

Digital Measures Track Progression in Early Parkinson's Disease Over 36 Months in the WATCH-PD Study

Varun Reddy

University of Rochester Medical Center

Yishu Gong

Takeda Pharmaceuticals

Peggy Auinger

University of Rochester Medical Center

Aaron Lerner

University of Rochester Medical Center

Thao-Uyen Dang

University of Rochester Medical Center

Tatiana Solodova

University of Rochester Medical Center

Thimaikhue Nguyen

University of Rochester Medical Center

Matthew Stephen

University of Rochester Medical Center

Courtney Roe

University of Rochester Medical Center

Tairmae Kangarloo

AbbVie Pharmaceuticals

Brian Tracey

Takeda Pharmaceuticals

Joshua Cosman

AbbVie Pharmaceuticals

David Anderson

Clinical Ink

Christopher Tarolli

University of Rochester Medical Center

Melissa A. Kostrzebski

University of Rochester Medical Center

Martijn L.T.M. Müller

Manchester Metropolitan University

Diane Stephenson

Critical Path Institute

E. Ray Dorsey

University of Rochester Medical Center

Jamie L. Adams

jamie_adams@urmc.rochester.edu

University of Rochester Medical Center

Parkinson Study Group WATCH-PD Study Investigators and Coordinators

Article

Keywords: Parkinson disease, smartphone, wearable electronic devices, observational study, gait, tremor, motor activity, speech

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11097406/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. Dr. Adams has received research support from the Michael J. Fox Foundation for Parkinson's Research, Critical Path for Parkinson's, NIH/NINDS, FDA, Biogen, the Huntington Study Group, PhotoPharmics, and Novartis; received compensation as a consultant/steering committee/advisory board member from the Huntington Study Group, Parkinson Study Group, AbbVie, VisualDx, BioSensics, Sana Biotechnology, Neuron23, Lundbeck, Biohaven, Wave Life Sciences, Clintrex, Vima Therapeutics, and the Michael J. Fox Foundation for Parkinson's Research; received honoraria for speaking from the Huntington Study Group, the Parkinson Study Group, and the American Neurological Association. Dr. Tairmae Kangarloo, Dr. Yishu Gong, Dr. Brian Tracey are/were employees of and own stock in Takeda Pharmaceuticals, Inc. Dr. Anderson is an employee of Clinical ink. Ms. Kostrzebski has received funding from the Michael J. Fox Foundation for Parkinson's Research, the NIH and the Department of Defense and she holds stock in Apple, Inc. Dr. Muller has received funding from the Michael J. Fox Foundation for Parkinson's Research Dr. Dorsey has received compensation for consulting services from Abbott, Abbvie, Acadia, Acorda, Bial-Biotech Investments, Inc., Biogen, Boehringer Ingelheim, California Pacific Medical Center, Caraway Therapeutics, Curasen Therapeutics, Denali Therapeutics, Eli Lilly, Genentech/Roche, Grand Rounds, Huntington Study Group, Informa Pharma Consulting, Karger Publications, LifeSciences Consultants, MCM Education, Mediflix, Medopad, Medrhythms, Merck, Michael J. Fox Foundation, NACCME, Neurocrine, NeuroDerm, NIH, Novartis, Origent Data Sciences, Otsuka, Physician's Education Resource, Praxis, PRIME Education, Roach, Brown,

McCarthy & Gruber, Sanofi, Seminal Healthcare, Spark, Springer Healthcare, Sunovion Pharma, Theravance, Voyager and WebMD; research support from Biosensics, Burroughs Wellcome Fund, CuraSen, Greater Rochester Health Foundation, Huntington Study Group, Michael J. Fox Foundation, National Institutes of Health, Patient-Centered Outcomes Research Institute, Pfizer, PhotoPharmics, Safra Foundation, and Wave Life Sciences; editorial services for Karger Publications; stock in Included Health and in Mediflix, and ownership interests in SemCap. Dr. Tarolli has received research support from the Michael J. Fox Foundation for Parkinson's Research, NIH/NINDS, Abbvie, Biogen, and Neuron23; received honoraria for speaking for the American Academy of Neurology; received honoraria as a workgroup member for the National Board of Medical Examiners. Dr. Wyant received compensation for consulting from Biovie, royalties from UpToDate, Medmastery, and research support from UCB BioPharma, Micheal J Fox Foundation, and PhotoPharmics. Dr. Espay has received grant support from the NIH and the Michael J Fox Foundation; personal compensation as a consultant/scientific advisory board member for Mitsubishi Tanabe Pharma America (formerly, Neuroderm), Amneal, Acadia, Acorda, Abbvie, Supernus (formerly, USWorldMeds), NeuroDiagnostics, Inc (SYNAPS Dx), Intrance Medical Systems, Inc., Praxis Precision Medicines, and Herantis Pharma; Data Safety Monitoring Board (chair) of AskBio; and publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer. He is co-inventor of the patent "Compositions and methods for treatment and/or prophylaxis of proteinopathies" with which he cofounded REGAIN Therapeutics to fund preclinical studies. He has no financial relationship with the company and relinquished the right to any personal income from future treatments. Dr. Espay owns no stock in any pharmaceutical company with which he has a scientific advisory relationship. Dr. Cloud has had clinical trial contracts (paid to her institution) with Cerevel, UCB, Amneal, Blue Rock Therapeutics, Biohaven, and Intra-Cellular Technologies. She has two inventions (Gastrointestinal Symptoms in Neurodegenerative Disease (GIND) scale, VCU Office of Technology Transfer # CLO-11-0R67 and Recognize and Deploy Vibration for Mitigation of Freezing of Gait, VCU Office of Technology Transfer provisional patent #PRE-21-137F (322203-8130), application filed July 2022, full patent application filed July 2023. She is 27% owner of Motion Medix, LLC.

Abstract

Digital tools can provide objective, sensitive, real-world measures of disease progression in Parkinson's disease (PD). However, few studies have evaluated their longitudinal performance beyond one year. We previously demonstrated significant changes in digital gait and tremor measures over 12 months in individuals with early PD in the multicenter WATCH-PD study. Here, we extend the follow up to 36 months in 43 participants in a remote study using smartphone and smartwatch-based assessments to determine which digital measures track disease progression.

Over 36 months, a composite gait score – derived from seven gait features (arm swing, cadence, double support time, gait speed, stance time, stride length, and turn duration) – increased from 3.0 (interquartile range 0.2-4.8) at baseline to 4.2 (2.2-6.3; $p=0.02$). At month 36, turn duration increased significantly while gait speed and arm swing decreased modestly. A digital speech composite also worsened over 36 months from 1.00 (-0.27-2.18) to 1.76 (-0.19-3.31; $p=0.02$). After a likely initial learning effect, psychomotor measures remained stable, and measures of activity and cognition showed no significant change. Time with tremor, measured passively by the smartwatch, increased during the first year but returned to baseline levels by month 36, possibly reflecting initiation of dopaminergic therapy.

Among numerous digital measures evaluated over 36 months, a composite gait measure showed the clearest signal of disease progression. Speech also demonstrated measurable longitudinal progression. Other digital measures were either relatively insensitive to change, altered by learning effects, or substantially influenced by symptomatic treatment. These findings support prioritizing composite digital gait measures, together with speech assessments, as candidate outcome measures for future disease-modifying trials in early PD.

INTRODUCTION

Parkinson's disease (PD) is a debilitating neurodegenerative condition characterized by a wide range of motor (e.g. tremor, bradykinesia, gait changes, hypophonia) and non-motor features (e.g. cognitive decline, speech/language dysfunction). The disease affects an estimated 11.8 million individuals worldwide, with rising prevalence, disability and deaths due to PD^{1,2}.

The development of disease modifying therapies for slowing the progression of PD has been a priority for researchers but has been hampered by many factors including a lack of sensitive markers for disease progression³. Current scales for characterization of disease progression like the Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) lack sensitivity to subtle changes in symptoms of early PD and have substantial intra- and inter-rater variability⁴⁻⁶.

Research into digital health technologies like smartphone applications and wearable sensors is a potential avenue to overcome the limitations of current clinical outcome assessments. These devices can provide sensitive measures of the disease through the collection of rater-independent, objective, and real-time data on PD progression^{7,8}. The WATCH-PD study explored how a smartwatch, smartphone, and

research-grade wearable sensors compared to traditional clinician-rated assessments in tracking early PD (n = 82 with PD diagnosis less than 2 years, Hoehn & Yahr stage 2 or lower) in clinic and at home. We previously reported the baseline and one-year results which showed that the digital devices detected significant differences in multiple domains (gait, tremor, fine motor, speech, cognition) compared to age-matched controls and measurable progression in metrics of gait and tremor. The correlations with traditional rating scales were variable⁹.

We extended the original WATCH-PD study in a remote, decentralized trial design (all visits and assessments were remote). Participants used a smartwatch and smartphone to complete active assessments and passive monitoring daily for seven days every three months. This extension compiled longitudinal data that can inform the use of digital endpoints in future, long duration disease-modifying trials. Here, we report results out to 36 months, focusing on the ability of the smartphone and smartwatch measures to detect change over this period. Given that most participants started dopaminergic medication by year 3, we also assessed changes in digital measures following initiation of therapy. Finally, a subset of participants remained off PD medications for the entire 36 months, which allowed for analysis of a medication-naïve population.

RESULTS

Study participants

Forty-eight individuals from the original WATCH-PD study consented to participate in the remote extension study from February through July 2023 and 43 were included in these analyses (five excluded due to: two with their first extension visit after 36 months; two with deep brain stimulation not considered early PD; one who did not withhold dopaminergic medication during study assessments) (**Figure 1**). As shown in **Table 1**, participants had a mean (standard deviation) age of 66.7 (8.3) years with the majority being white and college educated. At month 36, the mean disease duration was 3.9 (0.7) years, and 31 individuals (72%) had started dopaminergic medication. Twelve participants did not start dopaminergic therapy over the first three years of observation. Half of these medication-naïve individuals were female, and all were Hoehn and Yahr stage 1 or 2 at 36 months.

Data collection

Every three months, participants performed a battery of active assessments (gait, speech, fine motor, finger-tapping) completed daily in the OFF state for seven days and wore the smartwatch daily for continuous passive activity and tremor monitoring during the same seven days. MDS-UPDRS I-IV and patient-reported outcome measures were collected remotely every 6 months.

Availability of the gait task averaged 4.7 days per week (SD = 2.2) with a median of 5 days. All other active tasks and passive monitoring metrics have a median of 7 days per week.

Clinical measures

Total MDS-UPDRS scores, reported as median (interquartile range; n), worsened from 30.0 (23.0-39.0; 43) at baseline to 39.0 (30.0-44.5; 43) at month 36 ($p < 0.001$). Part I scores increased from 6.0 (3.0-8.0; 43) to 7.0 (3.0-11.0; 43) ($p = 0.18$), and Part II scores remained stable, with a median of 6.0 (2.0-9.0; 43) at baseline and 6.0 (3.0-9.5; 43) at month 36 ($p = 0.26$). Part III OFF scores worsened from 19.0 (12.5-24.5; 43) to 25.0 (16.5-32.5; 43) ($p = 0.004$; Table 2A). In the medication-naïve group, Part III OFF scores worsened from 15.0 (9.8-21.0; 12) to 29.0 (18.2-33.5; 12) ($p < 0.001$; Table 2B). For those taking medication, Part III OFF scores changed little from 21.0 (14.0-25.5; 31) to 23.0 (16.0-29.0; 31) ($p = 0.15$; Table 2C). After adjustment for levodopa equivalent daily dose (LEDD), the estimated 36-month change in Part III OFF scores worsened from 19.7 (14.2-23.0; 41) at baseline to 24.1 (18.6-27.5; 41) at month 36 ($p = 0.07$; Table 2D), with progression assessed under a common unmedicated reference state.

Gait

Digital measures of gait declined modestly among individuals with early PD (**Figure 2 and Table 2A**). As measured by the smartwatch, arm swing from the most affected side decreased from 22.5 degrees (16.5-31.4; 36) at baseline to 17.9 degrees (13.2-28.0; 37) at month 36 ($p = 0.11$). Among those who did not initiate dopaminergic therapy, the decline tended to be greater, from 30.5 degrees (14.2-41.0; 9) at baseline to 14.2 degrees (9.7-33.4; 10) at month 36 ($p = 0.29$) (**Table 2B**).

As measured by the smartphone for the entire cohort, median gait speed showed a small decline from 0.92 meters per second (m/s) (0.82-1.01; 37) at baseline to 0.89 m/s (0.79-0.98; 38) at month 36 ($p = 0.10$) (**Table 2A**). For the medication-naïve participants, gait speed slowed from 0.85 m/s (0.81-1.00; 9) to 0.77 m/s (0.66-0.96; 10) ($p = 0.18$) (**Table 2B**).

The group median turn duration increased from 2.30 seconds (2.05-2.50; 41) at baseline to 2.40 seconds (2.30-2.69; 38) by month 36 ($p = 0.02$). For medication-naïve participants, the turn duration increased from 2.29 seconds (2.15-2.60; 10) to 2.50 seconds (2.30-2.85; 10) ($p = 0.25$). The digital gait composite score (a standardized composite of seven gait metrics that differentiate PD patient from controls at baseline⁹) increased from 3.0 (0.2-4.8; 36) at baseline to 4.2 (2.2-6.3; 37) at month 36 ($p = 0.02$). Among those who did not initiate dopaminergic therapy, the increase in digital composite scores were larger, from 3.0 (-1.7-5.6; 9) at baseline to 5.4 (0.8-8.0; 10) at month 36 ($p = 0.10$). The medicated group has an increase in digital composite scores from 3.0 (0.8-4.6; 27) at baseline to 4.2 (2.3-5.3; 27) at month 36 ($p = 0.12$).

For the individual digital measures, we applied a rank transformation and calculated a standardized change¹⁰. The magnitude of this standardized change (in terms of worsening) was observed as 0.27 for arm swing, 0.25 for gait speed, and 0.50 for turn duration, and 0.36 for the digital gait composite score. For comparison, item 2.12 on the MDS-UPDRS (self-reported problems with walking and balance) showed a standardized change of 0.24, while item 3.10 (rater evaluation of gait) showed a standardized change of 0.58.

Tremor

As shown in **Figure 3A**, for the medication-naïve group, there was a consistent non-significant increase in percentage of time with tremor with the median increasing from 6.8% (0.6-14.9; 8) at baseline to 10.1% (2.6-20.2; 10) by month 36 ($p = 0.36$). In contrast as shown in **Figure 3B**, in the medication group, percentage of time with tremor increased from 17.9% (3.1-33.9; 18) at baseline to 30.7% (19.5-58; 12) at month 12; however, by month 36, these participants were all on dopaminergic medication and time with tremor fell to 9.8% (3.2-44.3; 26) with $p=0.62$ vs. baseline.

Activity

As assessed by the smartwatch, the median daily step count was 3044 (1477-4350; 21) at baseline and 3492 (1221-5274; 36) at month 36 ($p=0.56$). For medication-naïve participants, although non-significant, the median daily step count decreased from 5416 (3419-6576; 4) at baseline to 2440 (165-5190; 10) at month 36 ($p=0.77$). As shown in **Table 2A** for all participants, the median steps per hour was 280 (166-428; 21) at baseline and 283 (143-460; 36) at month 36 ($p=0.89$). The median steps per hour reduced non-significantly from 376 (217-536; 4) to 311 (36-638; 10) in the medication-naïve group ($p = 0.80$) (**Table 2B**).

Speech

The at-home digital speech composite scores (a standardized composite of four speech metrics that differentiate PD patients from controls at baseline⁹) worsened from 1.00 (-0.27-2.18; 35) at baseline to 1.76 (-0.19-3.31; 41) at month 36 ($p=0.02$). The median pause time during the reading task, a key component of the composite score, also increased during the same time period, from 0.57 seconds (0.49-0.69; 35) at baseline to 0.63 seconds (0.52-0.78; 41) at month 36 ($p=0.03$). In participants without medication, digital speech scores rose from 1.8 to 2.9 at month 36 ($p = 0.27$). In participants taking medication, scores rose from 0.98 to 1.71 at month 36 ($p = 0.08$). Overall, the composite score showed a greater standardized change in magnitude of 0.33 compared to 0.23 for item 2.01 (self-reported speech) and 0.22 for item 3.01 (rater evaluation of speech).

Psychomotor

Some psychomotor metrics showed significant changes over 36 months as shown in **Table 2A and Figure 4**. In the dominant hand, the median total taps increased from 126 (101-154; 41) at baseline to 171 (152-195; 24) at month 12 and remained at 168 (120-192; 41) at month 36 ($p < 0.001$ vs. baseline). A similar trend was observed in the nondominant hand, where the median total taps increased from 128 (102-159; 41) at baseline to 170 (128-190; 24) at month 12 and remained at 158 (132-184,41) at month 36 ($p < 0.001$ vs. baseline). For those who did not start dopaminergic therapy, finger tapping in the dominant hand increased from 144 (128-169; 10) at baseline to 185 (169-198; 8) at month 12, and reached 191 (156-199; 10) at month 36 ($p = 0.08$ vs. baseline); finger tapping in the nondominant hand increased from 140 (112-163; 10) at baseline to 182 (136-191; 8) at month 12 and remained at 180 (152-184; 10) at month 36 ($p = 0.01$ vs. baseline). This increase likely reflects a learning effect seen in the 12-month study as total taps remained at a comparable level from month 12 to month 36. The median inter-

tap interval measured at baseline was 164 milliseconds (ms) (82-213; 39) and at month 36 was 118 ms (101-156; 41) ($p=0.55$) for the dominant hand and 170 ms (92-221; 38) at baseline and 120 ms (107-147; 41) at month 36 ($p=0.04$) for the nondominant hand. For the fine motor task, the median number completed remained stable over 36 months in PD: dominant hand, 4.0 (3.0-5.0, 41) to 4.0 (3.0-5.0, 41) ($p=1.0$); nondominant hand, 4.0 (3.0-5.0, 41) to 4.0 (3.0-5.0, 41) ($p=1.0$).

Cognition

As shown in **Table 2A**, the median time to completion on Trails A was 22.7 seconds (s) (18.3-28.2; 42) at baseline and 21.9 s (17.7-27.2; 42) at month 36 ($p=0.40$), following an initial decrease to 18.9 s (16.8-24.1; 24) at month 12 consistent with a potential learning effect. Similarly, the median time to completion on Trails B was 22.0 s (19.5-30.8; 42) at baseline and 22.9 s (19.7-25.5; 42) at month 36 ($p=0.92$), after an initial decrease to 18.0 s (16.5-22.8; 24) at month 12. The difference between Trails B and A (Trails B – Trails A) was nonsignificant, 0.9 s (-2.2-3.8; 42) at baseline to 1.3 s (-1.0-2.4; 42) at month 36 ($p = 0.95$). For the medication-naïve participants, the changes observed were generally similar to the whole cohort (**Table 2B**).

DISCUSSION

Over 36 months, composite digital measures of gait and speech collected from a smartphone and smartwatch demonstrated measurable progression in individuals with early PD. Individual measures of gait (e.g., turn duration, gait speed, arm swing) also worsened over 36 months, especially in those who never started dopaminergic medication. Importantly, the magnitude of change observed was greater for some measures than that observed with traditional participant or expert scores. These findings support growing evidence that digital gait and speech markers may capture subtle progression earlier and not readily detected by conventional rating scales^{7,8,11,12}. Our study corroborates findings that increased turn duration may be an early marker of progressive gait dysfunction¹³. While arm swing amplitude and gait speed showed modest, non-significant reductions across 36 months, in other studies, these changes were more significant and therefore continue to be measures of interest^{14–18}. Our digital speech composite scores significantly increased from baseline to month 12 and month 36, underscoring emerging evidence which suggests that early changes to speech and language (e.g., dysprosody) may precede motor symptoms and show measurable change over time^{11,19}. These subtle changes may be undetectable by patients or expert raters, and our results demonstrated superiority of the digital speech composite to patient-reported and clinician-rated speech items on the MDS-UPDRS. Furthermore, gait and speech composite scores showed progression after accounting for LEDD. This adjustment was particularly important given the small number of participants who remained completely medication-naïve throughout the 36-month study period.

Inherently, PD is a heterogeneous condition with a diverse range of motor and non-motor impairments, making disease progression challenging to track with single-item measures. Indeed, the traditional scales typically used as primary endpoints in PD clinical trials are composites. Our gait and speech

composite measures showed the most significant changes across 36 months, indicating that the future development of models or composite measures derived from multiple digital parameters may show better reliability and sensitivity in assessing PD progression^{7,20,21}. Already digital composite measures via the Roche digital platform have shown high reliability and validity in measuring motor progression in Parkinson's and Huntington's disease and are being explored as endpoints in trials^{22,23}. Questions remain regarding weighting of items and meaningful change of composite scores, which deserve thoughtful study, but it is becoming evident that this may be the best path for digital endpoints in early PD.

Digital measures of tremor are also of interest in early PD and a recent study found that sensor-derived tremor measures from a wrist-worn device were more sensitive to progression over two years than clinical tremor scores in a medication-naïve cohort²⁴. In our study, the median proportion of time with tremor, as measured passively by the smartwatch, did not demonstrate significant change across 36 months for all participants. Participants who initiated medications between months 12 and 36 showed an initial increase in proportion of time with tremor at month 12 followed by a decrease to baseline levels by month 36. This trend contrasts with individuals in the small medication-naïve cohort, in which proportion of time with tremor remained stable across all months. Given that tremor is largely recognized as one of the most burdensome PD symptoms, individuals who defer medication may have a milder progression of tremor or no tremor, reducing the need to start dopaminergic medication earlier on,²⁵ which could have been a factor in our results. Based on our results from year 1 and the literature,²⁴ passively collected digital measures of tremor from a wrist worn device may be useful in medication-naïve cohorts. However, not all individuals with PD experience tremor, and tremor measures are substantially affected by the initiation of symptomatic therapy, limiting their utility in long-duration studies or medicated cohorts.

In addition to tremor, other passive measures were confounded by medication effects including measures of daily activity (e.g. median daily steps, steps per hour) which did not show significant changes over 36 months for the entire cohort. For the small group of medication-naïve participants, median daily step count showed a nonsignificant but greater than 50% reduction, highlighting important considerations for PD medication initiation to stabilize motor activity early on and alter longitudinal symptom trajectories. In general, the divergence between medication-naïve participants and those receiving dopaminergic therapy, particularly for passive measures, reinforces the importance of accounting for symptomatic treatment when validating digital progression markers. We had participants complete active tasks in the OFF state prior to morning medications to try to mitigate some of the confounding effect. This may be a reasonable strategy for trials using active digital monitoring; however, it would not account for long-lasting symptomatic benefit, which can persist for days after the last dose²⁶. More research on effects of symptomatic medication on digital measure performance is needed to best inform selection of measures for trials that include treated populations.

Psychomotor and cognitive measures in our study have likely been limited by learning effects causing an initial increase in performance. Following a suspected learning effect through month 12, the psychomotor metrics in the study, including median total taps, remained similar from month 12 to month 36 in both dominant and non-dominant hands. Other digital studies found digital finger tap measures including tap speed and tap count to correlate with in-clinic assessments of PD severity and present with high test-retest reliability, although these studies included participants with later stages of PD^{22,27,28}. Cognitive tasks conducted on the smartphone showed a similar suspected learning effect between baseline to month 12, which, however, was followed by a return toward baseline in performance from month 12 to month 36. The development of validated cognition-specific digital measures for longitudinal tracking of PD remains an underexplored area of interest^{12,29}. Mitigating learning effects, perhaps with a run-in period where tasks are completed for the first time before the baseline visit, would be wise if planning to use these measures in trials.

Our study aimed to compile a multi-year data set for use in identifying core digital measures that can function as clinical outcome assessments or monitoring biomarkers in future clinical trials. Despite large scale initiatives to establish validated real-world digital measures³⁰, there are limited studies assessing the use of digital health technologies beyond 12 months in early PD. Brzezicki et al. assessed the changes in motor features in *de novo* PD patients across two years using six wearable sensors. They identified multiple gait and upper limb bradykinetic features that progressed in early PD, with masking of progression on initiation of symptomatic medication¹⁶. A separate observational study identified multiple digital measures—Toe Tapping Amplitude, Sit-to-Stand Time, Velocity of Arms and Legs—that were significantly associated with UPDRS scores over 30–36 months and responsive to dopamine replacement therapy³¹. As wearable sensors may offer objective and sensitive measurement of therapeutic effectiveness, recent clinical trials have begun to assess digital measures as exploratory outcomes and even primary endpoints for PD^{32,33}. Based on our data, gait and speech measures appear to be the most promising.

A distinguishing feature of this study is the remote, decentralized study design, which allowed for clinical visits to be conducted remotely. This framework may prove useful in future trials by eliminating the need for multiple research sites, reducing patient travel/expenses, speeding up trial timelines through faster recruitment, and enabling long term follow up of clinical trial cohorts³⁴. This is particularly pertinent in neurodegenerative diseases with slow progression like PD, where the impact of a therapeutic to delay progression may not be detected during the initial 12- or 18-month trial. Participants in our study demonstrated strong, consistent engagement with the use of consecutive daily assessments following each three-month visit, with median data availability ranging from 5–7 days for the various tasks. The integration of repeated consecutive assessments across a minimum of 4 days have important implications in remote designs, allowing the separation of variability from short-term contextual and behavioral factors from long-term progression while minimizing participant fatigue and burnout^{35,36}. Maintaining compliance rates has previously posed a challenge for smartphone-based PD monitoring, with the mPower study serving as a striking example of low engagement with less than 10% of enrolled

participants completing five or more days of data collection³⁷. In WATCH-PD, we have not had problems with compliance, despite the long duration of the study, largely due to our study design and participant engagement.

Importantly, the WATCH-PD study has engaged with key stakeholders including investigators, regulatory bodies, the pharmaceutical industry, and individuals with PD. Aligning with the FDA's patient-focused drug development guidances, our study simultaneously conducted qualitative interviews with WATCH-PD participants evaluating change in prevalence and burden of early PD symptoms and relevance of digital measures over time³⁸. The qualitative study identified gait and balance as becoming more significantly bothersome over 3 years, consistent with the changes in digital gait measures we identified and further supporting their candidacy for meaningful progression markers²⁵. Overall, patients continue to rate the relevance of the WATCH-PD digital measures highly, suggesting that the current measures are targeting symptoms that matter most to individuals with early PD.

This study has several limitations. The sample size was modest, which may limit our ability to detect subtle yet meaningful changes. Caution should also be applied when interpreting standardized change of a categorical item. Furthermore, the lack of healthy controls beyond baseline and the small number of medication-naïve participants made it challenging to detect significant differences when comparing this group to participants on medications. Additionally, despite participants holding off on taking their PD medications until completion of active assessments each day, there remains a symptomatic benefit of PD medications (i.e., levodopa) which may persist for days after the last dose²⁶. Therefore, analyses of the entire cohort may understate true progression of PD as participants may perform assessments above their disease baseline. Maintaining a medication-naïve cohort across multiple years remains an ethical challenge and is difficult to implement in practice. Passive digital data collection faces challenges with varied environments as even differences in room size and arrangement of furniture have been identified as possible sources of gait variability in passive digital monitoring³⁹. Recently, an advisory committee, consisting of gait experts and patients with PD, developed a standardized framework to quantitatively assess gait using structured gait tasks, harmonize gait data collection across PD studies and allow for meaningful cross-study comparisons⁴⁰. Thus, future disease-modifying clinical trials may benefit from integrating standardized active assessments with passive digital monitoring in trial designs, leveraging the reproducibility of active standardized protocols with the ecological validity provided by passive real-world monitoring. The use of smart study designs (e.g. 7-day data collection bursts), improved technology, and pooled digital datasets for medication-naïve cohorts may further account for potential confounders.

Limitations notwithstanding, we conducted one of the longest digital measures studies in early PD to date, including both active and passive monitoring of motor and non-motor symptoms. Among the digital measures evaluated over three years, gait, particularly when summarized as a composite of complementary gait features, provided the most consistent signal of disease progression. These findings support prioritizing multidimensional digital gait endpoints, together with speech measures, in

future selection of digital outcome measures for clinical trials of symptomatic and putative disease-modifying therapies for PD.

METHODS

Study design, setting, participants

As described previously, WATCH-PD (Wearable Assessment in The Clinic and at Home in PD, NCT03681015) was a 12-month, multicenter observational study that evaluated the ability of digital devices to assess disease features and progression in persons with early, untreated PD compared to controls⁴¹. PD participants for this study were recruited from the original WATCH-PD cohort and evaluated remotely as part of a decentralized study design (Fig. 1). Remote visits occurred at screening/baseline and every three months thereafter. Inclusion criteria required prior completion of the original WATCH-PD study. Exclusion criteria included the presence of an alternative parkinsonian diagnosis or unreliable internet access.

Ethics

The University of Rochester Institutional Review Board approved the procedures used in the study, and there was full compliance with human experimentation guidelines. All participants provided written informed consent before study participation.

Data sources/measurement

As described previously⁴¹, this study used two devices: an Apple Watch S8, and an iPhone 12 (Apple, Inc.) running a smartphone application specifically for PD (BrainBaseline™). The smartphone application consisted of cognitive, speech, and psychomotor tasks including Trail Making Test, modified Symbol Digit Modalities Test, phonation, reading, diadochokinetic speech tasks, two-timed fine motor tests, and tremor, gait, and balance tasks.

Smartphone application tasks were conducted at each remote visit (day 1) and then for the next 6 days every three months (7 day bursts every 3 months). Participants were asked to complete the smartphone assessments in the OFF state and withhold PD medications until after the smartphone assessments were completed each day, in order to have consistency between measures derived from the active tasks from baseline where participants had not started PD medication. The smartphone was worn in a lumbar sport pouch during gait and balance tests. Participants wore the smartwatch on their more affected side and tracked symptoms on the smartphone daily for at least one week following each remote visit.

Gait features were extracted from the smartwatch and smartphone using a modified version of GaitPy⁴². The algorithm for extracting arm swing is a refactored version implemented in python based on

Warmerdam et al⁴³. Phonation and reading files were processed using custom Python code with features computed using the Parselmouth interface to Praat and the Librosa library. Common speech endpoints, such as jitter, shimmer, pitch statistics and Mel Frequency Cepstral Coefficients (MFCC), were computed.

Accelerometry data and tremor scores were collected from the smartwatch via Apple's Movement Disorders Application Programming Interface (API) during the passive monitoring periods⁴⁴. The Movement Disorders API generates tremor classification scores (none, slight, mild, moderate, strong, or unknown) for each one-minute period, and the fraction of time spent in each category was calculated for each participant⁴⁵.

Clinical measures

Every 6 months, participants remotely completed traditional rating scales including the MDS-UPDRS Parts I-IV (modified part III excluded rigidity and postural stability assessments), Montreal Cognitive Assessment, Modified Hoehn and Yahr, Geriatric Depression Scale, REM Sleep Behavior Disorder Questionnaire, Epworth Sleepiness Scale, Scale for Outcomes in Parkinson's Disease for Autonomic Symptoms, and the Parkinson's Disease Questionnaire-8.

Study size

The study was powered to detect a mean change over 36 months in MDS-UPDRS Part III scores and thus sufficient to detect changes in digital endpoints considered to have superior responsiveness to the MDS-UPDRS Part III. The mean change in part III from baseline to three years in combined untreated and treated OFF individuals with early PD in the PPMI study was 8.3 points with a standard deviation of 11.1. With a sample size of at least 40 participants, the study had more than 99% power to detect a true change of 8.3 points using a one-sample t-test and a two-tailed 5% significance.

Feature extraction

Speech features were extracted following the same methods as described in the month 12 manuscript⁹. Following that work, a composite speech score was formed by normalizing selected individual features using the normative cohort statistics at baseline (i.e., subtracting off the control baseline mean, dividing by the control baseline standard deviation, where a composite score of zero represents the average profile of a healthy control at baseline). The composite was formed using the four speech features (mean pause time in the reading task, monopitch features in the reading task, MFCC from the phonation task and cepstral peak prominence from the phonation task) that differed between individuals with and without PD at baseline⁹. Higher values indicate increasing clinical impairment. In this analysis, we applied this same methodology to calculate a gait composite score based on selected gait features,

including stride length, double support time, stance time, gait speed, arm swing, cadence, and turn duration which best differentiated those with PD from controls⁹. Step count statistics were acquired using the Forest Oak package implemented in Python^{46,47}.

Statistical methods

At month 36, the majority of participants were on PD medication, and assessments were conducted using a burst design in which tasks were performed in the OFF-medication state. This differs from the original WATCH-PD study⁹, where assessments were conducted biweekly at home over 12 months.

For each feature, repeated measurements were first summarized within each individual using the median to generate representative values at months 0, 3, 6, 9, 12, and 36, combining data from both the original and extension study phases. Population-level summaries were then obtained by averaging across individuals, with missing data omitted.

Longitudinal changes were assessed by comparing follow-up months with baseline (month 0) using the Hodges-Lehmann median shift, with values derived via bootstrap resampling (500 iterations). Standardized effect sizes were quantified using a rank-transformed standardized change by dividing the mean of the difference between month 36 and baseline assessments by the standard deviation of their difference¹⁰. Confidence intervals (95%) were estimated via nonparametric bootstrap resampling (100 iterations). As these approaches focus on different robust estimands (median vs. standardized change)—they do not necessarily correspond and may yield different statistical inferences.

To adjust for potential confounding introduced by concomitant medication and baseline disease severity, a baseline-adjusted quantile ANCOVA regression model adjusting for levodopa equivalent daily dose (LEDD), baseline MDS-UPDRS total score and sex was implemented for each digital biomarker feature. For each participant, baseline and Month 36 individual trajectory values were predicted under the unmedicated reference state. Statistical significance for 36-month disease progression was obtained directly from the quantile regression model coefficient for the 36-month time parameter. Standardized change (Cohen's d) and corresponding 95% confidence intervals were derived by scaling the model progression parameter and its standard error by the cohort baseline standard deviation. Statistical significance for 36-month disease progression was obtained directly from the quantile regression model coefficient for the 36-month time parameter. Standardized change (Cohen's d) and corresponding 95% confidence intervals were derived by scaling the model progression parameter and its standard error by the cohort baseline standard deviation. All statistical tests were two-sided with significance evaluated at $\alpha = 0.05$; given the exploratory nature of the study, P values were nominal and not adjusted for multiple comparisons.

Missing data

For active tasks, month 36 digital data were included for participants who were either medication-naïve or in the clinically defined OFF state (n = 43). Month 12 digital data are reported only for medication-naïve individuals to avoid the confounding effects of medication, as withholding was not required for active assessments at that time point. Participants were excluded from analysis if they had missing values for an outcome or required algorithm inputs; similarly, values of zero—indicating a task was not attempted—were excluded. For passive measures (e.g., tremor, step counts, and tremor-related MDS-UPDRS items), empirical values are reported as median (interquartile range; n) at baseline, month 12, and month 36. For step count data inclusion, we considered a participant’s data if there were a minimum of nine hours of wear time during the waking period (06:00–23:59). This nine-hour threshold is based on openly available step count data from the PPMI study, where step counts are collected using Verily Study Watches⁴⁸. The PPMI analysis revealed a Pearson’s correlation exceeding 0.9 between step counts derived from at least nine randomly selected hours within the wake period (6:00 am – 23:59 pm) and the overall step count for the entire waking period.

Declarations

Data Availability

The data are available to members of the Critical Path for Parkinson’s Consortium 3DT Initiative Stage 2. For those who are not a part of 3DT Stage 2, a proposal may be made to the WATCH-PD Steering Committee (via the corresponding author) for de-identified baseline datasets.

Code Availability

Custom Python code used for feature extraction and R code for statistical analysis is available from the authors upon request.

Acknowledgements

We thank Matthew Barrett, David Shprecher, Meredith Spindler, and Dmitri Volfson for their contribution and assistance in the conduct of the study.

Funding

Funding for the study was contributed by the Michael J. Fox Foundation for Parkinson’s Research (MJFF-022927).

Clinical Trial Registration

Clinical Trial number: not applicable

Author Contributions

J.L.A and B.T. had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: J.L.A. and E.R.D. Acquisition, analysis, or interpretation of the data: All authors. Drafting of the manuscript: J.L.A., V.R., A.L. T.K., Y.G., B.T., and P.A.. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Y.G., T.K., B.T., and P.A. Obtained funding: J.L.A. Administrative, technical, or material support: M.A.K. Supervision: J.L.A. and B.T.

Competing Interests

Dr. Adams has received research support from the Michael J. Fox Foundation for Parkinson's Research, Critical Path for Parkinson's, NIH/NINDS, FDA, Biogen, the Huntington Study Group, PhotoPharmics, and Novartis; received compensation as a consultant/steering committee/advisory board member from the Huntington Study Group, Parkinson Study Group, AbbVie, VisualDx, BioSensics, Sana Biotechnology, Neuron23, Lundbeck, Biohaven, Wave Life Sciences, Clintrex, Vima Therapeutics, and the Michael J. Fox Foundation for Parkinson's Research; received honoraria for speaking from the Huntington Study Group, the Parkinson Study Group, and the American Neurological Association. **Dr. Tairmae Kangarloo, Dr. Yishu Gong, Dr. Brian Tracey** are/were employees of and own stock in Takeda Pharmaceuticals, Inc. **Dr. Anderson** is an employee of Clinical ink. **Ms. Kostrzebski** has received funding from the Michael J. Fox Foundation for Parkinson's Research, the NIH and the Department of Defense and she holds stock in Apple, Inc. **Dr. Muller** has received funding from the Michael J. Fox Foundation for Parkinson's Research. **Dr. Dorsey** has received compensation for consulting services from Abbott, Abbvie, Acadia, Acorda, Bial Biotech Investments, Inc., Biogen, Boehringer Ingelheim, California Pacific Medical Center, Caraway Therapeutics, Curasen Therapeutics, Denali Therapeutics, Eli Lilly, Genentech/Roche, Grand Rounds, Huntington Study Group, Informa Pharma Consulting, Karger Publications, LifeSciences Consultants, MCM Education, Mediflix, Medopad, Medrhythms, Merck, Michael J. Fox Foundation, NACCME, Neurocrine, NeuroDerm, NIH, Novartis, Origent Data Sciences, Otsuka, Physician's Education Resource, Praxis, PRIME Education, Roach, Brown, McCarthy & Gruber, Sanofi, Seminal Healthcare, Spark, Springer Healthcare, Sunovion Pharma, Theravance, Voyager and WebMD; research support from Biosensics, Burroughs Wellcome Fund, CuraSen, Greater Rochester Health Foundation, Huntington Study Group, Michael J. Fox Foundation, National Institutes of Health, Patient-Centered Outcomes Research Institute, Pfizer, PhotoPharmics, Safra Foundation, and Wave Life Sciences; editorial services for Karger Publications; stock in Included Health and in Mediflix, and ownership interests in SemCap. **Dr. Tarolli** has received research support from the Michael J. Fox Foundation for Parkinson's Research, NIH/NINDS, Abbvie, Biogen, and Neuron23; received honoraria for speaking for the American Academy of Neurology; received honoraria as a workgroup member for the National Board of Medical Examiners.

Dr. Wyant received compensation for consulting from Biovie, royalties from UpToDate, Medmastery, and research support from UCB BioPharma, Micheal J Fox Foundation, and PhotoPharmics. **Dr. Espay** has received grant support from the NIH and the Michael J Fox Foundation; personal compensation as a consultant/scientific advisory board member for Mitsubishi Tanabe Pharma America (formerly, Neuroderm), Amneal, Acadia, Acorda, Abbvie, Supernus (formerly, USWorldMeds), NeuroDiagnostics, Inc

(SYNAPS Dx), Intrance Medical Systems, Inc., Praxis Precision Medicines, and Herantis Pharma; Data Safety Monitoring Board (chair) of AskBio; and publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer. He is co-inventor of the patent “Compositions and methods for treatment and/or prophylaxis of proteinopathies” with which he cofounded REGAIN Therapeutics to fund preclinical studies. He has no financial relationship with the company and relinquished the right to any personal income from future treatments. Dr. Espay owns no stock in any pharmaceutical company with which he has a scientific advisory relationship. **Dr. Cloud** has had clinical trial contracts (paid to her institution) with Cerevel, UCB, Amneal, Blue Rock Therapeutics, Biohaven, and Intra-Cellular Technologies. She has two inventions (Gastrointestinal Symptoms in Neurodegenerative Disease (GIND) scale, VCU Office of Technology Transfer # CLO-11-0R67 and Recognize and Deploy Vibration for Mitigation of Freezing of Gait, VCU Office of Technology Transfer provisional patent #PRE-21-137F (322203-8130), application filed July 2022, full patent application filed July 2023. She is 27% owner of Motion Medix, LLC.

THE PARKINSON STUDY GROUP WATCH-PD STUDY INVESTIGATORS AND COORDINATORS

Jamie L. Adams⁹, Christopher Tarolli⁹, Emma Waddell⁹, Stella Jensen-Roberts⁹, Kelvin Chou¹⁰, Kara J. Wyant¹⁰, Arjun Tarakad¹¹, Alessandro Di Rocco¹², Caroline Tanner¹³, Steven A. Gunzler¹⁴, Camilla Kilbane¹⁴, Cynthia Poon¹⁵, Tanya Simuni¹⁵, Alberto J. Espay¹⁶, Jeanne Feuerstein¹⁷, Michelle Fullard¹⁷, Andrew Feigin¹⁸, Odinachi Oguh¹⁹, Leslie J. Cloud²⁰

⁹University of Rochester, Rochester, NY, USA.

¹⁰University of Michigan, Ann Arbor, MI, USA.

¹¹Baylor College of Medicine, Houston, TX, USA.

¹²Northwell, New Hyde Park, NY, USA.

¹³University of California, San Francisco, CA, USA.

¹⁴University Hospitals Cleveland Medical Center, Cleveland, OH, USA.

¹⁵Northwestern University, Evanston, IL, USA.

¹⁶University of Cincinnati, Cincinnati, OH, USA.

¹⁷University of Colorado Denver, Denver, CO, USA.

¹⁸New York University, New York, NY, USA.

¹⁹Sentara Neurology Specialists, Virginia Beach, VA, USA.

References

1. WHO. *Parkinson disease*, %3C<https://www.who.int/news-room/fact-sheets/detail/parkinson-disease>%3E (2023).
2. Dorsey, E. R. *et al.* Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 17, 939–953 (2018). [https://doi.org/10.1016/S1474-4422\(18\)30295-3](https://doi.org/10.1016/S1474-4422(18)30295-3)
3. Vijiaratnam, N., Simuni, T., Bandmann, O., Morris, H. R. & Foltynie, T. Progress towards therapies for disease modification in Parkinson's disease. *Lancet Neurol* 20, 559–572 (2021). [https://doi.org/10.1016/S1474-4422\(21\)00061-2](https://doi.org/10.1016/S1474-4422(21)00061-2)
4. Regnault, A. *et al.* Does the MDS-UPDRS provide the precision to assess progression in early Parkinson's disease? Learnings from the Parkinson's progression marker initiative cohort. *J Neurol* 266, 1927–1936 (2019). <https://doi.org/10.1007/s00415-019-09348-3>
5. Evers, L. J. W., Krijthe, J. H., Meinders, M. J., Bloem, B. R. & Heskes, T. M. Measuring Parkinson's disease over time: The real-world within-subject reliability of the MDS-UPDRS. *Mov Disord* 34, 1480–1487 (2019). <https://doi.org/10.1002/mds.27790>
6. Kenny, L. *et al.* Inter-rater reliability of hand motor function assessment in Parkinson's disease: Impact of clinician training. *Clin Park Relat Disord* 11, 100278 (2024). <https://doi.org/10.1016/j.prdoa.2024.100278>
7. Czech, M. D. *et al.* Improved measurement of disease progression in people living with early Parkinson's disease using digital health technologies. *Commun Med (Lond)* 4, 49 (2024). <https://doi.org/10.1038/s43856-024-00481-3>
8. Mirelman, A. *et al.* Digital Mobility Measures: A Window into Real-World Severity and Progression of Parkinson's Disease. *Mov Disord* 39, 328–338 (2024). <https://doi.org/10.1002/mds.29689>
9. Adams, J. L. *et al.* Using a smartwatch and smartphone to assess early Parkinson's disease in the WATCH-PD study over 12 months. *NPJ Parkinsons Dis* 10, 112 (2024). <https://doi.org/10.1038/s41531-024-00721-2>
10. Conover, W. J. I., R. L. Rank Transformations as a Bridge between Parametric and Nonparametric Statistics. *The American Statistician* 35, 124–129 (1981). <https://doi.org/10.1080/00031305.1981.10479327>
11. Rusz, J., Krack, P. & Tripoliti, E. From prodromal stages to clinical trials: The promise of digital speech biomarkers in Parkinson's disease. *Neurosci Biobehav Rev* 167, 105922 (2024). <https://doi.org/10.1016/j.neubiorev.2024.105922>
12. Rabano-Suarez, P. *et al.* Digital Outcomes as Biomarkers of Disease Progression in Early Parkinson's Disease: A Systematic Review. *Mov Disord* 40, 184–203 (2025). <https://doi.org/10.1002/mds.30056>

13. Salarian, A. et al. Analyzing 180 degrees turns using an inertial system reveals early signs of progression of Parkinson's disease. *Annu Int Conf IEEE Eng Med Biol Soc* 2009, 224–227 (2009). <https://doi.org/10.1109/IEMBS.2009.5333970>
14. Monje, M. H. G. et al. Motor Onset Topography and Progression in Parkinson's Disease: the Upper Limb Is First. *Mov Disord* 36, 905–915 (2021). <https://doi.org/10.1002/mds.28462>
15. Patera, M. et al. Abnormal arm swing movements in Parkinson's disease: onset, progression and response to L-Dopa. *J Neuroeng Rehabil* 22, 47 (2025). <https://doi.org/10.1186/s12984-025-01589-w>
16. Brzezicki, M. A., Conway, N., Sotirakis, C., FitzGerald, J. J. & Antoniadou, C. A. Antiparkinsonian medication masks motor signal progression in de novo patients. *Heliyon* 9, e16415 (2023). <https://doi.org/10.1016/j.heliyon.2023.e16415>
17. Kirk, C. et al. Can Digital Mobility Assessment Enhance the Clinical Assessment of Disease Severity in Parkinson's Disease? *J Parkinsons Dis* 13, 999–1009 (2023). <https://doi.org/10.3233/JPD-230044>
18. Post, E. et al. Quantifying arm swing in Parkinson's disease: a method accounting for arm activities during free-living gait. *J Neuroeng Rehabil* 22, 37 (2025). <https://doi.org/10.1186/s12984-025-01578-z>
19. Cao, F., Vogel, A. P., Gharahkhani, P. & Renteria, M. E. Speech and language biomarkers for Parkinson's disease prediction, early diagnosis and progression. *NPJ Parkinsons Dis* 11, 57 (2025). <https://doi.org/10.1038/s41531-025-00913-4>
20. Yin, W. et al. Gait analysis in the early stage of Parkinson's disease with a machine learning approach. *Front Neurol* 15, 1472956 (2024). <https://doi.org/10.3389/fneur.2024.1472956>
21. Rehman, R. Z. U. et al. Selecting Clinically Relevant Gait Characteristics for Classification of Early Parkinson's Disease: A Comprehensive Machine Learning Approach. *Sci Rep* 9, 17269 (2019). <https://doi.org/10.1038/s41598-019-53656-7>
22. Lipsmeier, F. et al. Reliability and validity of the Roche PD Mobile Application for remote monitoring of early Parkinson's disease. *Sci Rep* 12, 12081 (2022). <https://doi.org/10.1038/s41598-022-15874-4>
23. Giboin, L. S. et al. A digital motor score for sensitive detection of progression in Huntington's disease. *Brain* 148, 3228–3238 (2025). <https://doi.org/10.1093/brain/awaf127>
24. Timmermans, N. A. et al. Daily-Life, Sensor-Derived Tremor Measures Are Sensitive to Progression in Early Parkinson's Disease. *Ann Neurol* 100, 282–294 (2026). <https://doi.org/10.1002/ana.78236>
25. Mammen, J. R. et al. Three years later: tracking bothersome symptoms and impacts for people with early Parkinson's disease. *J Neurol* 273, 93 (2026). <https://doi.org/10.1007/s00415-026-13615-5>
26. Fahn, S. et al. Levodopa and the progression of Parkinson's disease. *N Engl J Med* 351, 2498–2508 (2004). <https://doi.org/10.1056/NEJMoa033447>
27. Wilkins, K. B. et al. Quantitative Digitography Measures Motor Symptoms and Disease Progression in Parkinson's Disease. *J Parkinsons Dis* 12, 1979–1990 (2022). <https://doi.org/10.3233/JPD-223264>

28. Sahandi Far, M., Eickhoff, S. B., Goni, M. & Dukart, J. Exploring Test-Retest Reliability and Longitudinal Stability of Digital Biomarkers for Parkinson Disease in the m-Power Data Set: Cohort Study. *J Med Internet Res* 23, e26608 (2021). <https://doi.org/10.2196/26608>
29. Janssen Daalen, J. M. et al. Digital biomarkers for non-motor symptoms in Parkinson's disease: the state of the art. *NPJ Digit Med* 7, 186 (2024). <https://doi.org/10.1038/s41746-024-01144-2>
30. Yarnall, A. J. et al. The future of clinical trials? Monitoring mobility as an outcome measure in Parkinson's disease: The mobilise-D study. *J Parkinsons Dis* 16, 110–124 (2026). <https://doi.org/10.1177/1877718X251407026>
31. Di Lazzaro, G. et al. Technology-based therapy-response and prognostic biomarkers in a prospective study of a de novo Parkinson's disease cohort. *NPJ Parkinsons Dis* 7, 82 (2021). <https://doi.org/10.1038/s41531-021-00227-1>
32. Page, A. et al. A Smartphone Application as an Exploratory Endpoint in a Phase 3 Parkinson's Disease Clinical Trial: A Pilot Study. *Digit Biomark* 6, 1–8 (2022). <https://doi.org/10.1159/000521232>
33. Pagano, G. et al. Trial of Prasinezumab in Early-Stage Parkinson's Disease. *N Engl J Med* 387, 421–432 (2022). <https://doi.org/10.1056/NEJMoa2202867>
34. Schneider, R. B. et al. Design of a virtual longitudinal observational study in Parkinson's disease (AT-HOME PD). *Ann Clin Transl Neurol* 8, 308–320 (2021). <https://doi.org/10.1002/acn3.51236>
35. Prochnow, T. et al. Understanding Longitudinal Ecological Momentary Assessment Completion: Results From 12 Months of Burst Sampling in the TIME Study. *JMIR Mhealth Uhealth* 13, e67117 (2025). <https://doi.org/10.2196/67117>
36. Lavine, J. S. et al. Impacts on study design when implementing digital measures in Parkinson's disease-modifying therapy trials. *Front Digit Health* 6, 1430994 (2024). <https://doi.org/10.3389/fdgth.2024.1430994>
37. Bot, B. M. et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci Data* 3, 160011 (2016). <https://doi.org/10.1038/sdata.2016.11>
38. FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making, <<https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical%3E> (2025).
39. Roussos, G. et al. Identifying and characterising sources of variability in digital outcome measures in Parkinson's disease. *NPJ Digit Med* 5, 93 (2022). <https://doi.org/10.1038/s41746-022-00643-4>
40. Mancini, M. et al. A framework to standardize gait study protocols in Parkinson's disease. *J Parkinsons Dis* 15, 129–139 (2025). <https://doi.org/10.1177/1877718X241305626>
41. Adams, J. L. et al. Using a smartwatch and smartphone to assess early Parkinson's disease in the WATCH-PD study. *NPJ Parkinsons Dis* 9, 64 (2023). <https://doi.org/10.1038/s41531-023-00497-x>
42. Czech, M. D. P., S. GaitPy: An Open-Source Python Package for Gait Analysis Using an Accelerometer on the Lower Back. *Journal of Open Source Software* 4 (2019). <https://doi.org/10.21105/joss.01778>

43. Warmerdam, E. et al. Long-term unsupervised mobility assessment in movement disorders. *Lancet Neurol* 19, 462–470 (2020). [https://doi.org/10.1016/S1474-4422\(19\)30397-7](https://doi.org/10.1016/S1474-4422(19)30397-7)
44. Getting movement disorder symptom data. *Apple Developer Documentation* 2025. %3C<https://developer.apple.com/documentation/coremotion/getting-movement-disorder-symptom-data>%3E.
45. Powers, R. et al. Smartwatch inertial sensors continuously monitor real-world motor fluctuations in Parkinson's disease. *Sci Transl Med* 13 (2021). <https://doi.org/10.1126/scitranslmed.abd7865>
46. Simuni, T. et al. Longitudinal change of clinical and biological measures in early Parkinson's disease: Parkinson's Progression Markers Initiative cohort. *Mov Disord* 33, 771–782 (2018). <https://doi.org/10.1002/mds.27361>
47. Straczekiewicz, M., Huang, E. J. & Onnela, J. P. A "one-size-fits-most" walking recognition method for smartphones, smartwatches, and wearable accelerometers. *NPJ Digit Med* 6, 29 (2023). <https://doi.org/10.1038/s41746-022-00745-z>
48. Onnela, J.-P. et al. Beiwe: A data collection platform for high-throughput digital phenotyping. *Journal of Open Source Software* 6 <https://doi.org/10.21105/joss.03417>
49. Little, M. A. Smartphones for Remote Symptom Monitoring of Parkinson's Disease. *J Parkinsons Dis* 11, S49-S53 (2021). <https://doi.org/10.3233/JPD-202453>

Tables

Tables 1 and 2 are available in the supplementary files section.

Figures

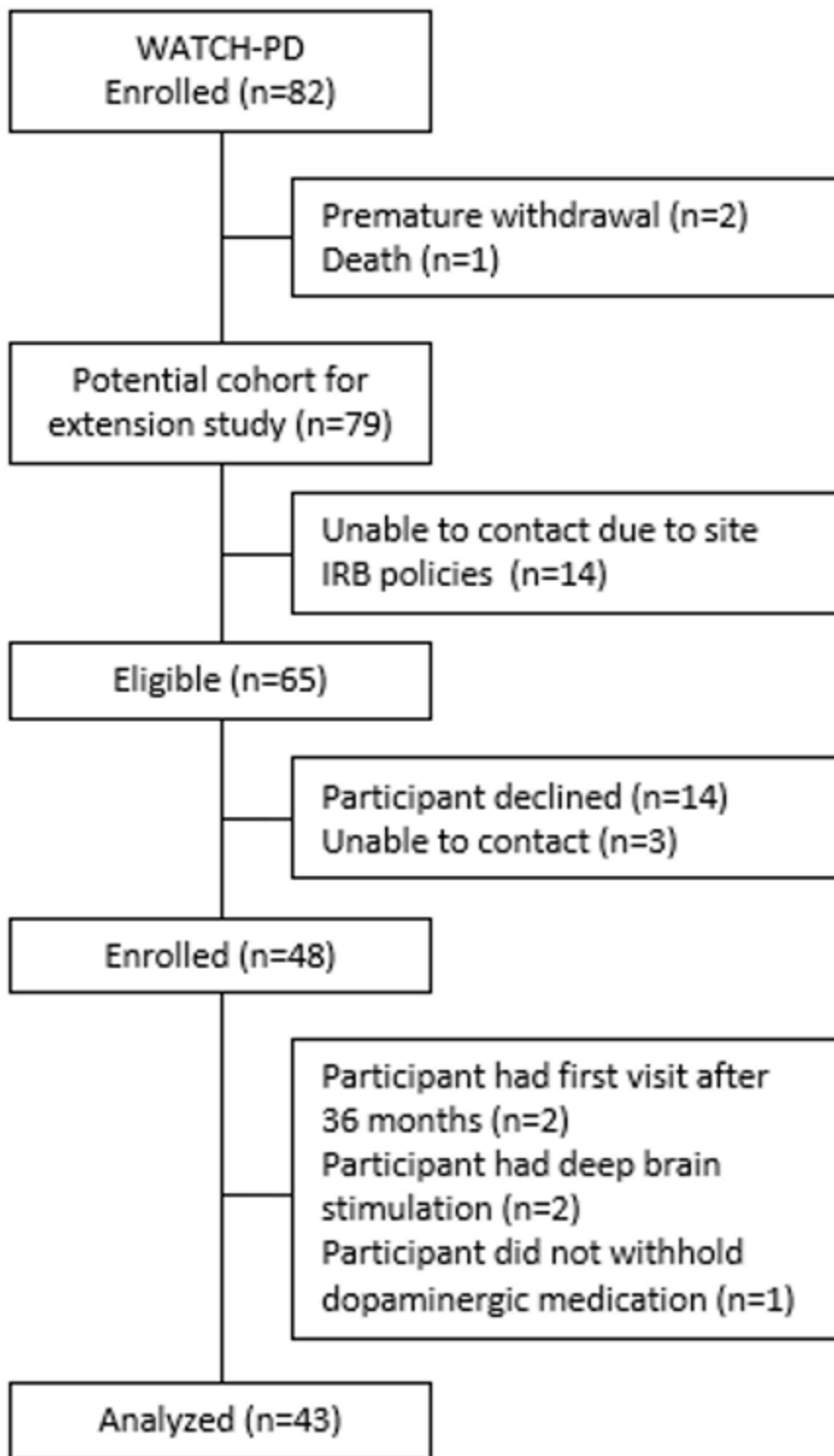


Figure 1

Flow of participants in the WATCH-PD study over 36 months

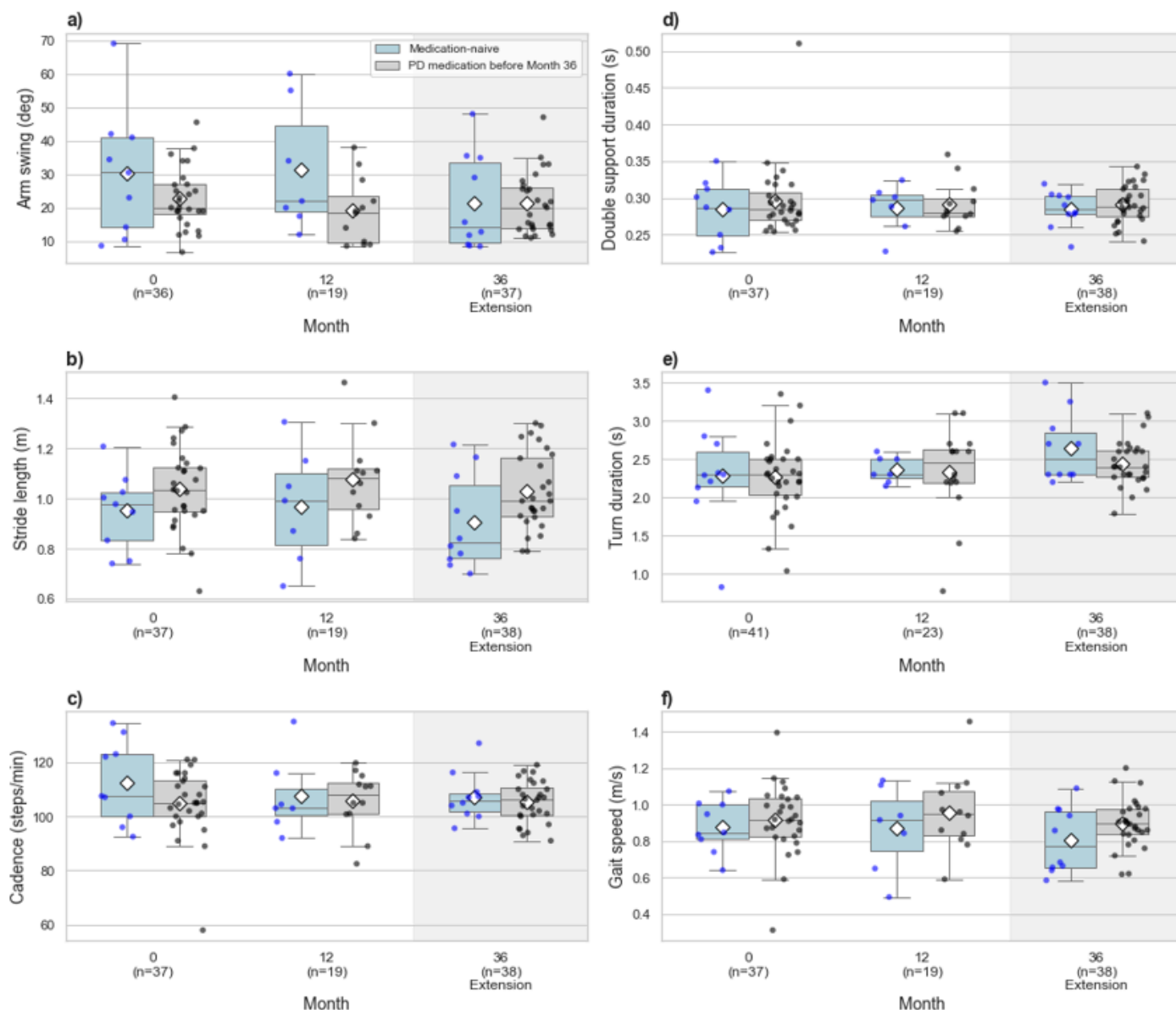


Figure 2

Comparisons of gait features derived from smartphone and smartwatch data during active tasks. Box plots show (a) arm swing (deg), (b) stride length (m), (c) cadence (steps/min), (d) double support duration (s), (e) turn duration (s), and (f) gait speed (m/s) at month 0, month 12, and month 36. Participants are grouped by medication status: *medication-naïve* (never received PD medication throughout the study) in blue and *participants on PD medication before month 36*; data included up to the point before starting medication, with month 36 assessments performed off medication for these participants) in grey. Dots show individual subject-level data, and diamonds indicate mean value.

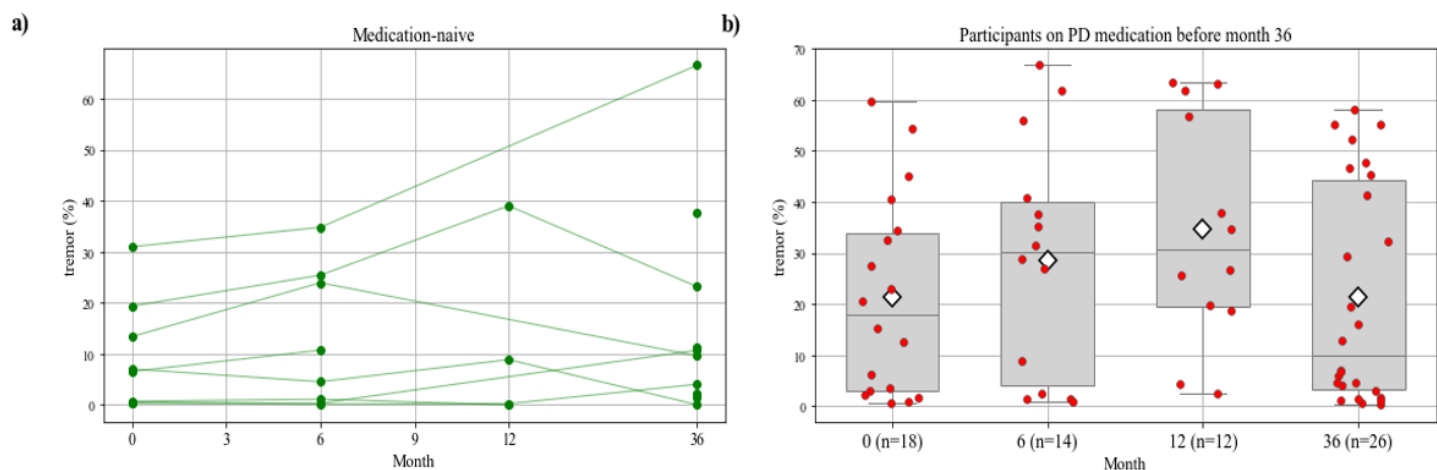


Figure 3

Percentage of time with tremor derived from the smartwatch. (a) Individual trajectories of medication-naïve participants. Each green line represents a subject's measurements over time; dots mark their individual values at each month. (b): *Participants on PD medication before month 36*. At month 0, these subjects are drug-naïve; by month 36, they are on medication. Box plots show distribution, and black dots indicate individual subject measurements. Diamonds indicate mean.

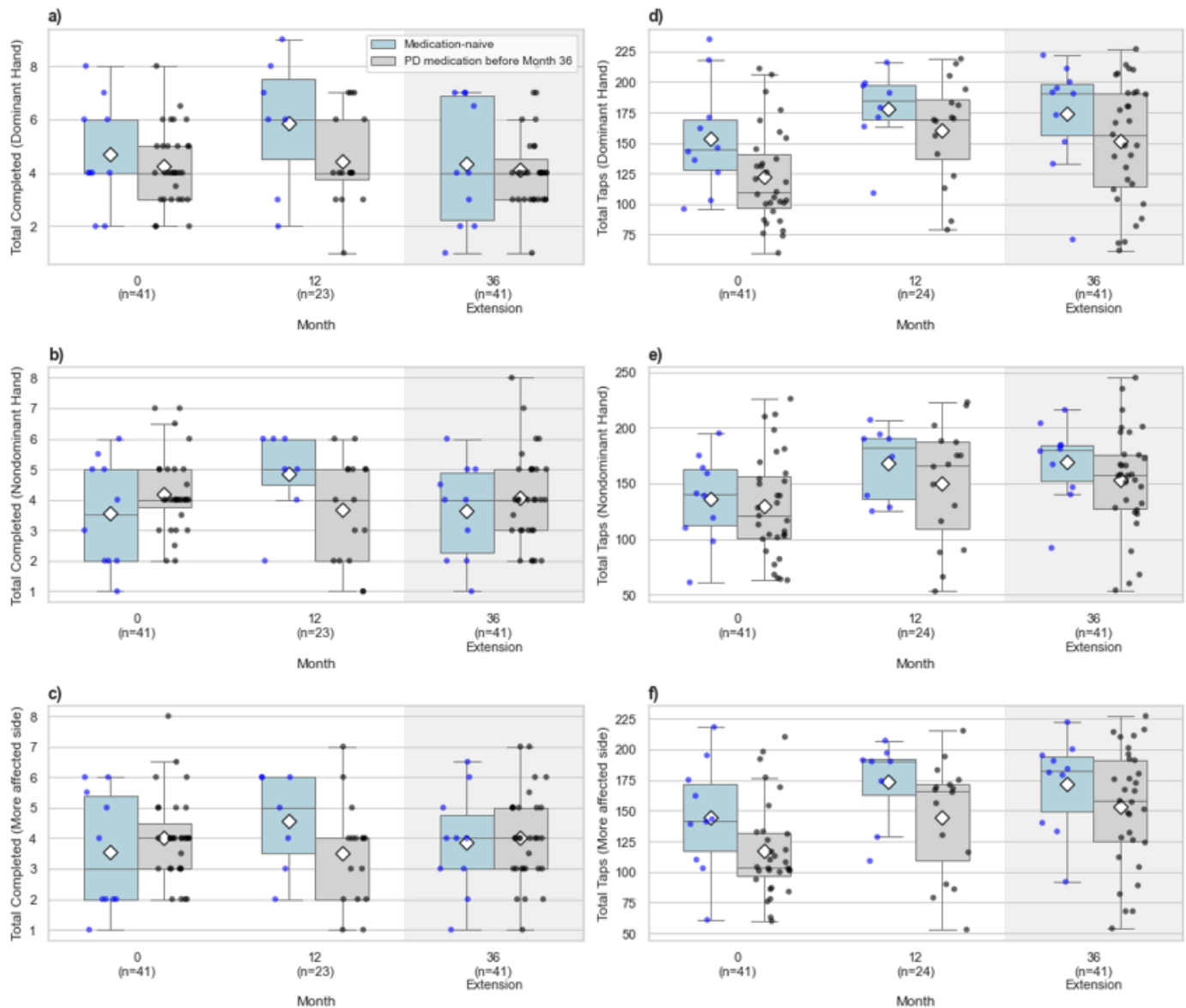


Figure 4

Psychomotor function derived from smartphone assessments. Box plots show (a) number of objects successfully manipulated with the dominant hand, (b) number of objects successfully manipulated with the nondominant hand, (c) number of objects successfully manipulated with the most affected hand, (d) total finger taps with the dominant hand, (e) total finger taps with the nondominant hand, (f) total finger taps with the most affected hand at month 0, month 12, and month 36. Specifically, month 0 and month 12 represent longitudinal tracking, while month 36 corresponds to the extension study. Participants are grouped by *medication-naïve* (never received PD medication throughout the study) in blue and *participants on PD medication before month 36*; data included up to the point before starting medication, with month 36 assessments performed off medication for these participants) in grey. Dots show individual subject-level data. Diamonds indicate mean value.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tables.docx](#)