for independent groups and analysis of 83.4% CI overlap were used to compare ICC between each group pair.

Results: Reliability of avDS measured through Garmin Vivosmart 4 for 5 consecutive days in unsupervised, free-living conditions was acceptable in the overall population with an ICC of 0.89 (0.85-0.92), SEM% below 10% and an MDC of 1580 steps/day (27% of criterion). In all investigated subgroups, reliability of avDS was also acceptable (ICC range 0.84-0.94). However, ICCs were significantly lower in PwPD with tremor (P=.030), with mild severity (P=.040) and earlier stage (P=.003). Moreover, SEM% was below 10% in PwPD with PIGD phenotype, without tremor, moderate disease severity and intermediate disease stage, with a MDC ranging from 1148 to 1687 steps/day (18-25% of criterion). Conversely, in PwPD with TD phenotype, tremor, mild disease severity and earlier disease stage, SEM was >10% of criterion and MDC values ranged from 1401 to 2263 steps/day (30- 33% of the criterion).

Conclusions: In mild-to-moderate PwPD, avDS measured through a consumer smartwatch in unsupervised, free-living conditions for 5 consecutive days are reliable irrespective of disease phenotype, stage, and severity. However, in PwPD with TD phenotype, tremor, mild disease severity and earlier disease stages, reliability could be lower. These findings could facilitate a broader and informed implementation of avDS as an index of ambulatory activity in PwPD.

(JMIR Preprints 16/06/2024:63153)
DOI: https://doi.org/10.2196/preprints.63153

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Edoardo Bianchini[1,2]*, Domiziana Rinaldi[1], Lanfranco De Carolis[1], Silvia Galli[1], Marika Alborghetti[1], Clint Hansen[2,3], Antonio Suppa[4,5], Marco Salvetti[1,5], Francesco E. Pontieri[1], Nicolas Vuillerme[2,6]*

[1] *Department of Neuroscience, Mental Health and Sensory Organs (NESMOS), Sapienza University of Rome, Rome, Italy*
[2] *AGEIS, Université Grenoble Alpes, Grenoble, France*
[3] *Department of Neurology, Kiel University, Kiel, Germany*
[4] *Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy*
[5] *IRCCS Neuromed Institute, Pozzilli, Italy*
[6] *Institut Universitaire de France, Paris, France*

Author contribution

EB, NV and FEP designed the study. EB, DR, LDC, SG and MA recruited participants and collected data. EB performed data analyses. EB and NV wrote the first draft of the manuscript. DR, LDC, SG, MA, CH, AS, MS and FEP reviewed manuscript draft. All authors contributed to the article and approved the submitted version.

Journal: Journal of Medical Internet Research (IF 7.4; https://www.jmir.org/)
Word count (main text): 3492
Word count (abstract): 440 (max 450)
Figures: 1
Tables: 3
References: 62

***Correspondence:**

Prof. Nicolas Vuillerme
AGEIS, Université Grenoble Alpes
Faculty of Medicine - Bâtiment Jean Roget
38706 La Tronche cédex, France
Phone: +33 4 7663 7104
E-mail: nicolas.vuillerme@univ-grenoble-alpes.fr

Dr. Edoardo Bianchini
Department of Neuroscience, Mental Health and Sensory Organs (NESMOS)
Sapienza University of Rome
Via di Grottarossa 1035, 00189 Roma RM, Italy
Phone +39 06 3377 5579
Fax +39 06 3377 5900
E-mail: edoardo.bianchini@uniroma1.it

Reliability of average daily steps measured through a consumer smartwatch in Parkinson's disease phenotypes, stages and severities: a cross-sectional study.

Abstract

Background: Average daily steps (avDS) could be a valuable indicator of real-world ambulation in people with Parkinson's disease (PwPD) and previous studies reported the validity and reliability of this measure. Nonetheless, no study to date has considered disease phenotype, stage and severity when assessing reliability of consumer wrist-worn devices to estimate daily step count in unsupervised, free-living conditions in PwPD.

Objective: To assess and compare the reliability of a consumer wrist-worn smartwatch (Garmin Vivosmart 4) in counting avDS in PwPD in unsupervised, free-living conditions among disease phenotypes, stages, and severity groups.

Methods: One-hundred-four PwPD were monitored through Garmin Vivosmart 4 for 5 consecutive days. Total daily steps for each day were recorded and avDS were calculated. PwPD were dichotomized into tremor dominant (TD) (N=39) or postural instability and gait disorder (PIGD) (N=65), presence (N=57) or absence (N=47) of tremor, and mild (N=65) or moderate (N=39) disease severity. Based on modified Hoehn and Yahr scale (mHY), PwPD were further dichotomized into earlier (mHY 1-2) (N=68) or intermediate (mHY 2.5-3) (N=36) disease stage. Intraclass correlation coefficient (ICC) (3, k), standard error of measurement (SEM) and minimum detectable change (MDC) were used to evaluate the reliability of avDS for each subgroup. The threshold for acceptability was set at an ICC ≥ 0.8 with a lower bound of 95% confidence interval (CI) ≥ 0.75. Student's t-tests for independent groups and analysis of 83.4% CI overlap were used to compare ICC between each group pair.

Results: Reliability of avDS measured through Garmin Vivosmart 4 for 5 consecutive days in unsupervised, free-living conditions was acceptable in the overall population with an ICC of 0.89 (0.85-0.92), SEM% below 10% and an MDC of 1580 steps/day (27% of criterion). In all investigated subgroups, reliability of avDS was also acceptable (ICC range 0.84-0.94). However, ICCs were significantly lower in PwPD with tremor (*P*=.030), with mild severity (*P*=.040) and earlier stage (*P*=.003). Moreover, SEM% was below 10% in PwPD with PIGD phenotype, without tremor, moderate disease severity and intermediate disease stage, with a MDC ranging from 1148 to 1687 steps/day (18-25% of criterion). Conversely, in PwPD with TD phenotype, tremor, mild disease severity and earlier disease stage, SEM was >10% of criterion and MDC values ranged from 1401 to 2263 steps/day (30- 33% of the criterion).

Conclusions: In mild-to-moderate PwPD, avDS measured through a consumer smartwatch in unsupervised, free-living conditions for 5 consecutive days are reliable irrespective of disease phenotype, stage, and severity. However, in PwPD with TD phenotype, tremor, mild disease severity and earlier disease stages, reliability could be lower. These findings could facilitate a broader and informed implementation of avDS as an index of ambulatory activity in PwPD.

Keywords: Gait; Parkinson’s disease; phenotype; wearable sensors; smartwatch; step count; reliability; activity monitor; digital health technology; digital outcome measures.

Introduction

Walking is a fundamental motor ability, and it is pivotal for functional independence and social well-being [1]. Gait disturbances are common in people with Parkinson's disease (PwPD), and include shuffling gait, shortened step length, altered automaticity, decreased arm swing and as freezing of gait [2]. These represent a particularly disabling group of symptoms, significantly hampering quality of life of PwPD [3] and increasing the risk of falls [4].

Daily steps are an easy-to-collect and useful measure of ambulatory activity and mobility [5]. Although this parameter could not provide details regarding subtle gait features, several evidence linked a reduced daily step count to overall mortality risk [6–9], as well as to range of health conditions, such as dementia [10], oncological and cardiovascular diseases [11,12]. Previous studies have also reported a negative correlation between daily steps and disease severity in PwPD [13] and proposed a minimum daily step goal of 4200 to match with physical activity recommendations in the early stages of the disease [14].

Wearable devices, including smartwatches, could represent a useful option to estimate daily steps in an unobtrusive, ecological way [15]. Moreover, wearables are widely available on the market, easily used by the general population and enable unobtrusive continuous long-term data collection [16]. However, since these devices are usually tested in healthy populations, knowledge on validity and reliability of collected data is generally limited when applied to different groups of patients. Therefore, a growing body of literature focused on validity and reliability of consumer wearable devices for step counting, with generally positive results [17–19].

In PD, motor and gait manifestations could render step detection and step count challenging and hence significantly diminish the validity and reliability of device algorithms [20–22]. Nevertheless, a prior study from our group involving 47 PwPD demonstrated a strong criterion validity in step counting using a consumer smartwatch (Garmin Vivosmart 4), when worn on the side least affected by the disease and under well-controlled pharmacological conditions [20]. Similar results were obtained in real-life setting by Ginis and colleagues [23] who demonstrated the criterion validity in estimating average daily steps (avDS) of two Fitbit devices (Fitbit Alta and Fitbit Inspire 3) in 28 PwPD at home, compared to a research-grade device (Dynaport Movemonitor, McRoberts, NL).

Besides criterion validity, we recently demonstrated that an optimal reliability of avDS recorded by Garmin Vivosmart 4 in real-life conditions could be achieved if the smartwatch is worn for a minimum of 4 days [24].

However, the clinical heterogeneity of PD when evaluating metrological characteristics of step-counting devices has been neglected so far. Indeed, the clinical presentation of PD is highly variable among individuals, thus, significant efforts have been made to identify distinct clusters and subtypes [25]. Several classifications were proposed over the years from clinical-based [26–28] to more recent biomarker-based classifications [29–31]. In this regard, one of the most used classification distinguishes PwPD with predominant features of tremor (ie. tremor dominant, TD) or gait, posture and balance issues (ie. postural instability and gait disorder PIGD) [26–28] based on sub-items scores of the Unified Parkinson Disease Rating Scale (UPDRS) and its revision (MDS-UPDRS) [32]. This classification could be relevant when measuring avDS in PwPD since tremor could increase noise-to-signal ratio, making step detection more challenging [21]. Similarly, in PIGD individuals, the higher degree of gait alterations could alter the performance of step detection algorithms [20,22].

Symptoms severity and disease stage could also represent other relevant parameters to be considered when assessing reliability of any wearable devices in counting avDS. Indeed, with disease progression, gait features increasingly deviate from normality [2,33], and tremor and bradykinesia could further alter the spatiotemporal and kinematic characteristics of walking, dampening, in turn,

algorithm performance in step detection [20–22].
Nevertheless, no study to date considered disease phenotype, stage and symptoms severity when assessing reliability of consumer wrist-worn devices for step counting in unsupervised, free-living conditions in mild-to-moderate PwPD. The present study was hence specifically designed to address this issue. We hypothesized that a reduced reliability might be observed in TD individuals and in PwPD with more severe symptoms and in more advanced disease stages due to the aforementioned increased signal noise due to tremor, the higher degree of motor symptoms and the more marked gait alterations.

Methods

Population

Participants were recruited at the Movement Disorder Outpatient Service of the Sant'Andrea University Hospital (Rome, Italy) in the period between March 2023 and March 2024. Inclusion criteria were: (i) diagnosis of idiopathic PD according to the MDS criteria (Postuma et al., 2015); (ii) age ≥ 18 years; (iii) disease stage <4 according to the modified Hoeh and Yahr scale (mHY)[34]; (iv) classification as TD or PIGD according to Stebbins et al. [27]. Exclusion criteria were: (i) cognitive impairment, defined by Montreal Cognitive Assessment (MoCA) [35] score < 21; (ii) orthopedic, rheumatologic, or systemic conditions affecting mobility as judged by the assessor.

Population

This cross-sectional study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments. Approval was granted by the local Ethical Committee of Sapienza, University of Rome (Ref. 0372/2022). Data collection and processing followed the current European regulation for data protection. All participants provided written informed consent before the beginning of measurements.

Demographic and clinical data

Participants were evaluated during scheduled visits. Demographics and anthropometric measures (including age, sex, weight, and height) were collected. Disease duration, disease stage according to mHY and Levodopa Equivalent Daily Dose (LEDD) [36] were also collected. MDS-UPDRS [32] part III was used to assess motor symptoms severity.
Patients were divided in 4 subgroup pairs. Based on MDS-UPDRS part II and III scores, participants were classified into TD or PIGD disease subtypes [27]. To evaluate the effect of tremor presence on device reliability, patients were also classified in those with and without tremor based on item 2.10 of MDS-UPDRS part II. Concerning disease severity, participants were grouped in those with mild or moderate disease severity based on MDS-UPDRS score as proposed by Martínez-Martín and colleagues [37]. Similarly, patients were dichotomized in earlier (mHY 1-2) or intermediate (mHY 2.5-3) stage, based on mHY score.

Experimental procedure

Participants received the smartwatch Garmin Vivosmart 4 to wear at home for 5 consecutive days, including at least one weekend day, on the wrist of the body side least affected by the disease [20]. We chose a 5-day period since we demonstrated previously that a minimum of 4 days of monitoring is needed to reliably estimate daily step count in PwPD [24]. Each smartwatch was configured according to the producer's recommendations and participants were asked to perform daily activities as usual. After 5 days, participants returned the smartwatch. Total daily number of steps for each day was recorded and avDS were calculated [24].

Data and statistical analysis

The statistical analyses were performed using JASP v0.18.3.0 (JASP Team, University of Amsterdam, The Nederlands), R v4.0.3 and RStudio v2023.12.0+369 for Windows (R Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics were calculated for the examined variables. Normality of distributions were assessed by histogram and residual plots inspection.

To evaluate the relative reliability for the 5-day monitoring period in the overall population and in each subgroup, a two-way intraclass correlation coefficient (ICC) with a fixed set of raters and averaged ratings was used [ICC (3,k), where k was the number of days of measurement], together with a custom R script. The following reference cut-off values for ICC interpretation were used [38]: Excellent: > 0.90; Good: 0.75 – 0.90; Moderate: 0.50 – 0.75; Poor: < 0.50. The a priori threshold for acceptable ICC was set at a point estimate ≥ 0.80 with a lower bound of 95% Confidence Interval (CI) ≥ 0.75 in accordance with a previous study from our group [24].

To compare ICCs between the four subgroups pairs, two methods were applied. First, standard errors and point estimates of ICCs were used to compute a t-statistics and perform independent groups Student's t-tests. Secondly, CI overlap between each group pair was graphically and numerically assessed. Non-overlapping CIs were considered indicators of significantly different ICCs [39]. Previous evidence underscored that 95% CI overlap assessment could inflate the risk of type II error and suggested that a 83.4% CI could be a more powerful option [39–41]. Therefore, we adopted this method for CI overlap evaluation.

Standard error of measurement (SEM) and minimal detectable change with a confidence interval of 95% (MDC) were used to compute the absolute reliability for the 5-day recordings in the overall population and in each subgroup [42]. SEM and MDC were reported as absolute value and percentage of criterion measure (SEM% and MDC%, respectively). The criterion was the average daily step count derived from the 5 days. For all analyses, the significance threshold was set at $\alpha < 0.05$. All data were reported as mean ± SD or median (Q1-Q3) for numerical data and N (%) for categorical variables.

Results

A total of 104 PwPD patients were included in the study. All patients were monitored through Garmin Vivosmart 4 at home for a period of 5 consecutive days. Participants took on average 5923 ± 3014 daily steps, ranging from 357 to 12620. Details of demographic, anthropometrics and clinical variables are shown in Table 1.

Table 1. Demographic, anthropometric and clinical characteristics of the study population. BMI: body mass index; F: females; LEDD: Levodopa equivalent daily dose; MDS-UPDRS-III: Movement Disorder Society Unified Parkinson Disease Rating Scale part III; mHY: modified Hoehn and Yahr scale; PIGD: postural instability and gait disorder; TD: tremor dominant.

	Overall (N=104)	**TD (N=39)**	**PIGD (N=65)**	**Tremor (N=57)**	**No tremor (N=47)**	**Mild (N=65)**	**Moderate (N=39)**	**mHY 1-2 (N=68)**	**mHY 2.5-3 (N=36)**
Age (years)	68.0 ± 8.4	66.4 ± 9.0	68.9 ± 8.0	69.5 ± 7.9	66.7 ± 8.8	66.8 ± 8.7	69.9 ± 7.7	65.3 ± 7.8	73.0 ± 7.3
Height (cm)	171 ± 9.0	173 ± 7.6	170 ± 9.6	170 ± 9.6	172 ± 8.4	173 ± 9.3	168 ± 7.6	173 ± 8.1	168 ± 9.8
Weight (kg)	75.7 ± 13.1	76.7 ± 13.0	75.1 ± 13.2	75.0 ± 12.1	76.3 ± 13.9	77.2 ± 13.7	73.2 ± 11.8	77.7 ± 13.8	71.9 ± 10.8
BMI (kg/m^2)	25.7 ± 3.4	25.5 ±	25.8 ± 3.4	25.8 ± 3.0	25.7 ± 3.8	25.7 ± 3.5	25.8 ± 3.4	25.9 ±	25.4 ±

		3.6						3.7	3.0
Sex (F)	34 (33%)	11 (28%)	23 (35%)	19 (33%)	15 (32%)	22 (34%)	12 (31%)	22 (32%)	12 (33%)
Disease duration (years)	6.4 ± 4.4	5.2 ± 4.4	7.0 ± 4.4	7.2 ± 3.9	5.7 ± 4.8	5.7 ± 4.5	7.5 ± 4.2	5.3 ± 4.3	8.4 ± 4.1
LEDD (mg)	553 ± 302	418 ± 247	634 ± 304	623 ± 271	495 ± 316	489 ± 289	659 ± 296	453 ± 248	741 ± 307
mHY	2 (2 - 2.5)	2 (1 - 2)	2 (2 - 2.5)	2 (2 - 2.5)	2 (2 - 2)	2 (2 - 2)	2.5 (2 - 3)	2 (2 - 2)	2.5 (2.5 - 3)
MDS-UPDRS-III	27 (21 - 32)	26 (18 - 31)	29 (22 - 33)	29 (22 - 33)	26 (21 - 32)	23 (19 - 29)	33 (29 - 37)	23 (18 - 29)	33 (29 - 36)
AvDS	5923 ± 3014	6654 ± 2733	5485 ± 3109	4594 ± 2612	7020 ± 2898	6512 ± 2857	4942 ± 3049	6838 ± 2908	4195 ± 2419

Reliability of avDS in PwPD

Relative reliability

AvDS collected during 5 consecutive days showed level of relative reliability above the threshold of acceptability, as indicated by an ICC point estimate ≥ 0.80 and a lower CI 95% limit ≥ 0.75, in the overall population and in all subgroups. Moreover, daily step count showed an excellent reliability in PwPD in intermediate disease stage. Details of ICC and CI limits are shown in Table 2.

Table 2. ICC (3,k) values with 95% CI for the overall population and each subgroup. For subgroups 83.4% CI to assess intervals overlap are also reported. CI: confidence interval; ICC: intraclass correlation coefficient; mHY: modified Hoehn and Yahr scale; PIGD: postural instability and gait disorder; TD: tremor dominant.

	Overall (N=104)	**TD (N=39)**	**PIGD (N=65)**	**Tremor (N=57)**	**No tremor (N=47)**	**Mild (N=65)**	**Moderate (N=39)**	**mHY 1-2 (N=68)**	**mHY 2.5-3 (N=36)**
ICC (3, k)	0.888	0.854	0.899	0.838	0.914	0.856	0.919	0.839	0.939
Lower 95%CI	0.850	0.767	0.855	0.760	0.868	0.793	0.871	0.769	0.900
Upper 95%CI	0.919	0.916	0.933	0.896	0.947	0.905	0.953	0.892	0.966
Lower 83.4%CI	-	0.797	0.869	0.786	0.883	0.814	0.887	0.792	0.914
Upper 83.4%CI	-	0.901	0.925	0.881	0.939	0.893	0.945	0.878	0.959

Absolute reliability

AvDS showed a SEM below 10% in the overall population with MDC of 1580 (26.7% of the criterion). AvDS also showed a SEM below 10% in PIGD disease subtype, in participants without tremor and in PwPD with a moderate disease severity and in an intermediate disease stage with a MDC ranging from 1148 to 1687 steps/day (18% to 25% of criterion) or . Conversely, in TD disease subtype, in participants with tremor and in PwPD with a mild disease severity and in an early disease stage, SEM was >10% of criterion and MDC values ranged from 1401 to 2263 steps/day (30% to 33% of the criterion). Details of SEM and MDC are shown in Table 3.

Table 3. Absolute and percentage values of SEM and MDC for the overall population and each subgroup; MDC: minimal detectable change; mHY: modified Hoehn and Yahr scale; PIGD: postural instability and gait disorder; SEM: Standard error of measurement; TD: tremor dominant.

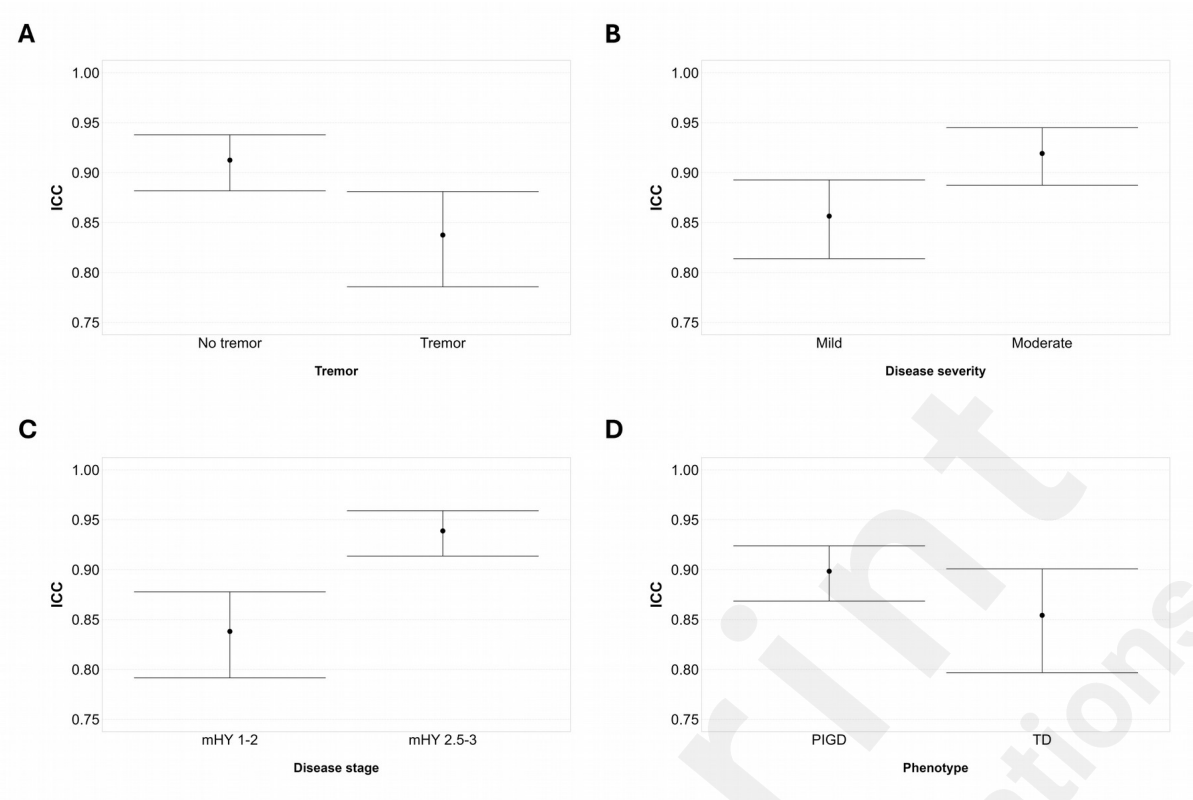
	Overall (N=104)	TD (N=39)	PIGD (N=65)	Tremor (N=57)	No tremor (N=47)	Mild (N=65)	Moderate (N=39)	mHY 1-2 (N=68)	mHY 2.5-3 (N=36)
SEM	570	742	495	506	609	701	414	817	267
SEM%	9.6	11.1	9.0	11.0	8.7	10.8	8.4	11.9	6.4
MDC	1580	2056	1372	1401	1687	1944	1148	2263	741
MDC%	26.7	30.9	25.0	30.5	24.0	29.8	23.2	33.1	17.7

Reliability comparison between subgroups

When comparing ICCs between subgroups pairs, Student's t-test for independent groups showed a significant difference between participants with and without tremor (t(102)=1.897; *P*=.030), between PwPD with mild and moderate disease severity (t(102)=1.765; *P*=.040) and between PwPD in early and intermediate disease stage (t(102)=2.817; *P*=.003). Conversely, no significant difference was found between TD and PIGD participants (t(102)=1.048; *P*=.149).

The analysis of CI 83.4% showed no overlap between interval limits between participants with and without tremor (Figure 1A), between PwPD in early and intermediate disease stage (Figure 1C) and only a negligible overlap between PwPD with mild and moderate disease severity (Figure 1B). Conversely, a degree of overlap of the two CI 83.4% was observed between TD and PIGD participants (Figure 1D). Details of CI 83.4% limits are shown in table 2.

Figure 1. ICC and 83.4% CI comparison for each subgroup pairs. Panel A: PwPD with reported presence or absence of tremor. Panel B: PwPD with mild and moderate disease severity as indicated by MDS-UPDRS. Panel C: PwPD with earlier (mHY 1-2) and intermediate (mHY 2.5-3) disease stage. Panel D: PIGD and TD phenotypes. ICC (3,k) point estimate is indicated by the black dot, CI limits are represented by the vertical bars. CI: confidence interval; ICC: intraclass correlation coefficient; MDS-UPDRS: International Movement Disorders Society Unified Parkinson Disease Rating Scale; mHY: modified Hoehn & Yahr stage; PIGD: postural instability and gait disorder; TD: tremor dominant.



Discussion

This cross-sectional study aimed at assessing and comparing the reliability of a consumer wrist-worn smartwatch (Garmin Vivosmart 4) in counting avDS in PwPD in unsupervised, free-living conditions for 5 consecutive days between disease phenotypes, stages and severity groups.

Overall, our results showed, for the first time, that avDS were acceptably reliable in PwPD irrespective of disease severity, stage, or phenotype. Nevertheless, our results further indicated a lower reliability in PwPD with TD phenotype, tremor, lower disease severity and earlier stage.

Reliability of avDS in the overall PwPD population

We found that relative reliability of avDS, measured in unsupervised, free-living conditions, by Garmin Vivosmart 4 for 5 consecutive days was within the a priori criteria for acceptability in the overall PwPD population. We identified only two studies investigating reliability of wearable devices in measuring avDS in PwPD [24,43]. Paul and colleagues [43] reported that 2 consecutive days of monitoring were sufficient to obtain an ICC > 0.9, using a research-grade, ankle-mounted, step counter (Step Activity Monitor, SAM) in 92 PwPD. One study from our group investigated the reliability of a wrist-worn consumer device in measuring avDS in PwPD before. In that study, we found an ICC (3,k) of 0.88 (0.82-0.93) for 4 days of monitoring in 56 mild-to-moderate PwPD using Garmin Vivosmart 4 [24]. The present results are consistent with our previous study, but also with other studies investigating the reliability of avDS through wrist-worn wearables in healthy elderly [43–46], and in people with various neurological conditions, such as multiple sclerosis [47] and stroke [48]. Regarding absolute reliability, we found that SEM% was below 10% (9.6%) with an MDC of 1580 steps/day (27% of criterion). Only one study, from our group, investigated SEM and MDC for avDS in PwPD [24]. The present results are consistent with our previous study in which

we reported a SEM% of 9.2% and a MDC of 1495 steps/day (or 26% of criterion) [24]. Taken together, these results confirm the reliability of avDS measured through Garmin Vivosmart 4 in unsupervised, free-living conditions in mild-to-moderate PwPD for 5 consecutive days.

Reliability of avDS in PwPD subgroups

To our knowledge, this is the first study to investigate and compare the reliability of a consumer smartwatch in measuring avDS in different subgroups of PwPD based on disease phenotype, severity, and stage. We found that in all investigated subgroups (i.e., TD vs PIGD, mild vs moderate, tremor vs non-tremor, and earlier vs intermediate), ICC values were within the a priori criteria for acceptability (ICC range 0.84-0.94). However, a significantly lower ICC was observed in PwPD with tremor, mild disease severity and earlier disease stage. Moreover, SEM% was below 10% in PwPD with PIGD phenotype, moderate disease severity, intermediate disease stage and without tremor, with a MDC ranging from 1148 to 1687 steps/day (18-25% of criterion). Conversely, in PwPD with TD phenotype, tremor, mild disease severity and earlier disease stage, SEM was >10% of criterion and MDC values ranged from 1401 to 2263 steps/day (30- 33% of the criterion).

MDC, defined as the minimal change that falls outside the measurement error of an instrument, is extremely relevant in study design since it allows to calculate the sample size of studies aiming to assess the effectiveness of interventions [49]. MDC could be also crucial to define the appropriateness and feasibility of using a determinate device to measure a given construct. A prior work from Handlery and colleagues [50] reported an increase in 1250 steps/day following a high-intensity physical activity intervention in PwPD measured through a research-grade wrist-worn device (Actigraph GT3X). In the present study, we found that in PwPD with TD phenotype, mild disease severity and earlier disease stage MDC was ~2000 steps/day. Although a direct comparison with the metrics reported in the work from Handlery and collaborators [50] could not be performed due the different devices, we could hypothesize that only large modification in avDS could be reliably measured through Garmin Vivosmart 4 in the aforementioned PwPD subgroups. To this end, future studies will be needed to define the minimal clinically important difference for avDS measured through consumer wrist-worn devices and to assess the attainability of avDS modifications sufficiently large to be reliably detected by these devices.

The reduced reliability in PwPD with tremor is in line with our hypothesis that tremor could reduce the performance of step detection algorithm. In fact, tremor could increase the noise-to-signal ratio in accelerometer signal and, in turn, render step detection more challenging [21]. Indeed, a previous study highlighted that tremor and dyskinesia together contributed to more than 19% of the variation in daily step counts when comparing measurements from waist-worn and wrist-worn devices in 46 PwPD with similar characteristics to those included in the present study [21]. In this regard, our study further supports the assumption that tremor could reduce step detection performance of wrist-worn devices in PwPD.

On the other hand, our hypothesis that a reduced reliability might be observed in PwPD with higher symptoms severity and more advanced stages was not supported by our results. Indeed, the reduced reliability of avDS observed in PwPD with mild disease severity and earlier disease stage, is somehow counterintuitive. In fact, previous evidence highlighted that step count was less accurate in people walking at slower gait speed and with shorter step length in several neurological and musculoskeletal conditions, including PD [20,22,51–54]. Reduced step length and slower walking speed are typical feature of parkinsonian gait, with a higher prevalence along the disease course [2,33]. Moreover, another typical characteristic of walking in PwPD is the reduced automaticity that leads to a more discontinuous and irregular gait pattern that can furtherly reduce device accuracy in step detection [2,22].

Despite these considerations, our results showed that avDS estimation was more reliable in PwPD

with moderate disease severity and intermediate disease stage, compared with individuals with a mild disease severity and earlier disease stage. In this regard, it must be considered that reliability is a measure of consistency and reproducibility of measurement and not a measure of accuracy [55]. Therefore a reduced accuracy could not directly translate into a reduced reliability. We could hypothesize that in more advanced stages of PD, variability of clinical presentation could be lower. In earlier stage, indeed, symptoms heterogeneity, both in terms of motor and non-motor features, could be extremely high [56–58]. However, this variability could decrease with disease progression since motor symptoms tends to consistently worsen along disease course and motor features such as gait and balance impairment become increasingly prevalent [59,60]. In addition, phenotype is dynamic along disease course and some researchers have proposed that the classification into PIGD/TD evolves over time [61,62]. One study, indeed, reported that over a period of 8 years, approximately 70% of TD individuals transitioned to PIGD, whereas only 4% of PIGD individuals transitioned to TD [61]. Another study reported that 45% of TD patients at baseline had a subtype shift along a 2-year follow-up while 85% of PIGD participants remained as PIGD [62]. This is mirrored in our study cohort, where 32/65 (49%) PwPD with mild disease severity were classified as TD, whereas only 7/39 (18%) in the moderate group. Similarly, 33/68 (49%) PwPD with earlier disease stage were classified as TD, whereas only 6/36 (17%) in the intermediate group. Therefore, a regression toward a more uniform motor impairment along disease course might be considered. Moreover, since we found that tremor could be a relevant factor in reducing avDS reliability, the different prevalence of TD phenotype could also contribute to explaining our results. However, it must be underlined that no study, to our knowledge, systematically compared the heterogeneity of PD features across early, intermediate and advanced disease stages. Therefore our hypothesis should be taken with caution and future studies are needed to confirm it.

In conclusion, our findings highlight that, although avDS were reliable across the examined subgroups, clinicians and researchers should consider disease phenotype, stage and severity when implementing wrist-worn wearables and interpreting mobility data collected through these devices.

Limitations

We acknowledge that the present study has some limitations. First, the PwPD included in our study displayed relatively preserved cognitive functions, due to our exclusion of participants with a MoCA<21. Additionally, those with more advanced disease stages or requiring walking aids were not included. This potentially limits the generalizability of our findings. However, it should be considered that the sample in our study can be seen as representative of the typical target for interventions using consumer-grade wearable technology. Moreover, including PwPD with a disease stage > 3, using walking aids, or with more severe cognitive impairments, poses significant challenges in the utilization of consumer technology and was beyond our scope. Nevertheless, future research incorporating PwPD with lower functional scores, higher disease severity, and more impactful cognitive impairments would be valuable. Finally, we used only PD subtyping based on clinical features, yet other classification methods and clustering techniques have been proposed incorporating instrumental and biological data. Future studies are warranted to investigate the reliability of consumer smartwatches in PwPD subgroups defined using multimodal biomarkers.

Conclusions

In mild-to-moderate PD patients, avDS measured through a consumer smartwatch in unsupervised, free-living conditions for 5 consecutive days are reliable irrespective of disease phenotype, stage, and severity. Researchers and clinicians who want to implement these instruments should consider that in PwPD with TD phenotype, tremor, mild disease severity and earlier disease stage reliability

could be lower and MDC could be higher. Future studies is needed to define the minimal clinically important difference for avDS measured through consumer wrist-worn devices and to assess the attainability of avDS modifications sufficiently large to be reliably detected by wrist-worn consumer devices. Taken together, we believe that these results could facilitate a broader implementation and an informed application of avDS as an index of ambulatory activity in PwPD and could be highly relevant to develop monitoring, preventive, educational and rehabilitation strategies for PD.

Acknowledgements

The authors sincerely thank all people that were involved in the data collection, from study participants to students who aided and assisted the assessments.

Data Availability

The data presented in this study are available on reasonable request from the corresponding author.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This research was performed as part of Edoardo Bianchini's PhD (Sapienza University of Rome, Italy & University Grenoble Alpes, France) and funded through Doctoral Bursary from Sapienza University of Rome (Student Number 1472455). The research was supported by Sapienza University of Rome as part of the program Sapienza Research Call 2020, grant n. AR120172B97AF626 and Research Call 2022, grant n. AR222181683A15AA and by the French National Research Agency (ANR) in the framework of the Investissements d'avenir program (ANR-10-AIRT-05 and ANR-15-IDEX-02), and the MIAI @ Grenoble Alpes (ANR-19-P3IA-0003). This work also forms part of a broader translational and interdisciplinary GaitAlps research program (N.V.)

Abbreviations

avDS: average daily steps
BMI: body mass index
CI: confidence interval
ICC: intraclass correlation coefficient
LEDD: Levodopa Equivalent Daily Dose
MDC: minimal detectable change
MDS-UPDRS: International Movement Disorders Society unified Parkins's Disease rating Scale
mHY: modified Hoeh and Yahr scale
MoCA: Montreal Cognitive Assessment
PD: Parkinson' disease
PIGD: postural instability and gait disorder
PwPD: people with Parkinson's disease
SEM: Standard error of measurement
TD: tremor dominant

References

1. MacKinnon CD. Sensorimotor anatomy of gait, balance, and falls. Handb Clin Neurol 2018;159:3–26. PMID:30482322

2. Mirelman A, Bonato P, Camicioli R, Ellis TD, Giladi N, Hamilton JL, Hass CJ, Hausdorff JM, Pelosin E, Almeida QJ. Gait impairments in Parkinson's disease. Lancet Neurol 2019 Jul;18(7):697–708. PMID:30975519

3. Lee M, Noh Y, Youm C, Kim S, Park H, Noh B, Kim B, Choi H, Yoon H. Estimating Health-Related Quality of Life Based on Demographic Characteristics, Questionnaires, Gait Ability, and Physical Fitness in Korean Elderly Adults. Int J Environ Res Public Health 2021 Nov 11;18(22):11816. PMID:34831575

4. Creaby MW, Cole MH. Gait characteristics and falls in Parkinson's disease: A systematic review and meta-analysis. Parkinsonism Relat Disord 2018 Dec;57:1–8. PMID:30041848

5. Bassett DR, Toth LP, LaMunion SR, Crouter SE. Step Counting: A Review of Measurement Considerations and Health-Related Applications. Sports Med 2017 Jul;47(7):1303–1315. PMID:28005190

6. Inoue K, Tsugawa Y, Mayeda ER, Ritz B. Association of Daily Step Patterns With Mortality in US Adults. JAMA Netw Open 2023 Mar 1;6(3):e235174. PMID:36976556

7. Lee I-M, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of Step Volume and Intensity With All-Cause Mortality in Older Women. JAMA Intern Med 2019 Aug 1;179(8):1105–1112. PMID:31141585

8. Paluch AE, Bajpai S, Bassett DR, Carnethon MR, Ekelund U, Evenson KR, Galuska DA, Jefferis BJ, Kraus WE, Lee I-M, Matthews CE, Omura JD, Patel AV, Pieper CF, Rees-Punia E, Dallmeier D, Klenk J, Whincup PH, Dooley EE, Pettee Gabriel K, Palta P, Pompeii LA, Chernofsky A, Larson MG, Vasan RS, Spartano N, Ballin M, Nordström P, Nordström A, Anderssen SA, Hansen BH, Cochrane JA, Dwyer T, Wang J, Ferrucci L, Liu F, Schrack J, Urbanek J, Saint-Maurice PF, Yamamoto N, Yoshitake Y, Newton RL, Yang S, Shiroma EJ, Fulton JE, Steps for Health Collaborative. Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. Lancet Public Health 2022 Mar;7(3):e219–e228. PMID:35247352

9. Saint-Maurice PF, Troiano RP, Bassett DR, Graubard BI, Carlson SA, Shiroma EJ, Fulton JE, Matthews CE. Association of Daily Step Count and Step Intensity With Mortality Among US Adults. JAMA 2020 Mar 24;323(12):1151–1160. PMID:32207799

10. Del Pozo Cruz B, Ahmadi M, Naismith SL, Stamatakis E. Association of Daily Step Count and Intensity With Incident Dementia in 78 430 Adults Living in the UK. JAMA Neurol 2022 Oct 1;79(10):1059–1063. PMID:36066874

11. Del Pozo Cruz B, Ahmadi MN, Lee I-M, Stamatakis E. Prospective Associations of Daily Step Counts and Intensity With Cancer and Cardiovascular Disease Incidence and Mortality and All-Cause Mortality. JAMA Intern Med 2022 Nov 1;182(11):1139–1148. PMID:36094529

12. Banach M, Lewek J, Surma S, Penson PE, Sahebkar A, Martin SS, Bajraktari G, Henein MY, Reiner Ž, Bielecka-Dąbrowa A, Bytyçi I. The association between daily step count and all-cause and cardiovascular mortality: a meta-analysis. Eur J Prev Cardiol 2023 Dec 21;30(18):1975–1985. PMID:37555441

13. Skidmore FM, Mackman CA, Pav B, Shulman LM, Garvan C, Macko RF, Heilman KM. Daily ambulatory activity levels in idiopathic Parkinson disease. J Rehabil Res Dev 2008;45(9):1343–1348. PMID:19319758

14. Handlery R, Stewart JC, Pellegrini C, Monroe C, Hainline G, Flach A, Handlery K, Fritz S. Physical Activity in De Novo Parkinson Disease: Daily Step Recommendation and Effects of Treadmill Exercise on Physical Activity. Phys Ther 2021 Oct 1;101(10):pzab174. PMID:34244805

15. Mattison G, Canfell O, Forrester D, Dobbins C, Smith D, Töyräs J, Sullivan C. The Influence of Wearables on Health Care Outcomes in Chronic Disease: Systematic Review. J Med Internet Res 2022 Jul 1;24(7):e36690. PMID:35776492

16. Bianchini E, Maetzler W. Wearable systems in movement disorders. International Review of Movement Disorders Academic Press; 2023. doi: 10.1016/bs.irmvd.2023.04.002

17. Chevance G, Golaszewski NM, Tipton E, Hekler EB, Buman M, Welk GJ, Patrick K, Godino JG. Accuracy and Precision of Energy Expenditure, Heart Rate, and Steps Measured by Combined-Sensing Fitbits Against Reference Measures: Systematic Review and Meta-analysis. JMIR Mhealth Uhealth 2022 Apr 13;10(4):e35626. PMID:35416777

18. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, Buote R, Van Heerden D, Luan H, Cullen K, Slade L, Taylor NGA. Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. PMID:32897239

19. Nakagata T, Murakami H, Kawakami R, Tripette J, Nakae S, Yamada Y, Ishikawa-Takata K, Tanaka S, Miyachi M. Step-count outcomes of 13 different activity trackers: Results from laboratory and free-living experiments. Gait Posture 2022 Oct;98:24–33. PMID:36030707

20. Bianchini E, Caliò B, Alborghetti M, Rinaldi D, Hansen C, Vuillerme N, Maetzler W, Pontieri FE. Step-Counting Accuracy of a Commercial Smartwatch in Mild-to-Moderate PD Patients and Effect of Spatiotemporal Gait Parameters, Laterality of Symptoms, Pharmacological State, and Clinical Variables. Sensors (Basel) 2022 Dec 25;23(1):214. PMID:36616812

21. Kim DW, Hassett LM, Nguy V, Allen NE. A Comparison of Activity Monitor Data from Devices Worn on the Wrist and the Waist in People with Parkinson's Disease. Mov Disord Clin Pract 2019 Nov;6(8):693–699. PMID:31745480

22. Wendel N, Macpherson CE, Webber K, Hendron K, DeAngelis T, Colon-Semenza C, Ellis T. Accuracy of Activity Trackers in Parkinson Disease: Should We Prescribe Them? Phys Ther 2018 Aug 1;98(8):705–714. PMID:29718452

23. Ginis P, Goris M, De Groef A, Blondeel A, Gilat M, Demeyer H, Troosters T, Nieuwboer A. Validation of Commercial Activity Trackers in Everyday Life of People with Parkinson's Disease. Sensors (Basel) 2023 Apr 21;23(8):4156. PMID:37112496

24. Bianchini E, Galli S, Alborghetti M, De Carolis L, Zampogna A, Hansen C, Vuillerme N, Suppa A, Pontieri FE. Four Days Are Enough to Provide a Reliable Daily Step Count in Mild to Moderate Parkinson's Disease through a Commercial Smartwatch. Sensors (Basel) 2023 Nov 4;23(21):8971. PMID:37960670

25. Outeiro TF, Alcalay RN, Antonini A, Attems J, Bonifati V, Cardoso F, Chesselet M-F, Hardy J, Madeo G, McKeith I, Mollenhauer B, Moore DJ, Rascol O, Schlossmacher MG, Soreq H, Stefanis L, Ferreira JJ. Defining the Riddle in Order to Solve It: There Is More Than One "Parkinson's Disease." Mov Disord 2023 Jul;38(7):1127–1142. PMID:37156737

26. Jankovic J, McDermott M, Carter J, Gauthier S, Goetz C, Golbe L, Huber S, Koller W, Olanow C, Shoulson I. Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort. The Parkinson Study Group. Neurology 1990 Oct;40(10):1529–1534. PMID:2215943

27. Stebbins GT, Goetz CG, Burn DJ, Jankovic J, Khoo TK, Tilley BC. How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. Mov Disord 2013 May;28(5):668–670. PMID:23408503

28. Zetusky WJ, Jankovic J, Pirozzolo FJ. The heterogeneity of Parkinson's disease: clinical and prognostic implications. Neurology 1985 Apr;35(4):522–526. PMID:3982637

29. Cardoso F, Goetz CG, Mestre TA, Sampaio C, Adler CH, Berg D, Bloem BR, Burn DJ, Fitts MS, Gasser T, Klein C, de Tijssen MAJ, Lang AE, Lim S-Y, Litvan I, Meissner WG, Mollenhauer B, Okubadejo N, Okun MS, Postuma RB, Svenningsson P, Tan LCS, Tsunemi T, Wahlstrom-Helgren S, Gershanik OS, Fung VSC, Trenkwalder C. A Statement of the MDS on Biological Definition, Staging, and Classification of Parkinson's Disease. Mov Disord 2024 Feb;39(2):259–266. PMID:38093469

30. Höglinger GU, Adler CH, Berg D, Klein C, Outeiro TF, Poewe W, Postuma R, Stoessl AJ, Lang AE. A biological classification of Parkinson's disease: the SynNeurGe research diagnostic criteria. Lancet Neurol 2024 Feb;23(2):191–204. PMID:38267191

31. Horsager J, Borghammer P. Brain-first vs. body-first Parkinson's disease: An update on recent evidence. Parkinsonism Relat Disord 2024 Mar 15;106101. PMID:38519273

32. 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):2129–2170. PMID:19025984

33. Schlachetzki JCM, Barth J, Marxreiter F, Gossler J, Kohl Z, Reinfelder S, Gassner H, Aminian K, Eskofier BM, Winkler J, Klucken J. Wearable sensors objectively measure gait parameters in Parkinson's disease. PLoS One 2017;12(10):e0183989. PMID:29020012

34. Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, Giladi N, Holloway RG, Moore CG, Wenning GK, Yahr MD, Seidl L, Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. Mov Disord 2004 Sep;19(9):1020–1028. PMID:15372591

35. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695–699. PMID:15817019

36. Jost ST, Kaldenbach M-A, Antonini A, Martinez-Martin P, Timmermann L, Odin P, Katzenschlager R, Borgohain R, Fasano A, Stocchi F, Hattori N, Kukkle PL, Rodríguez-Violante M, Falup-Pecurariu C, Schade S, Petry-Schmelzer JN, Metta V, Weintraub D, Deuschl G, Espay AJ, Tan E-K, Bhidayasiri R, Fung VSC, Cardoso F, Trenkwalder C, Jenner P, Ray Chaudhuri K, Dafsari HS, International Parkinson and Movement Disorders Society Non-Motor Parkinson Disease Study Group. Levodopa Dose Equivalency in Parkinson's Disease: Updated Systematic Review and Proposals. Mov Disord 2023 Jul;38(7):1236–1252. PMID:37147135

37. Martínez-Martín P, Rodríguez-Blázquez C, Mario Alvarez null, Arakaki T, Arillo VC, Chaná P, Fernández W, Garretto N, Martínez-Castrillo JC, Rodríguez-Violante M, Serrano-Dueñas M, Ballesteros

D, Rojo-Abuin JM, Chaudhuri KR, Merello M. Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale. Parkinsonism Relat Disord 2015 Jan;21(1):50–54. PMID:25466406

38. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016 Jun;15(2):155–163. PMID:27330520

39. Knol MJ, Pestman WR, Grobbee DE. The (mis)use of overlap of confidence intervals to assess effect modification. Eur J Epidemiol 2011 Apr;26(4):253–254. PMID:21424218

40. Austin PC, Hux JE. A brief note on overlapping confidence intervals. J Vasc Surg 2002 Jul;36(1):194–195. PMID:12096281

41. Cho J, Seo DM, Uh Y. Clinical Application of Overlapping Confidence Intervals for Monitoring Changes in Serial Clinical Chemistry Test Results. Ann Lab Med 2020 May;40(3):201–208. PMID:31858759

42. Balaguier R, Madeleine P, Vuillerme N. Intra-session absolute and relative reliability of pressure pain thresholds in the low back region of vine-workers: ffect of the number of trials. BMC Musculoskelet Disord 2016 Aug 18;17(1):350. PMID:27538914

43. Paul SS, Ellis TD, Dibble LE, Earhart GM, Ford MP, Foreman KB, Cavanaugh JT. Obtaining Reliable Estimates of Ambulatory Physical Activity in People with Parkinson's Disease. J Parkinsons Dis 2016 May 5;6(2):301–305. PMID:27164042

44. Yao J, Tan CS, Lim N, Tan J, Chen C, Müller-Riemenschneider F. Number of daily measurements needed to estimate habitual step count levels using wrist-worn trackers and smartphones in 212,048 adults. Sci Rep 2021 May 5;11(1):9633. PMID:33953288

45. Clemes SA, Griffiths PL. How many days of pedometer monitoring predict monthly ambulatory activity in adults? Med Sci Sports Exerc 2008 Sep;40(9):1589–1595. PMID:18685533

46. Sigmundová D, Vašíčková J, Stelzer J, Repka E. The influence of monitoring interval on data measurement: an analysis of step counts of university students. Int J Environ Res Public Health 2013 Jan 28;10(2):515–527. PMID:23358235

47. Norris M, Anderson R, Motl RW, Hayes S, Coote S. Minimum number of days required for a reliable estimate of daily step count and energy expenditure, in people with MS who walk unaided. Gait Posture 2017 Mar;53:201–206. PMID:28199925

48. Fini NA, Burge AT, Bernhardt J, Holland AE. Two Days of Measurement Provides Reliable Estimates of Physical Activity Poststroke: An Observational Study. Arch Phys Med Rehabil 2019 May;100(5):883–890. PMID:31030730

49. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005 Sep;84(9):719–723. PMID:16141752

50. Handlery R, Stewart JC, Pellegrini C, Monroe C, Hainline G, Flach A, Handlery K, Fritz S. Physical Activity in De Novo Parkinson Disease: Daily Step Recommendation and Effects of Treadmill Exercise on Physical Activity. Phys Ther 2021 Oct 1;101(10):pzab174. PMID:34244805

51. Treacy D, Hassett L, Schurr K, Chagpar S, Paul SS, Sherrington C. Validity of Different Activity Monitors to Count Steps in an Inpatient Rehabilitation Setting. Phys Ther 2017 May 1;97(5):581–588. PMID:28339904

52. Fokkema T, Kooiman TJM, Krijnen WP, VAN DER Schans CP, DE Groot M. Reliability and Validity of

Ten Consumer Activity Trackers Depend on Walking Speed. Med Sci Sports Exerc 2017 Apr;49(4):793–800. PMID:28319983

53. Chow JJ, Thom JM, Wewege MA, Ward RE, Parmenter BJ. Accuracy of step count measured by physical activity monitors: The effect of gait speed and anatomical placement site. Gait Posture 2017 Sep;57:199–203. PMID:28666177

54. Lamont RM, Daniel HL, Payne CL, Brauer SG. Accuracy of wearable physical activity trackers in people with Parkinson's disease. Gait Posture 2018 Jun;63:104–108. PMID:29729611

55. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010 May;19(4):539–549. PMID:20169472

56. Erro R, Vitale C, Amboni M, Picillo M, Moccia M, Longo K, Santangelo G, De Rosa A, Allocca R, Giordano F, Orefice G, De Michele G, Santoro L, Pellecchia MT, Barone P. The Heterogeneity of Early Parkinson's Disease: A Cluster Analysis on Newly Diagnosed Untreated Patients. PLoS One 2013 Aug 1;8(8):e70244. PMID:23936396

57. Wüllner U, Borghammer P, Choe C-U, Csoti I, Falkenburger B, Gasser T, Lingor P, Riederer P. The heterogeneity of Parkinson's disease. J Neural Transm (Vienna) 2023 Jun;130(6):827–838. PMID:37169935

58. Berg D, Borghammer P, Fereshtehnejad S-M, Heinzel S, Horsager J, Schaeffer E, Postuma RB. Prodromal Parkinson disease subtypes - key to understanding heterogeneity. Nat Rev Neurol 2021 Jun;17(6):349–361. PMID:33879872

59. Fasano A, Fung VSC, Lopiano L, Elibol B, Smolentseva IG, Seppi K, Takáts A, Onuk K, Parra JC, Bergmann L, Sail K, Jalundhwala Y, Pirtosek Z. Characterizing advanced Parkinson's disease: OBSERVE-PD observational study results of 2615 patients. BMC Neurol 2019 Apr 2;19(1):50. PMID:30940119

60. Stefani A, Tessitore A, Tambasco N, Cossu G, Ceravolo MG, Defazio G, Morgante F, Ramat S, Melzi G, Gualberti G, Merolla R, Onuk K, Lopiano L. Criteria for identification of advanced Parkinson's disease: the results of the Italian subgroup of OBSERVE-PD observational study. BMC Neurol 2022 Jan 28;22(1):41. PMID:35090406

61. Alves G, Larsen JP, Emre M, Wentzel-Larsen T, Aarsland D. Changes in motor subtype and risk for incident dementia in Parkinson's disease. Mov Disord 2006 Aug;21(8):1123–1130. PMID:16637023

62. Che N, Ou R, Li C, Zhang L, Wei Q, Wang S, Jiang Q, Yang T, Xiao Y, Lin J, Zhao B, Chen X, Shang H. Plasma GFAP as a prognostic biomarker of motor subtype in early Parkinson's disease. NPJ Parkinsons Dis 2024 Mar 1;10(1):48. PMID:38429295

Supplementary Files

Figures

ICC and 83.4% CI comparison for each subgroup pairs. Panel A: PwPD with reported presence or absence of tremor. Panel B: PwPD with mild and moderate disease severity as indicated by MDS-UPDRS. Panel C: PwPD with earlier (mHY 1-2) and intermediate (mHY 2.5-3) disease stage. Panel D: PIGD and TD phenotypes. ICC (3,k) point estimate is indicated by the black dot, CI limits are represented by the vertical bars. CI: confidence interval; ICC: intraclass correlation coefficient; MDS-UPDRS: International Movement Disorders Society Unified Parkinson Disease Rating Scale; mHY: modified Hoehn & Yahr stage; PIGD: postural instability and gait disorder; TD: tremor dominant.