




OPEN

Reliability and validity of the Roche PD Mobile Application for remote monitoring of early Parkinson's disease

Florian Lipsmeier^{1,7}, Kirsten I. Taylor^{1,7}, Ronald B. Postuma², Ekaterina Volkova-Volkmar¹, Timothy Kilchenmann¹, Brit Mollenhauer^{3,4}, Atieh Bamdadian¹, Werner L. Popp¹, Wei-Yi Cheng¹, Yan-Ping Zhang¹, Detlef Wolf¹, Jens Schjodt-Eriksen¹, Anne Boulay⁵, Hanno Svoboda¹, Wagner Zago⁶, Gennaro Pagano¹ & Michael Lindemann¹

Digital health technologies enable remote and therefore frequent measurement of motor signs, potentially providing reliable and valid estimates of motor sign severity and progression in Parkinson's disease (PD). The Roche PD Mobile Application v2 was developed to measure bradykinesia, bradyphrenia and speech, tremor, gait and balance. It comprises 10 smartphone active tests (with ½ tests administered daily), as well as daily passive monitoring via a smartphone and smartwatch. It was studied in 316 early-stage PD participants who performed daily active tests at home then carried a smartphone and wore a smartwatch throughout the day for passive monitoring (study NCT03100149). Here, we report baseline data. Adherence was excellent (96.29%). All pre-specified sensor features exhibited good-to-excellent test–retest reliability (median intraclass correlation coefficient = 0.9), and correlated with corresponding Movement Disorder Society–Unified Parkinson's Disease Rating Scale items (ρ : 0.12–0.71). These findings demonstrate the preliminary reliability and validity of remote at-home quantification of motor sign severity with the Roche PD Mobile Application v2 in individuals with early PD.

Clinical and drug development research in Parkinson's disease (PD) critically requires the precise quantification of clinical disease severity and its progression over time. This is a challenge in PD because of the fluctuating and slowly progressive nature of the disease. Moreover, the development of disease-modifying therapies focus on the earliest possible point in the course of the disease, when the least amount of neurodegeneration has taken place, and when disease severity and progression may be even more subtle¹. Digital health technologies (DHTs)², such as smartphones and smartwatches, may aid in overcoming these challenges, since they enable remote and therefore frequent measurement of motor signs to provide potentially more robust—i.e. reliable and valid—quantification of disease severity and its changes over time^{3–5}. Moreover, the inertial measurement unit sensors (e.g. accelerometers, gyroscopes) are highly sensitive to minute changes (even in consumer technologies) and therefore may detect motor changes not evident upon routine clinical examination. While not less important, non-motor signs remain largely inaccessible to sensor measurements via DHTs. Here, we describe a novel DHT, the Roche PD Mobile Application v2, which was designed to measure motor manifestations in early PD, and demonstrate its test–retest reliability and validity in a group of de novo diagnosed individuals with PD.

Clinical trials of future potentially disease-modifying therapies for PD⁶ are especially promising among individuals who have been recently diagnosed, when less neurodegeneration has occurred. Clearly, to detect potential treatment benefit, sensitive measures of changes in motor signs in the earliest stages of the disease are required. In the Parkinson's Progression Markers Initiative (PPMI) cohort⁷, Rasch Measurement Theory (RMT)

¹Roche Pharma Research and Early Development, pRED Informatics, Pharmaceutical Sciences, Clinical Pharmacology, and Neuroscience and Rare Diseases Discovery and Translational Area, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Grenzacherstrasse 124, 4070 Basel, Switzerland. ²Department of Neurology, McGill University, Montreal General Hospital, Montreal, QC, Canada. ³Paracelsus-Elena-Klinik, Kassel, Germany. ⁴Department of Neurology, University Medical Center Göttingen, Göttingen, Germany. ⁵Idorsia Pharmaceuticals Ltd, Allschwil, Switzerland. ⁶Prothena Biosciences Inc, South San Francisco, CA, USA. ⁷These authors contributed equally: Florian Lipsmeier and Kirsten I. Taylor. ✉email: florian.lipsmeier@roche.com

analyses of the Movement Disorder Society–Unified Parkinson’s Disease Rating Scale (MDS-UPDRS)⁸ Part II and III scores at screening, 12-month and 24-month visits (n = 384) were applied to explore potential biases in the ordinal MDS-UPDRS items comprising the subscale scores⁹. These revealed an apparent staged order of motor sign progression from unilateral bradykinesia and rigidity (first upper then lower extremities) to midline functions to bilateral bradykinesia and rigidity, and finally general movement problems¹⁰. These findings confirm the centrality of bradykinesia in early PD¹¹ and suggest that bradykinesia is a critical motor progression marker in early PD¹. However, the quantification of bradykinesia and other motor signs in early PD with rating scales such as the MDS-UPDRS remains a challenge: MDS-UPDRS Part II scores change little in early PD¹², i.e. less than the established minimal clinically meaningful difference¹³, and RMT analyses of MDS-UPDRS Part III item scores revealed multiple measurement irregularities in scores during the first 2 years of PPMI, including floor effects, item gaps for very mild motor signs and item misfits (i.e. items that did not fit the ordered progression pattern)^{10,14}. These findings highlight the urgent need for alternative methods of motor sign quantification in the earliest stages of PD.

Many smartphone- and smartwatch-based DHTs have been developed to estimate bradykinesia and other motor signs of PD in research settings, including the present DHT solution^{5,15–21}. The individual tasks and measurements are comparable, as described in the Discussion section. With respect to early PD, finger tapping is one of the most commonly used DHT measures of bradykinesia²¹. When sensor-based finger-tapping data are aggregated over 2-week periods, test–retest reliabilities increase⁵ and correlate with MDS-UPDRS Part III clinician ratings of finger-tapping performance in patients with early PD⁵. Additional bradykinesia tests implemented on smartphones include measures of hand turning and leg agility (by holding the phone on the thigh and lifting and stomping the foot), for example as implemented on CloudUPDRS^{16,17}. Smartwatches offer additional means to estimate bradykinesia during daily life, for example, by estimating the time taken to move a utensil from a plate to the mouth while eating¹⁵. Most DHT solutions for PD such as CloudUPDRS^{16,17}, HopkinsPD¹⁸, mPower¹⁹ and the Roche PD Mobile Application v1⁵ test not only bradykinesia, but also tremor, gait, and balance, thereby providing a profile of motor impairments for estimating and tracking PD motor severity. DHTs may additionally benefit from measures of cognition such as information processing speed (e.g. electronic Symbol Digit Modalities Test [eSDMT]) and speech, which are also affected in the earliest stages of PD²⁰.

The present report describes the reliability and validity of the Roche PD Mobile Application v2, a revision of v1 to include multiple novel measures of bradykinesia and cognition (eSDMT and speech tests), and to optimize existing tasks for the detection of PD motor signs. Three hundred and sixteen individuals with early-stage PD (<2 years; Hoehn and Yahr Stages I or II) who are participating in the Phase II PASADENA study (NCT03100149) were provided with a smartphone and smartwatch with the Roche PD Mobile Application v2 preinstalled, and requested to perform active tests daily on the smartphone (4 or 5 out of 10 tests each day, information processing speed once per fortnight), and to carry the smartphone and wear a smartwatch throughout the day to collect passive monitoring data. Pre-specified sensor features were calculated for each active test and for passive monitoring, and aggregated over the first two 2-week periods of the study. Adherence, test–retest reliability, and clinical validity (relationship to baseline clinical scales, known-groups validity) were quantified. These metrics represent the grounds for judging the potential utility of the Roche PD Mobile Application v2 to quantify and track progression of disease severity in early PD.

Results

Adherence. On average, daily remote active testing took a median of 5.3 (interquartile range [IQR] = 1.7) minutes on days without the SDMT, and 7.32 (median; IQR = 1.58) minutes on days with SDMT. Average adherence was high with 96.29% (median per participant) of all possible active tests performed during the first 4 weeks of the study (i.e. 26–27/27 days). Participants contributed a median of 8.6 h/day of study smartphone and a median of 12.79 h/day of study smartwatch passive monitoring data.











Reliability of sensor features. Reliability results are reported in Table 1. All 17 pre-specified sensor features demonstrated good-to-excellent²² test–retest reliability between the first two 2-week study periods (intra-class correlation coefficient [ICCs] ≥ 0.75) (Table 1)²³. The median sensor feature ICC was 0.9 (range, 0.75–0.95).

Clinical validity of sensor features. Clinical validity was assessed via Spearman’s correlations between the sensor features and corresponding MDS-UPDRS subscale and item scores (Fig. 1). Correlations with MDS-UPDRS item scores revealed that all sensor features correlated with their corresponding clinical items (Table 1 and Figs. 3, 4, 5). The numerically strongest relationships were observed with bradykinesia sensor features (i.e. Hand Turning and Finger Tapping) as well as postural and rest tremor sensor features, particularly for the more affected side, while the numerically weakest correlations were found with the Balance (comparator: MDS-UPDRS item 3.12 Postural stability) and Draw A Shape (comparator: MDS-UPDRS item 3.5 Hand movement) tests.

Cross-correlations between sensor features and MDS-UPDRS subscale scores demonstrated convergent and divergent validity of bradykinesia and tremor sensor features, with the strongest relationships between bradykinesia sensor features and MDS-UPDRS subscores, and between tremor sensor features and MDS-UPDRS subscores, compared with all other subscores (Fig. 2). The Postural Instability/Gait Disorders (PIGD) subscale score showed the strongest correlations with the U-turn and the Hand Turning test (most affected side). Rigidity subscale scores correlated weakly with all sensor features overall, showing a weak-to-moderate relationship with Hand Turning in the most affected side and U-turn speed. The novel ‘bulbar score’ (defined in the methods section below) correlated most strongly with the Speech Test sensor feature. Overall, MDS-UPDRS total scores were most strongly related to bradykinesia sensor features, as expected for this early-stage population¹.

Digital test	Sensor feature	MDS-UPDRS item	Test-retest reliability ICC (95% CI)	Spearman's correlation coefficient		MDS-UPDRS item score 0 versus 1					Hoehn and Yahr Stage I versus II				
				r_s	<i>P</i> value	N	U stat	<i>P</i> value	N 0/N 1	Percent difference on group medians	U stat	<i>P</i> value	N 1/N 2	Percent difference on group medians	Feature differences in expected direction
Draw A Shape	Spiral celerity (L)	3.5 Hand movement (L)	0.91 (0.88–0.93)	–0.15	0.0092	300	7061	0.0043	184/95	–14.06	6681	0.0047	74/226	–9.5	Yes
	Spiral celerity (M)	3.5 Hand movement (M)	0.92 (0.89–0.94)	–0.12	0.0334	300	3091	0.0981	49/144	–14.29	6463	0.0017	74/226	–15.2	Yes
Dexterity	Tapping variability (L)	3.4 Finger tapping (L)	0.84 (0.79–0.88)	0.32	<0.0001	300	6786	0.0006	152/116	27.33	6444	0.0015	74/226	27.16	Yes
	Tapping variability (M)	3.4 Finger tapping (M)	0.88 (0.84–0.91)	0.37	<0.0001	300	1261	0.0006	30/135	40.64	6691	0.005	74/226	17.95	Yes
Hand Turning	Median Hand Turning speed (L)	3.6 Pronation-supination (L)	0.92 (0.89–0.94)	–0.18	0.0015	300	7930	0.0361	192/95	–7.85	7268	0.0458	74/226	–7.18	Yes
	Median Hand Turning speed (M)	3.6 Pronation-supination (M)	0.95 (0.93–0.97)	–0.46	<0.0001	300	1912	0.0008	43/131	–12.37	7216	0.0385	74/226	–9.33	Yes
Speech	MFCC2 variability	3.1 Speech	0.77 (0.69–0.83)	–0.29	<0.0001	295	5692	<0.0001	126/138	–13.68	6395	0.0412	71/209	–4.47	Yes
Phonation	Voice jitter	3.1 Speech	0.88 (0.85–0.91)	0.18	0.0017	297	7381	0.0001	135/146	21.65	7248	0.073	73/224	12.87	Yes
Postural Tremor	Log median squared energy (L)	3.15 Postural tremor (L)	0.83 (0.77–0.87)	0.23	<0.0001	300	3240	0.0002	259/39	8.26	7111	0.0343	73/227	1.2	Yes
	Log median squared energy (M)	3.1 Postural tremor (M)	0.92 (0.90–0.94)	0.5	<0.0001	300	4779	<0.0001	150/116	12.45	8108	0.3918	73/227	1.28	Yes
Rest Tremor	Log median squared energy (L)	3.17 Rest tremor amplitude (L)	0.91 (0.88–0.93)	0.34	<0.0001	300	967	<0.0001	274/19	23.73	6614	0.0048	73/227	6.12	Yes
	Log median squared energy (M)	3.17 Rest tremor amplitude (M)	0.94 (0.91–0.95)	0.71	<0.0001	300	1565	<0.0001	124/71	19.27	7896	0.2731	73/227	3.59	Yes
Balance	Log sway jerk	3.12 Postural stability	0.82 (0.76–0.87)	0.14	0.0138	299	890	0.0071	287/11	10.6	7988	0.3870	72/227	1.29	Yes
U-turn	Median turn speed	3.14 Body bradykinesia	0.91 (0.88–0.93)	–0.22	0.0001	299	3113	0.0539	44/168	–5.39	6322	0.0019	72/227	–7.83	Yes
eSDMT	Number of correct responses	1.1 Cognitive impairment	0.75 (0.66–0.81)	–0.18	0.0014	303	3787.5	0.0012	248/43	–10	6538	0.0008	76/227	–7.14	Yes
Passive monitoring of gait	Median turn speed	3.14 Body bradykinesia	0.82 (0.77–0.89)	–0.22	0.0001	285	3130	0.154	43/162	–3.3	6270	0.0138	71/214	–4.65	Yes
Passive monitoring of non-gait arm movements	Median arm movement power (non-gait)	3.6 Pronation-supination, left hand	0.85 (0.80–0.89)	–0.43	<0.0001	269	3108	<0.0001	106/89	–35.73	5221	0.0242	60/209	–30.64	Yes

Table 1. Reliability and clinical validity of pre-specified Roche PD Mobile Application v2 active test and passive monitoring sensor features. “M” = most affected side; “L” = least affected side. *CI* confidence interval, *eSDMT* electronic Symbol Digit Modalities Test, *ICC* intraclass correlation coefficient, *L* less affected side, *M* more affected side, *MDS-UPDRS* Movement Disorder Society—Unified Parkinson’s Disease Rating Scale, *MFCC2* Mel Frequency Cepstral Coefficient 2.

ACTIVE TESTS									
Bradykinesia			Tremor/Bradykinesia	Tremor			Rigidity/Postural Instability		Cognition
Draw A Shape	Dexterity	Hand Turning	Speech	Phonation	Postural Tremor	Rest Tremor	Balance	U-Turn	Cognitive Test (eSDMT)
									
Bradykinesia Days (Every 2nd Day)			Alternating		Tremor and Stability Days (Every 2nd Day)			Fortnightly	




PASSIVE MONITORING		
Bradykinesia and Activities of Daily Living		
Gait	Arm Swing & Tremor	Mobility & Sociability
		
Daily	Daily	Daily

Figure 1. Roche PD Mobile Application v2 active tests and passive monitoring and schedule of assessments. eSDMT, electronic Symbol Digit Modalities Test; PD, Parkinson's disease.

Sensor feature sensitivity to early disease manifestations and Hoehn and Yahr stage. A total of 15/17 sensor features discriminated participants with scores of 0 versus 1 on associated MDS-UPDRS items (Table 1 and Figs. 3, 4, 5). Only sensor features from the Draw a Shape and U-turn tests showed borderline non-significant results. Participants in Hoehn and Yahr Stage I versus II were differentiated by 13 sensor features, i.e. all but Phonation test, Postural and Rest Tremor (most affected side) and the Balance test sensor features (Table 1).

Sensor feature sensitivity to side differences. Sensor features values from all lateralized tests demonstrated significant differences between the most and least affected sides (Supplementary Table 1). Moreover, sensor features and MDS-UPDRS scores measuring the same motor sign on the less affected (or more affected) side were more strongly correlated than sensor features and MDS-UPDRS scores measuring the same motor sign on opposite sides of the body (Fig. 6).

Discussion

Reliability and validity of the Roche PD Mobile Application v2. The Roche PD Mobile Application v1 was designed to measure the core motor signs of PD^{5,18}, and was recently revised to v2 to primarily include two new active tests of bradykinesia (Hand Turning, Draw A Shape), as well as a test of psychomotor slowing (eSDMT) and a speech test. In addition, the original gait task was revised to a U-turn test, and a smartwatch was incorporated into the remote passive monitoring procedure. Preliminary test-retest reliability scores for the pre-specified sensor features from all active tests except Speech and eSDMT, and for both passive monitoring measures, were in the 'excellent' range²². Preliminary clinical validity was established via correlations with corresponding MDS-UPDRS item scores. We note that these findings are reassuring considering the continuous (sensor feature) versus ordinal (MDS-UPDRS) nature of the two datasets, and the lack of conceptually comparable MDS-UPDRS items for some active test features (e.g. Draw A Shape). Cross-correlations between sensor features and MDS-UPDRS subscale scores supported the convergent and divergent validity of bradykinesia and tremor sensor features. Most active test sensor features demonstrated sensitivity for subtle manifestations, discriminating individuals who received MDS-UPDRS item scores of 0 from those with item scores of 1. Measures of upper limb bradykinesia demonstrated known-groups validity, differentiating individuals in Hoehn and Yahr Stage I versus II. All lateralized sensor features discriminated least versus the most affected sides of the body. The results from shared active tests and passive sensor features confirm previous findings with the Roche PD Mobile Application v1⁵. Taken together, these results indicate that the Roche PD Mobile Application v2 may prove suitable for quantifying motor disease severity and tracking disease progression in the earlier stages of PD.

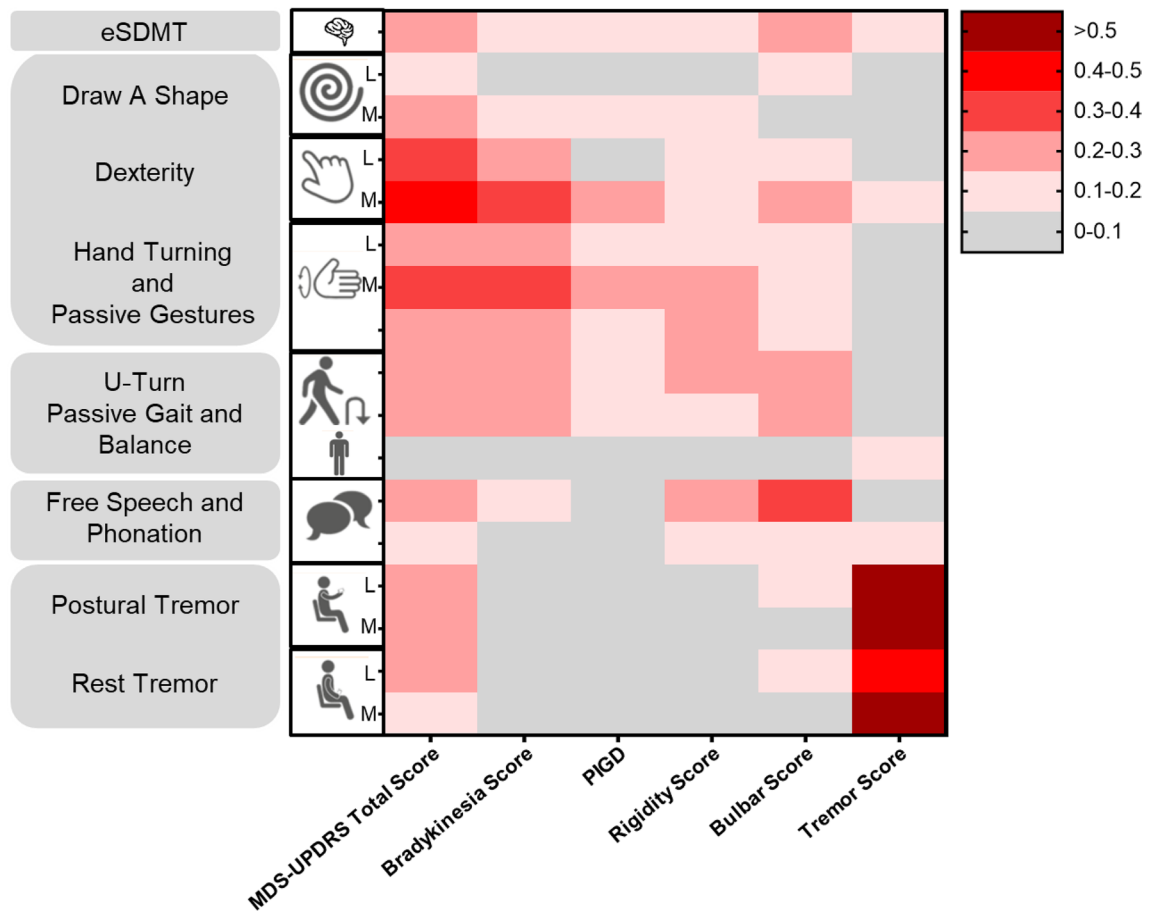


Figure 2. Absolute Spearman's correlations between baseline MDS-UPDRS Total and Subscores and Roche PD Mobile Application v2 active test and passive monitoring sensor features. *eSDMT* electronic Symbol Digit Modalities Test, *MDS-UPDRS* Movement Disorder Society—Unified Parkinson's Disease Rating Scale, *PD* Parkinson's disease, *PIGD* Postural Instability/Gait Disorders.

DHT measurement of bradykinesia. The Roche PD Mobile Application v2 contains three active tests designed to measure upper limb bradykinesia: Dexterity (finger tapping), Hand Turning (pronation/supination), and Draw A Shape. Pre-specified sensor features from all three tests correlated with their corresponding MDS-UPDRS upper limb bradykinesia item scores, and showed convergent and divergent validity in cross-correlations with MDS-UPDRS Part III subscale scores, correlating numerically most strongly with bradykinesia compared with all other subscale scores. These findings indicate that the Roche PD Mobile Application v2 bradykinesia tests indeed reflect the neurological concept of upper limb bradykinesia. Finger tapping and pronation/supination tasks are well-established assessments of upper limb bradykinesia as evidenced by their inclusion in both the UPDRS²⁴ and MDS-UPDRS⁸. Over the last decade, different digitized variants of finger tapping and pronation/supination tests have been developed²¹. Despite methodological differences, studies of these DHT tasks generally showed good correspondence between finger tapping sensor features and respective clinical ratings, as well as the ability to differentiate healthy controls from individuals with early PD, and individuals with early PD from individuals with later-stage PD^{5,18,25–28}, in line with the present findings. While the literature on digitized pronation/supination assessments is less rich than for finger tapping, available results also consistently demonstrate correlations with related clinical scores and the ability to differentiate healthy participants from individuals with PD^{16,23,29–31}. Spiral drawing is traditionally used in behavioral neurology to assess fine motor impairment including bradykinesia and tremor^{32–35}. DHT versions of spiral drawing demonstrated that time to completion correlated with clinician ratings of bradykinesia severity, and differentiated PD cases from controls³⁴. The majority of previous DHT spiral drawing tasks used pens/digital pens to draw on regular paper or tablets, a more challenging motor task compared with the present finger drawing on smaller smartphone touch screens. In the present study, celerity, i.e. accuracy/time to complete spiral shape tracing on the smartphone screen, was pre-specified to additionally consider the accuracy of directed fine motor movements in the unsupervised at-home setting. Spiral celerity correlated with MDS-UPDRS bradykinesia measures, and the strength of these correlations was numerically smaller compared with Finger Tapping and Hand Turning. This may be due to the relative difficulty of the latter two tasks compared with spiral drawing, which may have challenged individuals more, thereby revealing greater impairment. We note that additional sensor features (e.g. variability in drawing

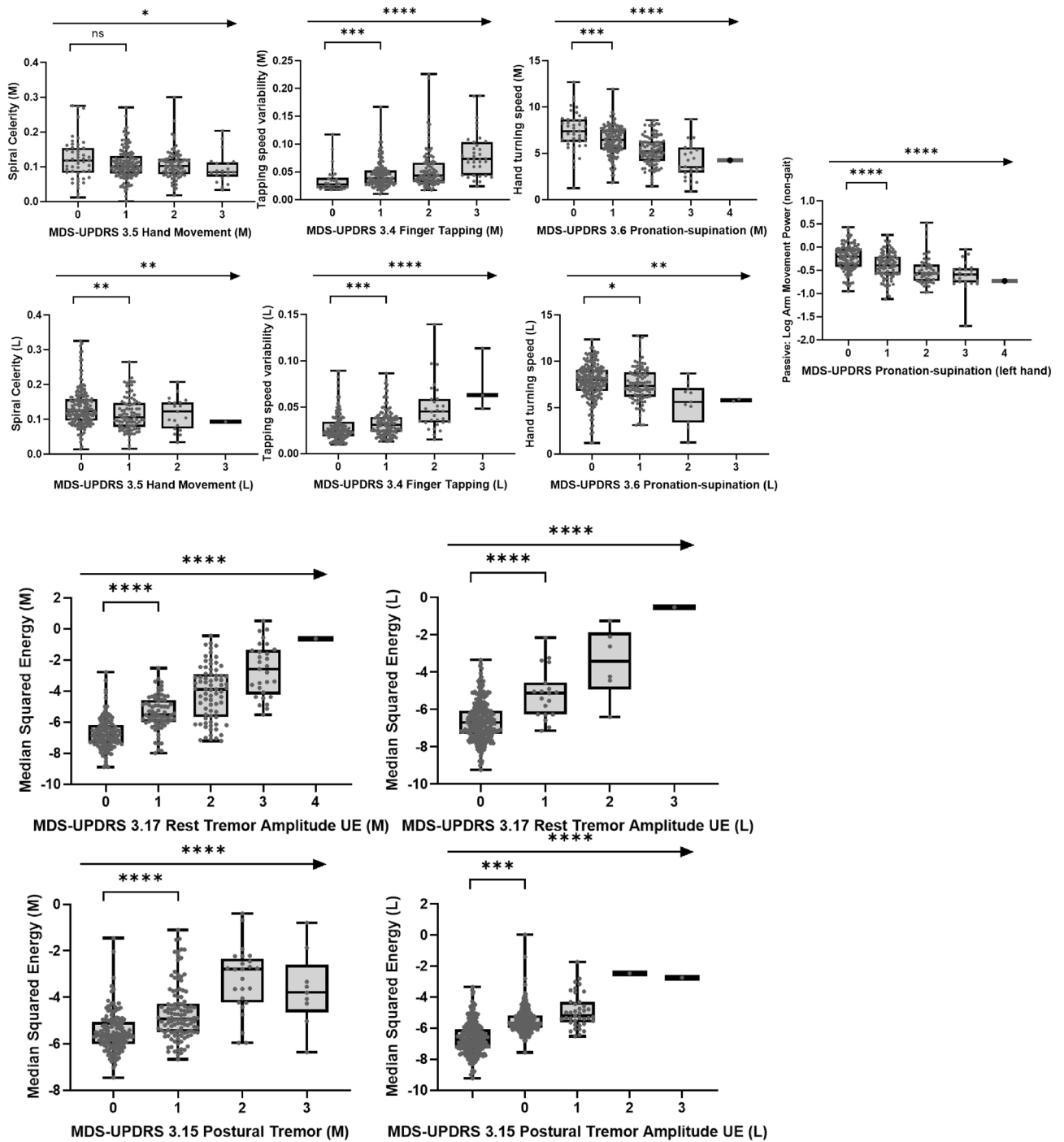


Figure 3. Association of sensor features from upper limb bradykinesia tests and upper limb tremor tests with corresponding clinical MDS-UPDRS measures at baseline (ns = $P > 0.05$; * = $P \leq 0.05$; ** = $P \leq 0.01$; *** = $P \leq 0.001$; **** = $P \leq 0.0001$). *L* less affected side, *M* more affected side, *MDS-UPDRS* Movement Disorder Society—Unified Parkinson’s Disease Rating Scale, *UE* Upper Extremity.

speed, hesitation), analyzed either individually or combined within and across shapes, are expected to provide additional meaningful information, as has been shown for PD and multiple sclerosis^{36,37}.

Passive monitoring with smartwatches. Passive monitoring with smartwatches provides a unique opportunity to explore slowing of upper limb movements during daily life. Here, sensor data segments during arm movements were identified from the circa 90% non-walking periods in the passive monitoring sensor data stream, using the squared magnitude of the accelerometer sensor movement as the sensor feature. This same feature has been related to decreased expressivity in patients with schizophrenia with negative symptoms³⁸. Here,

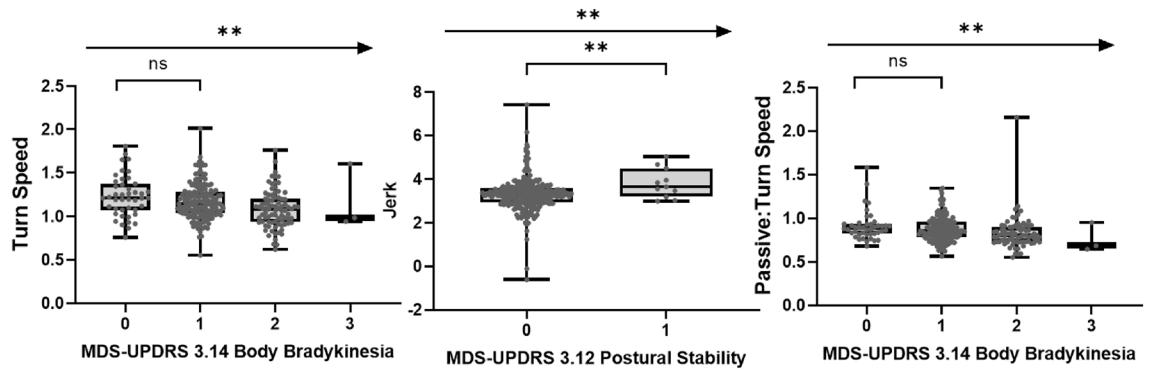


Figure 4. Association of sensor features from Phonation/Speech and eSDMT tests with corresponding clinical MDS-UPDRS measures at baseline (ns = $P > 0.05$; * = $P \leq 0.05$; ** = $P \leq 0.01$; *** = $P \leq 0.001$; **** = $P \leq 0.0001$). eSDMT electronic Symbol Digit Modalities test, MDS-UPDRS Movement Disorder Society—Unified Parkinson's Disease Rating Scale, MFCC2 Mel Frequency Cepstral Coefficient 2.

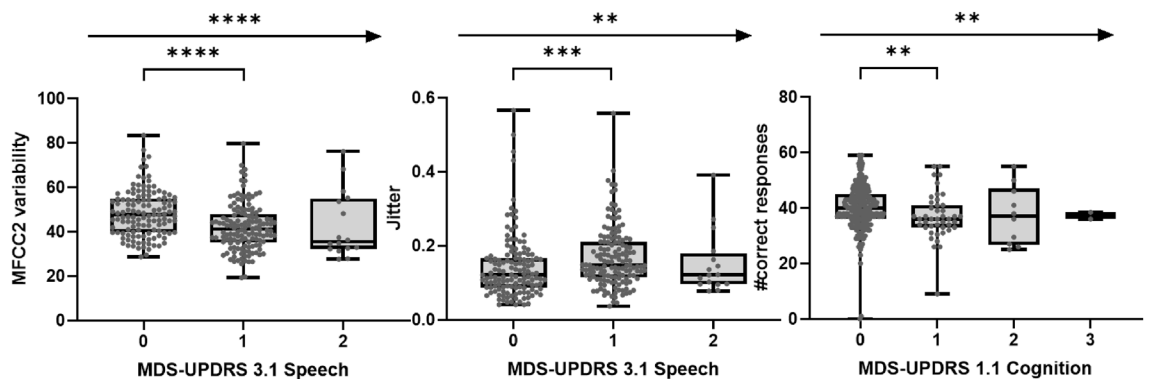


Figure 5. Association of sensor features from U-turn, Balance tests and Passive monitoring of gait with corresponding clinical MDS-UPDRS measures at baseline (ns = $P > 0.05$; * = $P \leq 0.05$; ** = $P \leq 0.01$; *** = $P \leq 0.001$; **** = $P \leq 0.0001$). MDS-UPDRS Movement Disorder Society—Unified Parkinson's Disease Rating Scale.

arm movement power was specifically related to the MDS-UPDRS bradykinesia subscore and item scores, as well as the rigidity subscore, and is in line with a slowing of hand movement in daily non-gait-related activities such as gesturing when speaking, eating, etc. These findings are consistent with previous research with wrist-worn wearables, which traditionally focused on arm swing during gait^{39–41}, as well as multi-sensor systems used to measure the impact of bradykinesia on activities of daily living^{15,42}. Thus, passively monitored motor behavior in daily life may facilitate our understanding of the effect and burden of PD on individuals' daily lives.

DHT measurement of bradyphrenia. The eSDMT⁴³ is commonly applied to measure psychomotor slowing, or bradyphrenia, one of the earliest cognitive signs in PD, appearing up to 5 years prior to a PD dementia diagnosis²⁰. However, as the test requires multiple cognitive functions, it is not surprising that it is sensitive to many forms of neurologic impairment⁴⁴. Indeed, while SDMT performance is reduced in PD⁴⁵, impairments are exacerbated in individuals with PD with concomitant vascular⁴⁶ and amyloid⁴⁷ imaging findings. A standard SDMT outcome measure, number of correct responses in 90 s, was pre-specified for the present analyses of the eSDMT, and showed 'good'²² test–retest reliability (ICC = 0.75). However, it correlated only weakly ($\rho = -0.18$) with the MDS-UPDRS item 1.1, assessing global cognitive impairment. This finding is surprising given the catch-all nature of both the eSDMT and MDS-UPDRS item 1.1., but may be accounted for by the fact that cognitive impairments were excluded during the screening process in the PASADENA study, leading to a truncation of range in both scores (see Supplementary Fig. 1). We note that we attempted to minimize the effect of bradykinesia on eSDMT scores by requiring a simple tap response on a number pad displayed at the bottom half of the smartphone screen. Nevertheless, to mitigate the risk of this confound, eSDMT performance could be controlled by a non-cognitively demanding motor test using a similar response format.

DHT measurement of voice and speech. Voice and speech impairments in PD are varied and generally summarized under the term dysarthria, and include resonatory, articulatory, phonatory, prosodic and respiratory components⁴⁸. This symptomatology and its relevance to patients' daily lives motivated the inclusion of a Sustained Phonation task in the suite of active tests, and the development of the novel Speech test. Voice jitter was pre-selected as a proxy of disordered vocal fold function for the sustained phonation test. In line with previ-

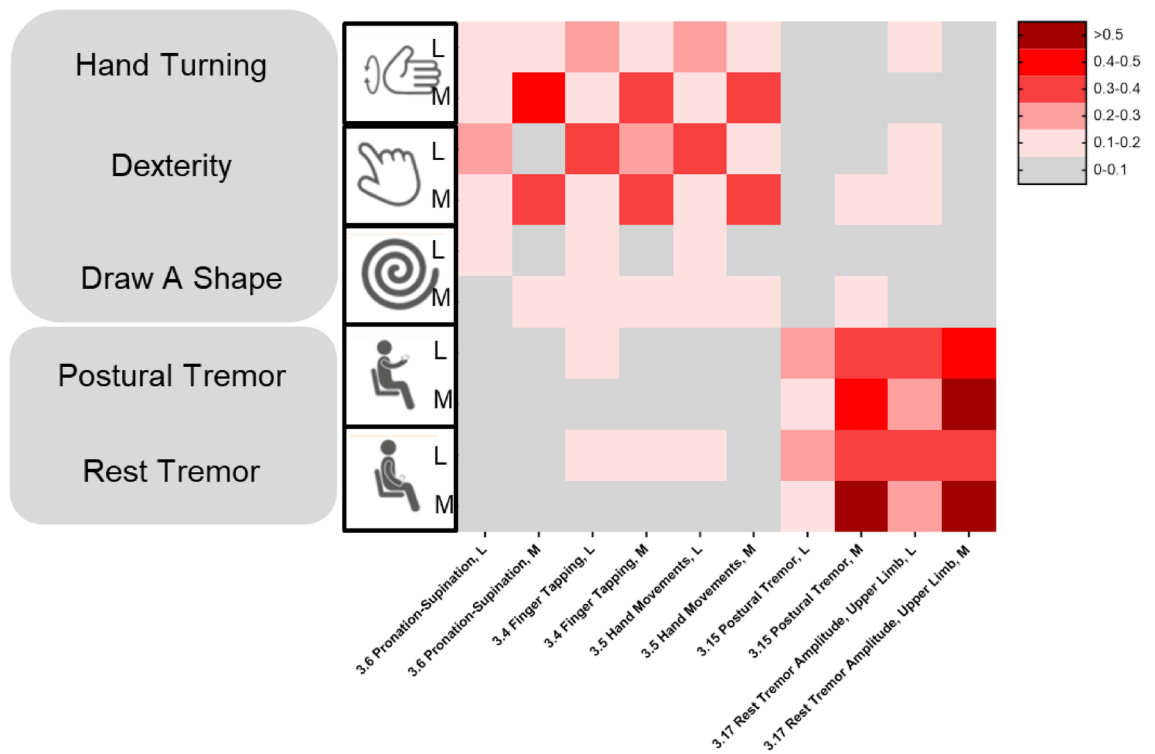


Figure 6. Correlations between active test sensor features from the more and less affected sides (M and L, respectively) with MDS-UPDRS Part III item scores evaluating M and L. *L* less affected side, *M* more affected side, *MDS-UPDRS* Movement Disorder Society—Unified Parkinson's Disease Rating Scale.

ous research⁴⁹, increased voice jitter correlated weakly with MDS-UPDRS 3.1. (Speech) scores, and differentiated individuals with slight speech disturbances (MDS-UPDRS 3.1. score of 1) from those with no perceivable speech impairment at the site visit (MDS-UPDRS 3.1. score of 0). In the Speech active test, monotonicity (i.e. Mel Frequency Cepstral Coefficient 2 [MFCC] 2 fundamental frequency variability) was selected as the sensor feature of prosodic deficits based on previous research demonstrating that this feature differentiated individuals with PD from healthy controls⁴⁸. In the present study, MFCC2 variability correlated with MDS-UPDRS 3.1. (Speech) scores, and differentiated participants with MDS-UPDRS 3.1. scores of 0 and 1. The bulbar MDS-UPDRS Part III composite item score was designed to gauge the severity of motor impairments in body parts involved in speech production. Despite a truncation of range in this score (average < 3/20 points), MFCC2 variability correlated with the bulbar score, indicating that this feature may estimate the severity of motor impairments in the speech apparatus. Future research will investigate further richly multi-faceted aspects of speech function to better understand motor and cognitive behavior in PD.

DHT measurement of tremor, turning and balance. The Roche PD Mobile Application v2 aims to assess the broad array of motor signs in PD and related movement disorders. Thus, besides bradykinesia, speech, voice, and psychomotor slowing, tremor (rest, postural), turning during gait, and balance were also assessed. The rest and postural tremor active test features corresponded most strongly to the respective MDS-UPDRS concepts of tremor, as demonstrated by the highest correlation overall with any MDS-UPDRS item and subscale scores. This is consistent with similar DHT reports^{5,25,50}. The novel U-turn test (which instructed individuals, if safe to do so, to walk several paces and make a U-turn at least five times) and the identification of turning while walking throughout the day in passive monitoring sensor data, were motivated by findings that turning is particularly impaired in PD^{5,51,52}. For example, a 360 degree walking turn and instrumented timed-up-and-go test showed strong reliability and discriminated controls from PD participants^{53,54}. Similarly, sensor-based measures of turn speed in daily life differentiated PD individuals from controls⁵⁵. In the present study, turn speed measured in both the active test and passive setting correlated with MDS-UPDRS 3.14. body bradykinesia item scores, but was not specifically related to MDS-UPDRS PIGD relative to other subscores. While neither measure of turn speed differentiated between less and more affected individuals on MDS-UPDRS body bradykinesia scores of 0 versus 1, both differentiated between individuals in Hoehn and Yahr Stage I versus II. Although participants were not instructed to 'turn as fast as possible' to ensure a safe conduct of the active test, the U-turn test showed numerically higher correlations with body bradykinesia compared with passive turning speed, in line with similar profile of performance (active testing) versus capacity (passive monitoring) scores previously demonstrated for gait speed⁵⁶. In the balance active test, the jerk sensor feature correlated with the MDS-UPDRS 3.12. postural stability item score, similar to previous reports^{5,57}, and differentiated individuals with MDS-UPDRS item 3.12 scores of 0 versus 1, but failed to differentiate individuals in Hoehn and Yahr Stage I versus II. We speculate that

this negative finding may reflect the low levels of gait and postural instability impairments in the present cohort (mean PIGD = 1).

DHT composite scores. A composite summary score of individual features across diverse assessments is expected to provide a more robust measure of global PD severity and progression, especially given the heterogeneous nature of PD. Several DHT solutions besides the Roche PD Mobile Application v2 administer different motor active tests, and some additionally collect passive monitoring data^{5,17,18}. Supplementary Table 2 provides a high-level comparison of these DHT solutions. All solutions contain active tests for tremor and tapping, but vary with respect to the inclusion of other upper limb, postural stability/gait, cognition, and voice/speech tests, and whether passively monitored motor data are collected. The power of combining different features across the tests in these DHTs has been shown via machine learning models that predict MDS-UPDRS total scores (Roche PD Mobile Application v1)⁵⁸ or lead to a new score based on differentiation of ON and OFF L-dopa states⁵⁹, and distinguished between healthy controls, idiopathic Rapid Eye Movement and PD^{16,60}. A machine learning approach was also used to combine different HopkinsPD baseline sensor features to predict clinically significant events (e.g. falls, functional impairment) at the 18-month follow-up⁶¹. In contrast to data-driven approaches to composite score development, a clinical outcomes assessment approach could be applied whereby information from individuals with PD informs the selection of sensor features such that they optimally reflect what matters most to patients⁶².

Limitations. Several facets of the present study limit the generalizability of the findings. Firstly, all individuals' disease duration was < 2 years, and individuals were in Hoehn and Yahr Stages I or II. Thus, the applicability of the present findings to later-stage or prodromal PD is unknown. The reduced range of disease severities also appeared to limit the ranges of some DHT and clinical measures, which consequently limited the possibility to detect relationships between the two (Supplementary Fig. 1). Also, further research is necessary to better understand the suitability of this remote monitoring approach for later-stage patients with more severe cognitive or visual impairments. Second, since Roche PD Mobile Application v2 data are not yet available from neurologically normal individuals, sensor feature cut-off values differentiating normal from impaired motor behavior could not yet be calculated. It should be also noted that comparisons between DHT measures and clinical measures such as the MDS-UPDRS can also be affected by limitations in the clinical measures; if an active test is not adequately reflected by a clinical measure, the ability to detect meaningful correlations is reduced. Finally, only two continuous 2-week periods of DHT data were analyzed; thus, the long-term adherence to the remote monitoring procedure and ability of sensor features to detect changes over time remain to be established. Towards this end, it is critical to quantify and report test–retest reliabilities of sensor feature scores towards assessing a sensor feature's potential to detect changes over time⁶³ and any deviation from normal progression as a function of e.g. pharmacological interventions.

The Roche PD Mobile Application v2 was designed to measure the severity of early PD core motor signs and to provide information complementary to established clinical outcome measures. This remote monitoring approach enables high-frequency (i.e. daily) assessments with low average daily burden. The frequent measurement coupled with the high sensitivity of smartphone/smartwatch sensors may increase signal-to-noise of digital outcome measures for clinical research and provide novel insights into patients' functioning in daily life.

Methods

Participants. Baseline Roche PD Mobile Application v2 data from 316 dopaminergic-treatment-naïve individuals recently diagnosed with dopamine transporter imaging with single-photon emission computed tomography-confirmed PD (Hoehn and Yahr Stages I–II, diagnosis ≤ 2 years) were analyzed (see Table 2 for demographic and clinical characteristics, and Supplementary Table 6 for non-parametric descriptive statistics). All individuals were enrolled in an ongoing randomized, double-blind, placebo-controlled, Phase II clinical trial (PASADENA Part 1; NCT03100149) of prasinezumab (RO7046015/PRX002), an anti- α -synuclein monoclonal antibody (see Pagano et al. 2021)⁶⁴.

All of the 59 PASADENA sites received approval from their institutional review boards or ethics committees and collected data used for the present analyses, and written informed consent was provided by all participants. The study is being conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice.

Roche PD Mobile Application v2. The Roche PD Mobile Application v2 consists of dedicated applications installed on a provisioned smartphone and smartwatch (see Fig. 1). The PD Mobile Application prompted participants to perform the active tests described below. All unilateral tests were performed twice, once with each side of the body.

1. *Draw A Shape*³⁷ participants were instructed to trace six different shapes (two diagonal lines [once drawn up, once drawn downwards], and a square, circle, figure of 8, and spiral) on the smartphone screen with their index finger, as quickly and as accurately as possible (timeout: 30 s);
2. *Dexterity* participants were instructed to alternately tap two touchscreen buttons with their index finger as quickly and regularly as possible (20 s per hand);
3. *Hand Turning* holding the smartphone in the outstretched hand while seated, participants were instructed to rotate the hand as quickly and fully as possible such that the phone faced up and down (10 s per hand);
4. *Speech* participants were provided with three open-ended questions, one after the other, and instructed to read each out loud and answer each question out loud (20 s per question);

Characteristic	Participants (N = 316)
Age, mean (SD), years	59.9 (9.10)
Male, n (%)	213 (67%)
MDS-UPDRS Part IA Non-Motor Aspects of Experience of Daily Living (rater), mean (SD)	1.16 (1.62)
MDS-UPDRS Part IB Non-Motor Aspects of Experience of Daily Living (patient/caregiver), mean (SD)	3.45 (2.78)
MDS-UPDRS Part II Motor Aspects of Experiences of Daily Living (patient), mean (SD)	5.33 (4.04)
MDS-UPDRS Part III Motor Examination (rater), mean (SD)	21.47 (9.00)
MDS-UPDRS Part VI Motor Complications, mean (SD)	NA
MDS-UPDRS Total Score, mean (SD)	31.41 (12.78)
PIGD score, mean (SD)	1.04 (0.90)
Bradykinesia score, mean (SD)	10.50 (5.71)
Rigidity score, mean (SD)	4.58 (2.95)
Tremor score, mean (SD)	5.56 (4.01)
Speech score, mean (SD)	2.72 (2.19)
Axial symptoms, mean (SD)	2.04 (1.71)
Hoehn and Yahr stage, %	
Stage I	25
Stage II	75
Time since diagnosis, mean (SD), months	10.1 (6.5)
MoCA, mean (SD)	29.19 (1.01)

Table 2. Baseline characteristics of participants. *MDS-UPDRS* Movement Disorder Society—Unified Parkinson’s Disease Rating Scale, *MoCA* Montreal Cognitive Assessment, *NA* not applicable, *PIGD* Postural Instability/Gait Disorders, *SD* standard deviation.

5. *Phonation* participants were instructed to make a single, continuous “aaaah” sound for as long as possible with one breath and in a steady pitch and volume while the phone was held at the ear (timeout: 30 s);
6. *Postural tremor* participants were instructed to sit with their eyes closed, and to hold the smartphone in an outstretched hand while counting down out loud from a pre-specified number that differed for each test administration (15 s per hand);
7. *Rest tremor* participants were instructed to sit with their eyes closed and to hold the phone in the palm of their hand, with their forearm resting on their thigh, and to count down out loud from a pre-specified number that differed for each test administration (15 s per hand);
8. *Balance* while standing with the smartphone in a running belt at waist height with the phone placed at the front of the body, participants were instructed to stand still with their arms at their side (30 s);
9. *U-turn* participants were instructed to place the smartphone in a running belt with the phone placed at the front of the body, and to walk between two points at least four steps apart at normal speed, completing at least five turns in 60 s;
10. *eSDMT*⁴³ participants were instructed to match a sequence of displayed symbols to the respective numbers using a displayed coding key as quickly and as accurately as possible (90 s).

The Roche PD Mobile Application v2 additionally administered questionnaires, which are not the focus of the present report. For passive monitoring, participants were instructed to carry their smartphone (e.g. in their trouser pocket or in the pouch of a provided running belt) and wear their smartwatch as they conducted the daily active tests and their normal daily activities. No active tests were administered directly via the smartwatch.

Procedure. Participants were provided with an Android smartphone (Galaxy S7, Samsung, Seoul, South Korea) and smartwatch (Moto G 360 2nd Gen Sport; Motorola, Chicago, USA) during a screening visit at the latest 7 days prior to the baseline clinical visit, and trained on the use of the devices and the Roche PD Mobile Application v2. Participants were instructed to open the application on the provisioned smartphone every morning. Active tests were scheduled automatically such that half of the motor tests were presented on alternating days, and the eSDMT every 2 weeks (Fig. 1), with a total expected testing time (per day) including transitions and test-start countdowns between tests of 5–10 min (including eSDMT). All data were stored in encrypted files on the smartphone and sent by WiFi to a cloud storage facility each time the smartphone connected to the Internet.

Baseline clinical assessments included the MDS-UPDRS⁸, from which subscale scores were generated (i.e. *PIGD*, *bradykinesia*, *rigidity*, *tremor*)²⁸ The MDS-UPDRS was administered according to standardized procedures, and all MDS-UPDRS raters completed online MDS-UPDRS training by the MDS. Additionally, a ‘bulbar score’ was defined as the composite sum of MDS-UPDRS items 2.1 Speech; 2.2 Saliva and Drooling; 2.3 Chewing and Swallowing; 3.1 Speech; and 3.2 Facial expression.

Sensor data processing. The raw sensor data from the smartphone and smartwatch were extracted and processed using a dedicated internally developed backend infrastructure. Custom algorithms implemented in Python were applied on quality-controlled sensor data (e.g. for correct test execution) and converted data into pre-defined ‘sensor features’, one per active test performed and side of body (if applicable) and one for passive monitoring. Features were selected based on previous literature and their relevance to PD (Supplementary Table 3).

1. *Draw a shape:* The feature *Spiral celerity* combines drawing accuracy and drawing speed (accuracy/speed) of the spiral drawing.
2. *Dexterity: Tapping variability* quantifies the variability of tapping speed as measured by the standard deviation of the time between consecutive tap events. This feature has already shown positive results in a previous study in PD⁵.
3. *Hand turning: Median hand turning speed* is the median turn speed over all segmented hand rotations to estimate bradykinesia.
4. *Speech:* The MFCC2 is the ratio between vocal tract resonance of the high and vocal fold vibration of the low Mel-frequency bands affected in PD. MFCC values have already shown case/control differences in other studies of PD⁵. For this novel feature, MFCC2s of consecutively voiced speech segments (i.e. parts of speech that are longer than 200 ms and are acoustically distinguishable) were calculated and averaged over each segment. MFCC2s are interpreted as a measure of speech monotonicity.
5. *Phonation: Voice jitter* is defined as a mean of the absolute differences between the period of adjacent pitch cycles, normalized by the mean pitch period, multiplied by 100. This definition of jitter is also referred to as jitter:local⁴⁹. Jitter represents the variability of the speech fundamental frequency (pitch period) from one cycle to another and is a measure of micro-instability of vocal fold vibration, where higher values of the feature indicate higher instability of vocal fold vibration.
6. *Rest and postural tremor: Log median squared energy* measures the average acceleration magnitude, which is a proxy for the average amplitude during tremor induced by hand movements when trying to hold the hand still. A similar feature showed clinical validity for PD in a previous study⁵.
7. *Balance: Log sway jerk* describes the jerkiness (i.e. irregular, non-smooth accelerations) of movements when trying to stand still and may be a marker of disease progression in PD⁶⁵.
8. *U-turn: Median turn speed* describes the average turn speed of all turns completed during the U-turn test. The same feature correlated with clinically assessed gait impairment in previous studies in individuals with PD and multiple sclerosis⁵.
9. *SDMT: Number of correct responses* is the standard feature also reported in the traditional paper-based in-clinic SDMT⁴³.
10. *Passive monitoring—gait: Median turn speed in passive monitoring* is the average turn speed of all turns detected over a given day of smartphone sensor recording, and had previously demonstrated discriminability between individuals with PD and control participants⁶⁶.
11. *Passive monitoring—non-gait arm movements⁶⁷: Median arm movement power (non-gait)* is the median of the integrated squared acceleration magnitude (i.e. power) over all identified arm movements during non-gait data segments in a given day of sensor recording with the smartwatch. As such, it measures the intensity of arm movements during activities of daily living (gesturing when speaking, grabbing something, etc.) that do not occur during periods of walking (i.e. does not reflect arm swing while walking). Here, we hypothesize that a reduced intensity of arm movements is associated with bradykinesia.

All features reported in this manuscript are based on smartphone sensor data, with the exception of ‘arm movement power (non-gait)’, which leverages passively acquired sensor data collected with the smartwatch.

Data underwent quality control (QC) checks to ensure that the tests had been performed properly. Towards this end, QC metrics were generated. For example, one QC metric quantified the amount of energy from the accelerometer during the Hand Turning test to estimate whether the smartphone was lying still (e.g. on a table) or moving during the test. 0.3% (n = 179/56,786) of digital active test data did not meet the pre-specified QC thresholds and were therefore excluded from the analyses.

Statistical analyses. Sensor features from passive monitoring and each active test performed were summarized (median) over 2-week intervals starting at the baseline visit (Weeks 1 and 2) and in the 2-week period thereafter, provided that ≥ 3 data points were available during each 2-week testing interval. Where applicable, sensor features were assigned to less/more affected side (for definition see Supplementary Material). For convergent/divergent validity (i.e. degree of association with related/unrelated symptom domains), the averaged (median) sensor data collected during the first two study weeks were compared with clinical data collected at the baseline visit (Day 1) using Spearman’s correlations. Adherence and test–retest metrics were calculated for aggregated sensor features for the first two 2-week study periods. Adherence was defined as the number of fully completed active testing sessions relative to the number of all possible active testing sessions. For passive monitoring, also calculated over the first two 2-week study periods, the average number of hours per day participants carried the provisioned study smartphone with them and wore the study smartwatch was calculated. Sensor feature test–retest reliabilities were quantified with the ICC between averaged values of the first and second contiguous 2-week periods. To investigate the sensitivity of sensor features to subtle or very early symptoms, sensor features from participants receiving MDS-UPDRS item scores of 0 versus 1 were compared using Mann–Whitney U tests. For known-groups validity (i.e. differences between pre-defined groups where a difference is

prima facie expected), sensor features were compared between participants in Hoehn and Yahr Stage I versus II, and by comparing sensor feature values from less and more affected sides, both using Mann–Whitney U tests.

Ethics declarations. Participants were identified for potential recruitment using site-specific recruitment plans prior to consenting to take part in this study. Recruitment materials for participants had received Institutional Review Board or Ethics Committee approval prior to use. The following Institutional Review Boards ruled on ethics of the PASADENA study: Ethikkommission der Medizinischen Universität Innsbruck, Innsbruck, Austria; Comité de Protection des Personnes (CPP) Ouest IV, Nantes, France; Ethikkommission der Universität Leipzig and Geschäftsstelle der Ethikkommission an der medizinischen Fakultät der Universität Leipzig, Leipzig, Germany; Ethikkommission der Fakultät für Medizin der Technischen Universität München, München, Germany; Ethikkommission der Universität Ulm (Oberer Eselsberg), Ulm, Germany; Landesamt für Gesundheit und Soziales Berlin and Geschäftsstelle der Ethik-Kommission des Landes Berlin, Berlin, Germany; Ethikkommission des FB Medizin der Philipps-Universität Marburg, Marburg, Germany; Ethikkommission an der Medizinischen Fakultät der Eberhard-Karls-Universität und am Universitätsklinikum Tübingen, Tübingen, Germany; Ethikkommission an der medizinischen Fakultät der HHU Düsseldorf, Düsseldorf, Germany; Ethikkommission der LÄK Hessen, Frankfurt, Germany; CEIm Hospital Universitari Vall d’Hebron, Barcelona, Spain; Copernicus Group Independent Review Board, Puyallup, Washington, USA; Western Institutional Review Board, Puyallup, Washington, USA; the University of Kansas Medical Center Human Research Protection Program, Kansas City, Kansas, USA; Oregon Health & Science University Independent Review Board, Portland, Oregon, USA; Northwestern University Institutional Review Board, Chicago, Illinois, USA; Spectrum Health Human Research Protection Program, Grand Rapids, Michigan, USA; the University of Vermont Committees on Human Subjects, Burlington, Vermont, USA; Beth Israel Deaconess Medical Center Committee on Clinical Investigations, New Procedures and New Forms of Therapy, Boston, Massachusetts, USA; Vanderbilt Human Research Protection Program Health Sciences, Nashville, Tennessee, USA; University of Maryland, Baltimore Institutional Review Board, Baltimore, Maryland, USA; University of Southern California Institutional Review Board, Los Angeles, California, USA; Columbia University Medical Center Institutional Review Board, New York, New York, USA; University of Southern California San Francisco Institutional Review Board, San Francisco, California, USA; University of Pennsylvania Institutional Review Board, Philadelphia, Philadelphia, USA; HCA—HealthOne Institutional Review Board, Denver, Colorado, USA. All Institutional Review Boards gave ethical approval of the study.

Data availability

Data will be made available upon reasonable request. Please direct your request to dbm.datarequest@roche.com. For further details on Roche’s Global Policy on the Sharing of Clinical Information and how to request access to related clinical trial documents, see https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

Code availability

The code used to perform all of the statistical analyses described herein are available upon request to dbm.datarequest@roche.com.

Received: 31 January 2022; Accepted: 30 June 2022

Published online: 15 July 2022

References

- Maetzler, W., Liepelt, I. & Berg, D. Progression of Parkinson’s disease in the clinical phase: Potential markers. *Lancet Neurol.* **8**, 1158–1171. [https://doi.org/10.1016/S1474-4422\(09\)70291-1](https://doi.org/10.1016/S1474-4422(09)70291-1) (2009).
- Sacks, L. & Kunkoski, E. Digital health technology to measure drug efficacy in clinical trials for Parkinson’s disease: A regulatory perspective. *J. Parkinsons Dis.* **11**, S111–S115. <https://doi.org/10.3233/JPD-202416> (2021).
- Espay, A. J. *et al.* A roadmap for implementation of patient-centered digital outcome measures in Parkinson’s disease obtained using mobile health technologies. *Mov. Disord. Off. J. Mov. Disord. Soc.* **34**, 657–663. <https://doi.org/10.1002/mds.27671> (2019).
- Espay, A. J. *et al.* Technology in Parkinson’s disease: Challenges and opportunities. *Mov. Disord. Off. J. Mov. Disord. Soc.* **31**, 1272–1282. <https://doi.org/10.1002/mds.26642> (2016).
- Lipsmeier, F. *et al.* Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson’s disease clinical trial. *Mov. Disord. Off. J. Mov. Disord. Soc.* **33**, 1287–1297. <https://doi.org/10.1002/mds.27376> (2018).
- Dawson, V. L. & Dawson, T. M. Promising disease-modifying therapies for Parkinson’s disease. *Sci. Transl. Med.* <https://doi.org/10.1126/scitranslmed.aba1659> (2019).
- Parkinson Progression Marker Initiative. The Parkinson progression marker initiative (PPMI). *Prog. Neurobiol.* **95**, 629–635. <https://doi.org/10.1016/j.pneurobio.2011.09.005> (2011).
- Goetz, C. *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov. Disord. Off. J. Mov. Disord. Soc.* **23**, 2129–2170. <https://doi.org/10.1002/mds.22340> (2008).
- Andrich, D. & Marais, I. *A Course in Rasch Measurement Theory: Measuring in the Educational, Social and Health Sciences* (Springer, 2019).
- 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. <https://doi.org/10.1007/s00415-019-09348-3> (2019).
- Postuma, R. B. *et al.* MDS clinical diagnostic criteria for Parkinson’s disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* **30**, 1591–1601. <https://doi.org/10.1002/mds.26424> (2015).
- 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. <https://doi.org/10.1002/mds.27361> (2018).

13. Horvath, K. *et al.* Minimal clinically important differences for the experiences of daily living parts of movement disorder society-sponsored unified Parkinson's disease rating scale. *Mov. Disord. Off. J. Mov. Disord. Soc.* **32**, 789–793. <https://doi.org/10.1002/mds.26960> (2017).
14. Sheng, Y., Zhou, X., Yang, S., Ma, P. & Chen, C. Modelling item scores of Unified Parkinson's Disease Rating Scale Part III for greater trial efficiency. *Br. J. Clin. Pharmacol.* **87**, 3608–3618 (2021).
15. Kyritsis, K. *et al.* Assessment of real life eating difficulties in Parkinson's disease patients by measuring plate to mouth movement elongation with inertial sensors. *Sci. Rep.* **11**, 1632. <https://doi.org/10.1038/s41598-020-80394-y> (2021).
16. Jha, A. *et al.* The CloudUPDRS smartphone software in Parkinson's study: Cross-validation against blinded human raters. *NPJ Parkinson's Dis.* **6**, 36. <https://doi.org/10.1038/s41531-020-00135-w> (2020).
17. Stamate, C. *et al.* The cloudUPDRS app: A medical device for the clinical assessment of Parkinson's Disease. *Pervasive Mob. Comput.* **43**, 146–166 (2018).
18. Arora, S. *et al.* Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism Relat. Disord.* **21**, 650–653. <https://doi.org/10.1016/j.parkreldis.2015.02.026> (2015).
19. Bot, B. M. *et al.* The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci. Data* **3**, 160011. <https://doi.org/10.1038/sdata.2016.11> (2016).
20. Johnson, D. K., Langford, Z., Garner-Villarreal, M., Morris, J. C. & Galvin, J. E. Onset of mild cognitive impairment in Parkinson disease. *Alzheimer Dis. Assoc. Disord.* **30**, 127–133 (2016).
21. Hasan, H., Athauda, D. S., Foltynie, T. & Noyce, A. J. Technologies assessing limb bradykinesia in Parkinson's disease. *J. Parkinsons Dis.* **7**, 65–77 (2017).
22. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012> (2016).
23. Rabelo, A. G. *et al.* Objective assessment of bradykinesia estimated from the wrist extension in older adults and patients with Parkinson's disease. *Ann. Biomed. Eng.* **45**, 2614–2625. <https://doi.org/10.1007/s10439-017-1908-3> (2017).
24. Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. *Mov. Disord. Off. J. Mov. Disord. Soc.* **18**, 738–750. <https://doi.org/10.1002/mds.10473> (2003).
25. Kassavetis, P. *et al.* Developing a tool for remote digital assessment of Parkinson's disease. *Mov. Disord. Clin. Pract.* **3**, 59–64. <https://doi.org/10.1002/mdc3.12239> (2015).
26. Lee, C. Y. *et al.* A validation study of a smartphone-based finger tapping application for quantitative assessment of bradykinesia in Parkinson's disease. *PLoS ONE* **11**, e0158852. <https://doi.org/10.1371/journal.pone.0158852> (2016).
27. Lalvay, L. *et al.* Quantitative measurement of Akinesia in Parkinson's disease. *Mov. Disord. Clin. Pract.* **4**, 316–322 (2016).
28. Orozco-Arroyave, J. R. *et al.* Apkinson: The smartphone application for telemonitoring Parkinson's patients through speech, gait and hands movement. *Neurodegener. Dis. Manag.* **10**, 137–157 (2020).
29. Delrobaei, M., Tran, S., Gilmore, G., McIsaac, K. & Jog, M. Characterization of multi-joint upper limb movements in a single task to assess bradykinesia. *J. Neurol. Sci.* **4**, 337–342. <https://doi.org/10.1016/j.jns.2016.07.056> (2016).
30. Mentzel, T. *et al.* Reliability and validity of an instrument for the assessment of bradykinesia. *Psychiatry Res.* **238**, 189–195. <https://doi.org/10.1016/j.psychres.2016.02.011> (2016).
31. di Biase, L. *et al.* Quantitative analysis of bradykinesia and rigidity in Parkinson's disease. *Front. Neurol.* **9**, 121. <https://doi.org/10.3389/fneur.2018.00121> (2018).
32. Saunders-Pullman, R. *et al.* Validity of spiral analysis in early Parkinson's disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* **23**, 531–537. <https://doi.org/10.1002/mds.21874> (2007).
33. Zham, P., Kumar, D., Dabnichki, P., Poosapadi Arjunan, S. & Raghav, S. Distinguishing different stages of Parkinson's disease using composite index of speed and pen-pressure of sketching a spiral. *Front. Neurol.* **8**, 435. <https://doi.org/10.3389/fneur.2017.00435> (2017).
34. Danna, J. *et al.* Digitalized spiral drawing in Parkinson's disease: A tool for evaluating beyond the written trace. *Hum. Mov. Sci.* **65**, 80–88 (2019).
35. San Luciano, M. *et al.* Digitized spiral drawing: A possible biomarker for early Parkinson's disease. *PLoS ONE* **11**, e0162799. <https://doi.org/10.1371/journal.pone.0162799> (2016).
36. Memedi, M. *et al.* Automatic spiral analysis for objective assessment of motor symptoms in Parkinson's disease. *Sensors (Basel)* **15**, 23727–23744. <https://doi.org/10.3390/s150923727> (2015).
37. Creagh, A. P. *et al.* Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. *Physiol. Meas.* **41**, 054002. <https://doi.org/10.1088/1361-6579/ab8771> (2020).
38. Umbrecht, D., Cheng, W. Y., Lipsmeier, F., Bamdadian, A. & Lindemann, M. Deep learning-based human activity recognition for continuous activity and gesture monitoring for schizophrenia patients with negative symptoms. *Front. Psychol.* **11**, 574375. <https://doi.org/10.3389/fpsy.2020.574375> (2020).
39. Rincon, D. *et al.* Wristbands containing accelerometers for objective arm swing analysis in patients with Parkinson's disease. *Sensors (Basel)* <https://doi.org/10.3390/s20154339> (2020).
40. Zampier, V. C. *et al.* Gait bradykinesia and hypometria decrease as arm swing frequency and amplitude increase. *Neurosci. Lett.* **687**, 248–252. <https://doi.org/10.1016/j.neulet.2018.09.051> (2018).
41. Huang, X. *et al.* Both coordination and symmetry of arm swing are reduced in Parkinson's disease. *Gait Posture* **35**, 373–377. <https://doi.org/10.1016/j.gaitpost.2011.10.180> (2012).
42. Thorp, J. E., Adamczyk, P. G., Ploeg, H. L. & Pickett, K. A. Monitoring motor symptoms during activities of daily living in individuals with Parkinson's disease. *Front. Neurol.* **9**, 1036. <https://doi.org/10.3389/fneur.2018.01036> (2018).
43. Smith, A. *Symbol Digit Modalities Test Manual* (Western Psychological Services, 1973).
44. Lezak, M. D., Howieson, D. B., Bigler, E. D. & Tranel, D. *Neuropsychological Assessment* 5th edn. (Oxford University Press, 2012).
45. Pascoe, M., Alamri, Y., Dalrymple-Alford, J., Anderson, T. & MacAskill, M. The symbol-digit modalities test in mild cognitive impairment: Evidence from Parkinson's disease patients. *Eur. Neurol.* **79**, 206–210. <https://doi.org/10.1159/000485669> (2018).
46. Linortner, P. *et al.* White matter hyperintensities related to Parkinson's disease executive function. *Mov. Disord. Clin. Pract.* **7**, 629–638. <https://doi.org/10.1002/mdc3.12956> (2020).
47. Fiorenzato, E. *et al.* Brain amyloid contribution to cognitive dysfunction in early-stage Parkinson's Disease: The PPMI dataset. *J. Alzheimer's Dis. JAD* **66**, 229–237. <https://doi.org/10.3233/JAD-180390> (2018).
48. Magee, M., Copland, D. & Vogel, A. P. Motor speech and non-motor language endophenotypes of Parkinson's disease. *Expert Rev. Neurother.* **19**, 1191–1200. <https://doi.org/10.1080/14737175.2019.1649142> (2019).
49. Rusz, J., Cmejla, R., Ruzickova, H. & Ruzicka, E. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J. Acoust. Soc. Am.* **129**, 350–367. <https://doi.org/10.1121/1.3514381> (2011).
50. Scanlon, B. K. *et al.* An accelerometry-based study of lower and upper limb tremor in Parkinson's disease. *J. Clin. Neurosci.* **20**, 827–830. <https://doi.org/10.1016/j.jocn.2012.06.015> (2013).
51. Mellone, S., Mancini, M., King, L. A., Horak, F. B. & Chiari, L. The quality of turning in Parkinson's disease: A compensatory strategy to prevent postural instability? *J. Neuroeng. Rehabil.* **13**, 39. <https://doi.org/10.1186/s12984-016-0147-4> (2016).
52. Crenna, P. *et al.* The association between impaired turning and normal straight walking in Parkinson's disease. *Gait Posture* **26**, 172–178. <https://doi.org/10.1016/j.gaitpost.2007.04.010> (2007).

53. Salarian, A. *et al.* iTUG, a sensitive and reliable measure of mobility. *IEEE Trans. Neural Syst. Rehabil. Eng.* **18**, 303–310. <https://doi.org/10.1109/TNSRE.2010.2047606> (2010).
54. Soke, F. *et al.* Reliability and validity of the timed 360 degrees turn test in people with Parkinson's disease. *Eur. Geriatr. Med.* **11**, 417–426. <https://doi.org/10.1007/s41999-019-00285-y> (2020).
55. Mancini, M. *et al.* Continuous monitoring of turning in Parkinson's disease: Rehabilitation potential. *NeuroRehabilitation* **37**, 3–10. <https://doi.org/10.3233/NRE-151236> (2015).
56. Atrsaei, A. *et al.* Gait speed in clinical and daily living assessments in Parkinson's disease patients: Performance versus capacity. *NPJ Parkinsons Dis.* **7**, 24. <https://doi.org/10.1038/s41531-021-00171-0> (2021).
57. Mancini, M. *et al.* ISway: A sensitive, valid and reliable measure of postural control. *J. Neuroeng. Rehabil.* **9**, 59. <https://doi.org/10.1186/1743-0003-9-59> (2012).
58. Chen, O. Y. *et al.* Building a machine-learning framework to remotely assess Parkinson's disease using smartphones. *IEEE Trans. Biomed. Eng.* **67**, 3491–3500. <https://doi.org/10.1109/TBME.2020.2988942> (2020).
59. Zhan, A. *et al.* Using smartphones and machine learning to quantify Parkinson disease severity: The mobile Parkinson disease score. *JAMA Neurol.* **75**, 876–880. <https://doi.org/10.1001/jamaneuro.2018.0809> (2018).
60. Arora, S. *et al.* Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. *Neurology* **91**, e1528–e1538. <https://doi.org/10.1212/WNL.0000000000006366> (2018).
61. Lo, C. *et al.* Predicting motor, cognitive & functional impairment in Parkinson's. *Ann. Clin. Transl. Neurol.* **6**, 1498–1509. <https://doi.org/10.1002/acn3.50853> (2019).
62. Taylor, K. I., Staunton, H., Lipsmeier, F., Nobbs, D. & Lindemann, M. Outcome measures based on digital health technology sensor data: Data- and patient-centric approaches. *NPJ Digit. Med.* **3**, 97. <https://doi.org/10.1038/s41746-020-0305-8> (2020).
63. Dodge, H. H. *et al.* Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. *PLoS ONE* **10**, e0138095. <https://doi.org/10.1371/journal.pone.0138095> (2015).
64. Pagano, G. *et al.* A phase II study to evaluate the safety and efficacy of prasinezumab in early Parkinson's disease (PASADENA): Rationale, design, and baseline data. *Front. Neurol.* **12**, 705407. <https://doi.org/10.3389/fneur.2021.705407> (2021).
65. Mancini, M. *et al.* Postural sway as a marker of progression in Parkinson's disease: A pilot longitudinal study. *Gait Posture* **36**, 471–476. <https://doi.org/10.1016/j.gaitpost.2012.04.010> (2012).
66. Cheng, W. Y. *et al.* Smartphone-based continuous mobility monitoring of Parkinson's disease patients reveals impacts of ambulatory bout length on gait features. In *IEEE Life Sciences Conference* 166–169. <https://doi.org/10.1109/LSC.2017.8268169> (2017).
67. Cheng, W.-Y. *et al.* Large-scale continuous mobility monitoring of Parkinson's disease patients using smartphones. In *International Conference on Wireless Mobile Communication and Healthcare* 12–19. https://doi.org/10.1007/978-3-319-98551-0_2 (2018).

Acknowledgements

We gratefully acknowledge the participants of the PASADENA study and their families, as well as the PASADENA Investigators. We also acknowledge the Roche PD Mobile application v2 development team, especially Hans van Wesenbeeck, Juraj Korcek, Monnika Broennimann for their contributions in the development and deployment of the solution. The authors thank Kiran Verma, of Chrysalis Medical Communications UK, for providing editorial support and Megan Speakman, of MediTech Media UK, for medical editing assistance.

Author contributions

F.L., K.I.T., R.B.P., T.K., B.M., and M.L. designed the study. T.K., B.M., A.B., W.L.P., and J.S.E., were involved in data collection. F.L., K.I.T., E.V.V., W.L.P., W.C., Y.Z., D.W., and M.L. analysed the data. F.L., K.I.T., R.B.P., E.V.V., H.S., W.Z., G.P. and M.L. were involved in data interpretation. F.L., K.I.T., and M.L. were involved in drafting the work. All authors revised and gave input on the article. All authors approved the final content and are accountable for all aspects of the publication.

Funding

The authors declare that the study is funded by F. Hoffmann-La Roche Ltd and Prothena Inc. F. Hoffmann-La Roche Ltd and Prothena Inc were involved in the study design, collection, analysis, interpretation of data, the writing of this article and the decision to submit it for publication. Medical editorial support were funded by F. Hoffmann-La Roche, in accordance with Good Publication Practice (GPP3) guidelines (<http://www.ismpp.org/gpp3>).

Competing interests

The Authors declare no Competing Non-Financial Interests but the following Competing Financial Interests: FL, KIT, EVV, TK, AB, WLP, WC, YZ, DW, JSE, ABo, HS and GP reports personal fees from F. Hoffmann-La Roche Ltd. BM reports personal fees and other from F. Hoffmann-La Roche Ltd, during the conduct of the study; personal fees from Biogen, personal fees from Servier, non-financial support from Amprion, Grants and personal fees from Michael J. Fox Foundation for Parkinson's Research, Grants from DFG, Grants from EU (Horizon2020), Grants from Parkinson Fonds Deutschland, Grants from Hilde Ulrich Stiftung. WZ reports personal fees from F. Hoffmann-La Roche Ltd and Prothena Ltd. RBP reports Grants and personal fees from Fonds de la Recherche en Sante, Grants from Canadian Institute of Health Research, Grants from The Parkinson Society of Canada, Grants from Weston-Garfield Foundation, Grants from Michael J. Fox Foundation, Grants from Webster Foundation, Grants from National Institute of Health, Grants and personal fees from Roche, personal fees from Takeda, personal fees from Teva Neurosciences, personal fees from Biogen, personal fees from Boehringer Ingelheim, personal fees from TheraNexus, personal fees from GE HealthCare, personal fees from Jazz Pharmaceuticals, personal fees from AbbVie, personal fees from Janssen, personal fees from Otsuko, personal fees from Phytopharmica, personal fees from Inception Sciences, other from Parkinson Canada, personal fees from Paladin. ML reports personal fees from F. Hoffmann-La Roche Ltd and has a patent US20190200915A1 pending to Hoffmann-La Roche Inc., a patent EP3701542A2 pending to F. Hoffmann-La Roche AG, a patent WO2019215230A1 pending to F. Hoffmann-La Roche AG, and a patent WO2021055443A1 pending to F. Hoffmann-La Roche AG.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-15874-4>.

Correspondence and requests for materials should be addressed to F.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022