

# UCSF

## UC San Francisco Previously Published Works

### Title

Postural instability in Parkinson's disease assessed with clinical "pull test" and standardized postural perturbations: effect of medication and body weight support

### Permalink

<https://escholarship.org/uc/item/8mf537hg>

### Journal

Journal of Neurology, 270(1)

### ISSN

0340-5354

### Authors

Lu, Chiahao  
Louie, Kenneth H  
Stutz, Amber M  
[et al.](#)

### Publication Date

2023

### DOI

10.1007/s00415-022-11375-6

Peer reviewed



Published in final edited form as:

*J Neurol.* 2023 January ; 270(1): 386–393. doi:10.1007/s00415-022-11375-6.

## Postural instability in Parkinson’s disease assessed with clinical “pull test” and standardized postural perturbations: effect of medication and body weight support

Chiahao Lu<sup>1</sup>, Kenneth H. Louie<sup>2,3</sup>, Amber M. Stutz<sup>1,4</sup>, Colum D. MacKinnon<sup>1</sup>, Scott E. Cooper<sup>1</sup>

<sup>1</sup>Department of Neurology, University of Minnesota, 516E. 717 Delaware Building. 717 Delaware St. SE, Minneapolis, MN 55414, USA

<sup>2</sup>Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN 55414, USA

<sup>3</sup>Present Address: Department of Neurological Surgery, University of California, San Francisco, CA 94143, USA

<sup>4</sup>Present Address: Neurology, Sanford Brain and Spine Center, Fargo, ND 58103, USA

### Abstract

**Objective**—This experiment tested if balance performance differed between a standardized treadmill surface perturbation task and a clinical pull test and was affected by medication or the presence of body weight support in people with Parkinson’s disease (PD).

**Methods**—Twenty-seven individuals were tested (14 PD in both ON- and OFF-medication states). Clinical pull test and rapid forward (backward fall) translations of the support surface were applied to induce postural reactions requiring at least 1 step to restore balance. The effects of pull type (clinical vs. treadmill), partial bodyweight support (0 vs 20% body weight) and group (control, PD ON-meds and PD OFF-meds) on reactive stepping as well as practice/learning effect were examined. The number of steps taken and the first step duration were entered in linear repeated-measures mixed-effect models separately.

**Results**—The effects of pull type, group, and bodyweight support were all significant in both metrics, as was ON- vs. OFF-medication. A significant interaction term (group x pull type) was found in the first step duration, showing that the group difference was greater in treadmill compared to the clinical pull test. A significant practice effect was also observed within and across testing sessions.

---

Chiahao Lu, luxxx214@umn.edu.

**Conflicts of interest** The authors declare that they have no conflict of interest.

**Ethical standard** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The protocol was approved by the University of Minnesota Institutional Review Board (No. 1508M77143)

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00415-022-11375-6>.

Supplier

Treadmill: Motek medical, Vleugelboot 14, 3991 CL Houten, Netherlands.

**Conclusions**—A standardized treadmill perturbation performed slightly better than the classical pull test in distinguishing between groups, and partial weight support did not substantially degrade the test's performance to detect the balance deficits in people with PD.

### Keywords

Parkinson's disease; Reactive stepping; Postural instability; Pull test; Postural perturbation

---

### Introduction

Postural instability is a common motor symptom in people with Parkinson's disease (PD). Approximately half of all people with PD will develop severe gait and balance problems which are often associated with falls, loss of independent community ambulation and eventually decreased quality of life [1, 2].

Postural instability in PD is clinically evaluated with the Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [3] pull test (item 3.12 "Postural Stability"). In this test, an examiner stands behind the patient and pulls abruptly on the patient's shoulders. The pull must be strong enough that the patient requires one or more steps backward to recover balance ("reactive stepping"). Despite its widespread use in clinical settings and clinical trials, the test suffers from a major disadvantage, the variability of the clinician-imposed perturbation in terms of method of delivery, force and duration. Even highly trained examiners in a clinical trial do not perform the test consistently [4]. Such variability of the perturbation will introduce both random noise, and examiner bias which undermines the validity of clinical trials.

Standardized perturbations have been used to study postural instability in laboratory settings, such as surface translation with movable platforms or a direct pull from mechanical actuators [5-9]. These methods are considered to generate more accurate and reliable measurements, but it is unclear if these techniques discriminate postural deficits between people with PD and matched controls better than the simple pull test. Here, we compared balance recovery performance between a group with PD (OFF- and ON-medication) and matched controls in response to a rapid anterior translation of the support surface on a treadmill (imposed backward fall) and during standard clinical pull test. Many balance assessment systems use partial body weight support (BWS) systems to reduce testing difficulty or ensure safety, but BWS might also compromise test sensitivity. Thus, we also examined whether the presence of partial BWS affected performance during the two perturbation tasks. Furthermore, we examined how repeated exposure to a perturbation affected the reactive stepping performance, since a prominent "first trial" effect has been reported for reactive stepping [10-13].

We hypothesized: (1) the performance of reactive stepping would systematically differ between PD and controls in both perturbation types, (2) differences between the PD and controls would be smaller during ON-medication (ON-meds) state compared to OFF-medication (OFF-meds) state, (3) the surface translation test would distinguish PD vs. Control, and ON-meds vs. OFF-meds *at least* as well as the standard clinical test, (4) partial BWS would improve the reactive stepping performance of both PD and controls, (5) PD

vs. Control and ON-meds vs. OFF-meds effects would be detectable even with partial body weight support, and (6) reactive stepping would improve with repeated exposure within a testing session as well as from one session to the next session.

## Methods

We studied 27 individuals including 14 diagnosed with PD by their treating neurologist (3 females) and 13 controls (8 females; see Table 1 for summary demographics and Table S1 for detailed demographics of individuals with PD). No participants had been treated surgically for PD. Exclusion criteria for all participants included medical conditions other than PD that significantly impaired use of lower limbs and dementia or cognitive impairment including documented Montreal Cognitive Assessment (MoCA) [14] or Folstein [15] score < 24. The PD participants were asked to complete two visits, ON-meds and OFF-meds, in random order. For the OFF-meds visit, the PD participants were tested after overnight withdrawal from immediate release PD medications and 24-h withdrawal from sustained release PD medications. Eleven PD participants completed two visits with at least one week apart.

Participants stood on a stationary treadmill (C-Mill, Motek, Amsterdam) with safety harness loosely suspended from an overhead sliding track and an experienced movement disorders specialist (S.C.) behind the participant. Postural perturbation inducing backward reactive stepping was induced using two methods: (1) a rapid anterior translation of the support surface on a treadmill<sup>a</sup> (“treadmill” condition) or (2) a standard clinical pull test (“clinical” condition). The treadmill perturbation has been previously described [8]. Briefly, after a verbal warning, the treadmill was turned on following a variable time interval to produce servocontrolled anterior belt movement sufficient to evoke 1–2 steps from neurologically normal individuals. A tachometer directly measured belt movement to verify the standard trajectory. All participants required at least one step to recover balance in all treadmill trials. For the clinical perturbation, the pull test was performed by administering a quick tug backwards on the shoulders by the movement disorders specialist in conformance with the MDS-UPDRS specifications.

Falls were prevented using a safety harness loosely suspended from an overhead sliding track and the examiner behind the participant. Safety bars were located on each side of the participant, who was, however instructed not to grab them to regain balance during the testing. Participants started each trial looking at a fixation target (2.7 m away at head level) and with weight evenly distributed between the feet (verified by monitoring real-time center of pressure trajectory). The participants were permitted to step with either foot. Most participants stepped consistently with the same foot on all trials. A total of 48 trials were collected in 3 sets. Each set included 16 trials, beginning with 3 trials of clinical perturbation, followed by 10 trials of treadmill perturbation and ending with 3 trials of the clinical perturbation. The first and third sets were without BWS, while BWS of 20% total body weight was provided during the second (middle) set. Participants were videotaped at 60 fps from the waist down in the frontal plane, and a blinded rater counted the number of steps taken and noted the video frame numbers of liftoff, and touchdown for the first step. The duration of the first step was then calculated from these frame numbers. A total of

1767 trials were collected; 57 trials were missing due to equipment failure and one subject was unable to complete the protocol. A separate analysis of step lengths, extracted from the videos using machine-learning techniques, is the subject of a separate paper [16].

Linear repeated-measures mixed-effect models were applied to the number of steps taken and the first step duration respectively with fixed factors of Group (PD OFF-meds, PD ON-meds, Control), Pull Type (Treadmill vs. Clinical) and BWS (0% BWS vs. 20% BWS) and subject as random factor with age and sex included as covariates. Post-hoc analysis for Group was performed with Tukey Test for multiple pairwise comparisons. The group differences in age and sex were examined using *t*-test and chi-square, respectively. All statistical analyses were done in R [17], with significance testing by ANOVA (F test, Satterthwaite's degrees of freedom method) using lme4 and lmeTest libraries. Significance level was set to 0.05 with Holm–Bonferroni correction for multiple tests using “p.adjust” function.

## Results

There were no significant differences in age ( $t_{(25)} = -0.92$ ,  $p = 0.37$ ) and sex ( $\chi^2 = 2.98$ ,  $p = 0.08$ ) between the PD and control groups.

### Number of steps

Group ( $F_{(2,44.5)} = 17.08$ ,  $p < 0.001$ ), Pull Type ( $F_{(1,1737.9)} = 19.51$ ,  $p < 0.001$ ) and BWS ( $F_{(1,1737.9)} = 119.41$ ,  $p < 0.001$ ) effects were all statistically significant. Post hoc showed PD patients took fewer steps (mean:  $-0.34$  steps) during ON-meds than OFF-meds ( $p < 0.001$ ) and control participants took fewer steps (mean  $-0.65$  steps) than PD patients who were OFF-meds ( $p = 0.04$ ) independent of Pull Type and BWS. Both covariates, age ( $F_{(1,23.3)} = 0.07$ ,  $p = 0.80$ ) and sex ( $F_{(1,23.5)} = 3.48$ ,  $p = 0.08$ ) were not significant. These results are shown graphically in Fig. 1a.

Repeating the analysis with a Poisson link function (i.e., treating number of steps as an ordinal variable) in a generalized linear mixed model (instead of the standard Gaussian) provided the same result as above, including that PD patients took fewer steps during ON-meds than OFF-meds ( $p < 0.01$ ). However, with the Poisson link, the difference between control participants and PD patients while OFF-meds did not reach significance ( $p = 0.06$ ).

The direction of the medication effect was a normalizing one, since PD patients in the OFF-meds state took more steps than Controls, and medication reduced this difference. The PD ON-meds vs. Control comparison was non-significant in all analyses.

To further assess the effects of Pull Type and BWS, we constructed two augmented models, (1) with an additional Group X BWS term, and (2) with an additional Group X Pull Type term. Neither interaction was significant ( $F_{(2,1735.8)} = 1.57$ ,  $p = 0.21$ ;  $F_{(2,1735.7)} = 0.05$ ,  $p = 0.95$ ), so we retained the null hypothesis that neither Pull Type nor BWS changed the effect of medication on the number of steps. The same result was obtained with the model using the Poisson linking function.

### First step duration

Similar results were observed in the first step duration (Fig. 1b) with Group ( $F_{(2,43.7)} = 5.11$ ,  $p = 0.01$ ), Pull Type ( $F_{(1,1737.4)} = 265.85$ ,  $p < 0.001$ ) and BWS ( $F_{(1,1737.4)} = 98.02$ ,  $p < 0.001$ ) all statistically significant. PD patients had a longer first step duration (mean: + 0.01 s) during ON-meds than OFF-meds ( $p < 0.001$ ) independent of BWS. In the augmented model, Group X BWS was non-significant ( $F_{(2,1735.3)} = 0.43$ ,  $p = 0.65$ ); however, the Group X Pull Type interaction was significant ( $F_{(2,1735.3)} = 9.15$ ,  $p < 0.001$ ) reflecting greater Group difference with the Treadmill than with the Clinical perturbation. Post hoc showed significant differences in PD patients between ON-meds and OFF-meds only in the Treadmill ( $p < 0.01$ ). Both covariates, age ( $F_{(1,23.0)} = 0.89$ ,  $p = 0.35$ ) and sex ( $F_{(1,23.1)} = 1.78$ ,  $p = 0.20$ ) were not significant (see supplementary material for analysis in only PD patients including with clinical characteristics as additional covariates).

### Relationship between number of steps and first step duration

To examine whether the first step duration was related to number of steps taken, we augmented the main model of number of steps taken with a term for the first step duration. The result showed a significant relationship in both Gaussian ( $F_{(1,1749.5)} = 161.98$ ,  $p < 0.001$ ) and Poisson models ( $\chi^2 = 65.54$ ,  $p < 0.001$ ) independent of the other terms (Age, Sex, BWS, Pull Type, Group). This is shown graphically in Fig. 2 (note that the relationship is only apparent in the PD group because there was very little variance in the number of steps for the Control group).

### Habituation effect

We frequently observed participants requiring more steps to recover balance initially compared to subsequent trials. This has also been noted by others [10]. To quantify this effect, we augmented our main model with a term for TrialNumber and a TrialNumber X Group interaction term. The TrialNumber was significant ( $F_{(1,1736.7)} = 60.99$ ,  $p < 0.001$ ; coefficient value:  $-0.013$ ) with a non-significant TrialNumber X Group interaction ( $F_{(2,1734.7)} = 2.13$ ,  $p = 0.12$ ). This indicates that participants required more steps initially in a series of trials (Fig. 3a), independent of Group. Pull Type ( $F_{(1,1736.84)} = 19.78$ ,  $p < 0.001$ ), BWS ( $F_{(1,1736.9)} = 123.64$ ,  $p < 0.001$ ) and Group ( $F_{(1,44.42)} = 17.24$ ,  $p < 0.001$ ) remained significant. Results were similar with the Poisson link function (TrialNumber coefficient:  $-0.010$ ).

Since PD participants were tested on two separate days (for both ON-meds and OFF-meds), we were able to evaluate long-term (weeks) persistence of a practice/learning effect. To examine the effect, we applied the same model on number of steps with a term for session number (i.e., whether the measurements were made at the first or second session) including only PD participants and an interaction for session order and medication state. The session number was significant ( $F_{(1,1011.1)} = 18.94$ ,  $p < 0.001$ ; coefficient:  $-0.31$ ). The added interaction term was not significant ( $F_{(1,7.01)} = 1.30$ ,  $p = 0.29$ ). That is, PD participants tended to take fewer steps on the second session, independent of medication state. This suggests a persistent practice effect (shown graphically in Fig. 3b). Additionally, Pull Type ( $F_{(1,1011.2)} = 9.99$ ,  $p = 0.001$ ), BWS ( $F_{(1,1011.2)} = 63.97$ ,  $p < 0.001$ ) and medication state

( $F_{(1,1011.2)} = 28.60, p < 0.001$ ) remained significant. Results were unchanged with the Poisson link function (session number coefficient:  $-0.15$ ).

## Discussion

This study examined postural instability in PD using two reactive stepping tasks: the standard clinical “pull test” and an imposed translation of the support surface on a treadmill. For both tasks, number of steps taken by the PD patients in the OFF-meds condition was greater than the controls. This finding is consistent with previous studies [6, 9] and demonstrates that both perturbation types produce increased instability in people with PD compared to controls. PD patients also took a shorter duration first step consistently compared to controls. The two metrics (number of steps, duration of first step) were negatively correlated, and behaved very similarly, across all analyses.

The performance of people with PD on both metrics was significantly closer to controls’ in the ON-meds compared to the OFF-meds condition. Dopaminergic medication (primarily levodopa) is first-line therapy to treat PD [18]; however, whether medication improves postural instability in PD remains equivocal [8, 9, 19-23]. In our study, a clear improvement in reactive stepping due to medication was observed. This finding contrasts with the results of de Kam et al. [9] who found no significant improvement in reactive stepping performance with medication. One possible explanation is that our platform perturbation acceleration was greater ( $5 \text{ m/s}^2$  vs.  $1.5 \text{ m/s}^2$ ), thus producing greater imbalance and potentially a larger effect size. Discrepancies might also come from differences in the PD cohorts since our group had a slightly higher daily dose of medications (average levodopa equivalent dose 881 mg vs. 740 mg) and were also slightly older (mean: 68 years vs. 65 years). An important possible explanation for why we detected a medication effect is the larger quantity of data in our study (48 measurements in each medication condition, compared to only 4 in de Kam et al.). It is expected that this would increase statistical power to detect a medication effect. In addition, we were able to replicate our surface perturbation results with the standard clinical pull test, which was not reported in de Kam et al., which allows us to generalize our results beyond the specifics of the platform perturbation protocol.

Partial body weight support is widely available in rehabilitation settings, and can make reactive stepping tests easier, safer, and possibly less stressful for the patient. However, easier testing could translate into a “less sensitive” evaluation, i.e., a ceiling effect. In this study, all participants required fewer steps with 20% BWS, indicating the task was easier, but the difference between PD vs. Controls was preserved as was the difference between ON-meds vs. OFF-meds. This suggests that 20% BWS can be used without significantly affecting the sensitivity of the test. Body weight has a strong relationship with postural stability [24]. Thus, providing the extra weight support may reduce the biomechanical constraints due to weight and increase the dynamic stability range [25].

The clinical pull test has the great advantage of requiring no equipment. However, the correct technique is difficult to maintain even for trained and experienced examiners in clinical trials [4]. A recent study [26] using wearable sensors during a clinical pull test showed that postural responses between hydrocephalus patients from healthy controls could

be distinguished when the magnitude of the applied perturbation was systematically varied across multiple trials. Similar to our findings, a relationship between pull intensity and step length was observed and the slope of this relationship was significantly different between groups. This approach requires multiple trials of various pull intensities and, as we observed, may require accounting for the practice/learning effect over trials.

A treadmill perturbation does not require a trained and experienced examiner and/or additional wearable sensors and may allow for a more consistent perturbation. While the “clinical standard” pull test performed well compared to the treadmill perturbation, the treadmill perturbation may have performed better than the pull by the shoulders. Specifically, although both types of perturbation performed similarly for the number-of-steps metric, the secondary metric (i.e., first step duration) showed a difference between ON-meds vs. OFF-meds with the treadmill perturbation, but not the pull test. This may reflect underlying biomechanical differences between the types of perturbation, but it could also be due to the greater consistency (lower variability) of the treadmill perturbation. Testing variability may have been further reduced by the larger number of treadmill trials than shoulder-pull trials; however, the clinical standard pull test does not allow an indefinite increase in the number of trials [3]. Note that in our protocol, the pull test was always performed by the same experienced movement disorder specialist and 6 sets of 3 trials were tested. Thus, we are conservatively underestimating the variability of the clinical pull test when the single-examiner constraint is not enforced and only the usual smaller number of trials are performed. Substituting a mechanical actuator for a human examiner is a natural way to standardize the pull test perturbation, and our results support its effectiveness. We used a treadmill as our testing modality because treadmills are widely available in gait labs and even some clinical (rehabilitation) settings, which requires no specialized or specially constructed equipment. Our results should be generalized with caution since we used one particular make and model of treadmill. The protocols and perturbation characteristics provided by other manufacturers and models of treadmill may be different and may not yield the same results.

We compared the “first step duration” to the classical “number of steps” metric and found similar performance. Both metrics were obtainable from video, but the first step duration is potentially more automatable using a motion analysis system, pressure-sensitive shoe insert, or forceplate (see our previous study [27]). We expected greater sensitivity with the more nearly continuous “duration of the first step” metric than with the coarser, integer “number of steps.” In fact, the duration metric performed equally well as the number of steps. This may reflect that our “number of steps” was averaged over large number of individual trials, giving a quantity which was more fine-grained than the individual integer values. Given its good performance, especially in combination with the treadmill perturbation, the duration of first step metric can serve as an alternative when averaging a large number of integer count of steps measurements is not feasible. One study limitation is that our main analysis omitted step latency for the shoulder perturbation because video recordings were from the waist down. This was aimed to maximize resolution of the lower extremities and avoid capturing participants' faces in order to preserve participant's anonymity. As a result, shoulder perturbation onset was not available from video data. Nonetheless, these experiments were done with an accelerometer on the dorsal surface of the examiner's hand,



allowing us to detect the timing of onset of the experimenter's application of force to the shoulder, in a total of 508 trials. For an analysis of shoulder perturbation step latency, see Supplementary Material and Fig. S1.

Previous studies have shown that PD patients and controls generate greater postural responses in the first trial of perturbation [10, 11]. This effect was also observed in the current study. Nanhoe-Mahabier et al. suggested that the habituation rate is slower in PD compared to controls [10]. We were not able to replicate this finding since the Group X TrialNumber interaction was not significant. Several factors could account for this, including differences in the perturbation protocol, outcome measures and demographics of the sample population. Importantly, the current study demonstrates that group differences in reactive stepping performance can be detected when controlling for the first step response (significant Group effect even with TrialNumber included in the model). Furthermore, we demonstrated a practice/learning effect over a longer time window (weeks apart) in people with PD (significant effect of visit number independent of medication state). This result supports the practice of randomizing the order of ON/OFF—medication testing.

## Conclusions

In summary, people with Parkinson's disease performed reactive stepping better in the ON-medication compared to the OFF-medication state, over a range of measurement techniques, including the standard clinical pull test. Both the clinical pull test and treadmill surface translation protocols induced an increased number of steps and reduced first step duration in PD, but the surface translation task performed slightly better in differentiating between conditions (OFF vs ON-meds). The additional 20% body weight support did not substantially degrade test performance. Step counts and first step duration were both useful metrics for evaluating reactive balance.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding

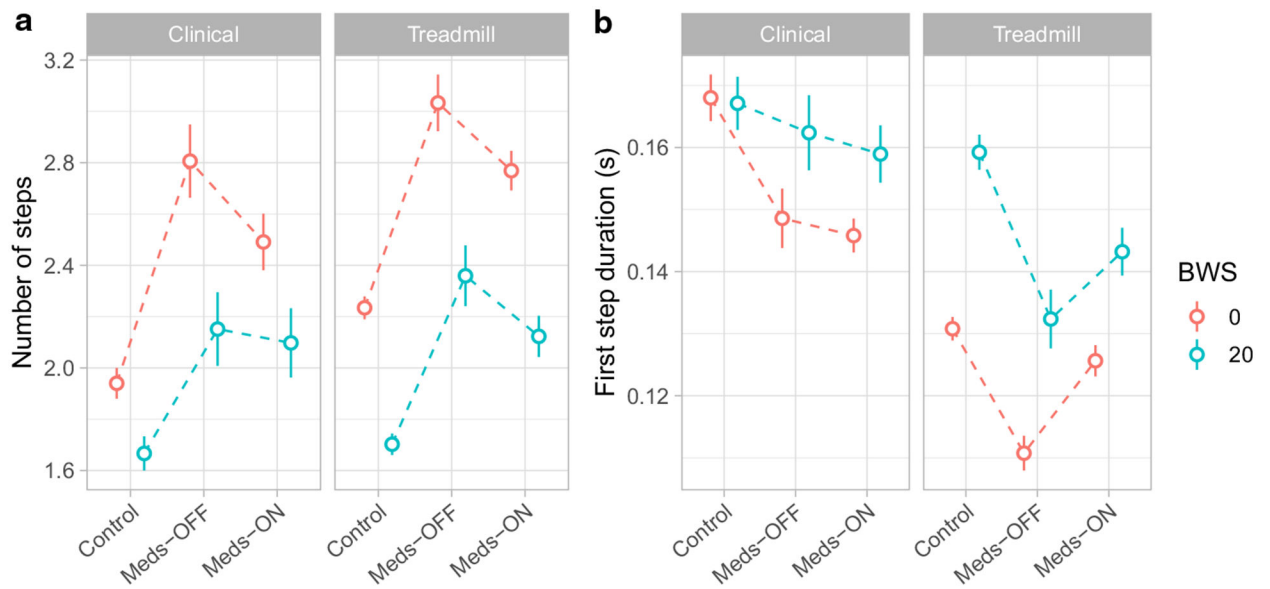
National Center for Advancing Translational Sciences (Grant No. UL1TR000114); National Institute of Neurological Disorders and Stroke (Grant Nos. P50 NS098573, P50 NS123109, R01 NS088679, R01 NS124814-01A1); State of Minnesota (Grant No. MnDRIVE).

## References

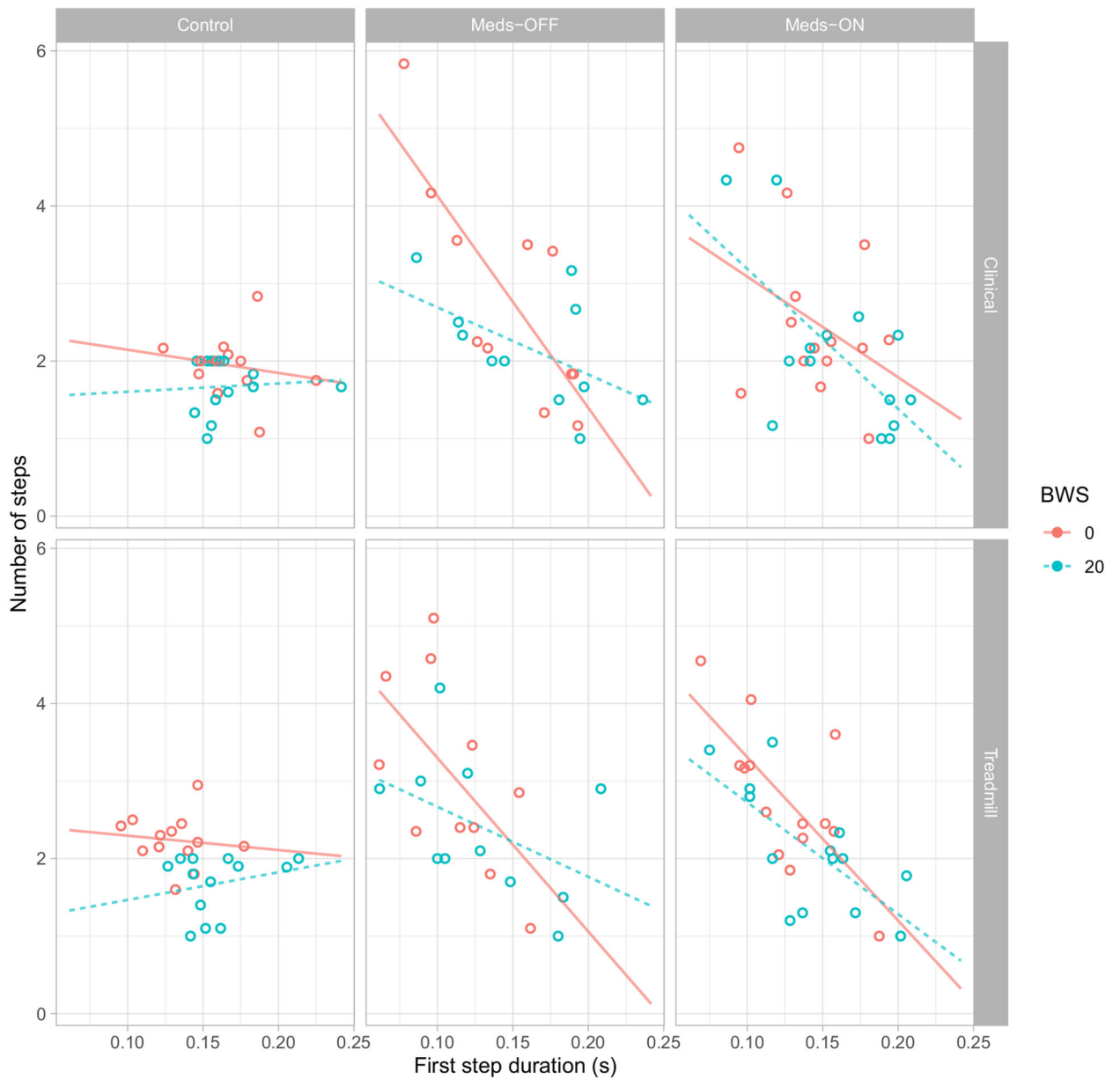
1. Williams DR, Watt HC, Lees AJ (2006) Predictors of falls and fractures in bradykinetic rigid syndromes: a retrospective study. *J Neurol Neurosurg Psychiatry* 77:468–473. 10.1136/jnnp.2005.074070 [PubMed: 16543524]
2. Schrag A, Jahanshahi M, Quinn N (2000) What contributes to quality of life in patients with Parkinson's disease? *J Neurol Neurosurg Psychiatry* 69:308–312. 10.1136/jnnp.69.3.308 [PubMed: 10945804]
3. Goetz CG, Tilley BC, Shaftman SR et al. (2008) Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 23:2129–2170. 10.1002/mds.22340 [PubMed: 19025984]

4. Munhoz RP, Li J-Y, Kurtinecz M et al. (2004) Evaluation of the pull test technique in assessing postural instability in Parkinson's disease. *Neurology* 62:125–127. 10.1212/WNL.62.1.125 [PubMed: 14718714]
5. Carpenter MG, Allum JHJ, Honegger F et al. (2004) Postural abnormalities to multidirectional stance perturbations in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 75:1245–1254. 10.1136/jnnp.2003.021147 [PubMed: 15314109]
6. McVey MA, Amundsen S, Barnds A et al. (2013) The effect of moderate Parkinson's disease on compensatory backwards stepping. *Gait Posture* 38:800–805. 10.1016/j.gaitpost.2013.03.028 [PubMed: 23607994]
7. Horak FB, Dimitrova D, Nutt JG (2005) Direction-specific postural instability in subjects with Parkinson's disease. *Exp Neurol* 193:504–521. 10.1016/j.expneurol.2004.12.008 [PubMed: 15869953]
8. Peterson DS, Horak FB (2016) The effect of levodopa on improvements in protective stepping in people with Parkinson's disease. *Neurorehabil Neural Repair* 30:931–940. 10.1177/1545968316648669 [PubMed: 27162165]
9. de Kam D, Nonnekes J, Oude Nijhuis LB et al. (2014) Dopaminergic medication does not improve stepping responses following backward and forward balance perturbations in patients with Parkinson's disease. *J Neurol* 261:2330–2337. 10.1007/s00415-014-7496-3 [PubMed: 25228002]
10. Nanhoe-Mahabier W, Allum JHJ, Overeem S et al. (2012) First trial reactions and habituation rates over successive balance perturbations in Parkinson's disease. *Neuroscience* 217:123–129. 10.1016/j.neuroscience.2012.03.064 [PubMed: 22542872]
11. Campbell AD, Squair JW, Chua R et al. (2013) First trial and StartReact effects induced by balance perturbations to upright stance. *J Neurophysiol* 110:2236–2245. 10.1152/jn.00766.2012 [PubMed: 23945786]
12. Sutter K, Nonnekes J, Dibilio V et al. (2016) Does the StartReact effect apply to first-trial reactive movements? *PLoS ONE* 11:1–11. 10.1371/journal.pone.0153129
13. Tan JL, Perera T, McGinley JL et al. (2018) Neurophysiological analysis of the clinical pull test. *J Neurophysiol* 120:2325–2333. 10.1152/jn.00789.2017 [PubMed: 30110235]
14. Nasreddine ZS, Phillips NA, Bédirian V et al. (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53:695–699. 10.1111/j.1532-5415.2005.53221.x [PubMed: 15817019]
15. Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189–198. 10.1016/0022-3956(75)90026-6 [PubMed: 1202204]
16. Herbers C, Schroeder J, Lu C et al. (2022) Quantifying clinical evaluations of parkinsonian postural instability through deep learning based video analysis (unpublished data)
17. R Core Team (2018) A language and environment for statistical computing. *R Found. Stat. Comput*
18. Stowe RL, Ives NJ, Clarke C et al. (2008) Dopamine agonist therapy in early Parkinson's disease. *Cochrane Database Syst Rev*. 10.1002/14651858.CD006564.pub2
19. Nova IC, Perracini MR, Ferraz HB (2004) Levodopa effect upon functional balance of Parkinson's disease patients. *Parkinsonism Relat Disord* 10:411–415. 10.1016/j.parkreldis.2004.04.004 [PubMed: 15465397]
20. Horak FB, Frank J, Nutt J (1996) Effects of dopamine on postural control in parkinsonian subjects: scaling, set, and tone. *J Neurophysiol* 75:2380–2396. 10.1152/jn.1996.75.6.2380 [PubMed: 8793751]
21. Bloem BR, Beckley DJ, van Dijk JG et al. (1996) Influence of dopaminergic medication on automatic postural responses and balance impairment in Parkinson's disease. *Mov Disord* 11:509–521. 10.1002/mds.870110506 [PubMed: 8866492]
22. Revilla F, Larsh T, Mani A et al. (2013) Effect of dopaminergic medication on postural sway in advanced Parkinson's disease. *Front Neurol* 4:202 [PubMed: 24379799]
23. Mancini M, Rocchi L, Horak FB, Chiari L (2008) Effects of Parkinson's disease and levodopa on functional limits of stability. *Clin Biomech (Bristol, Avon)* 23:450–458. 10.1016/j.clinbiomech.2007.11.007 [PubMed: 18155331]

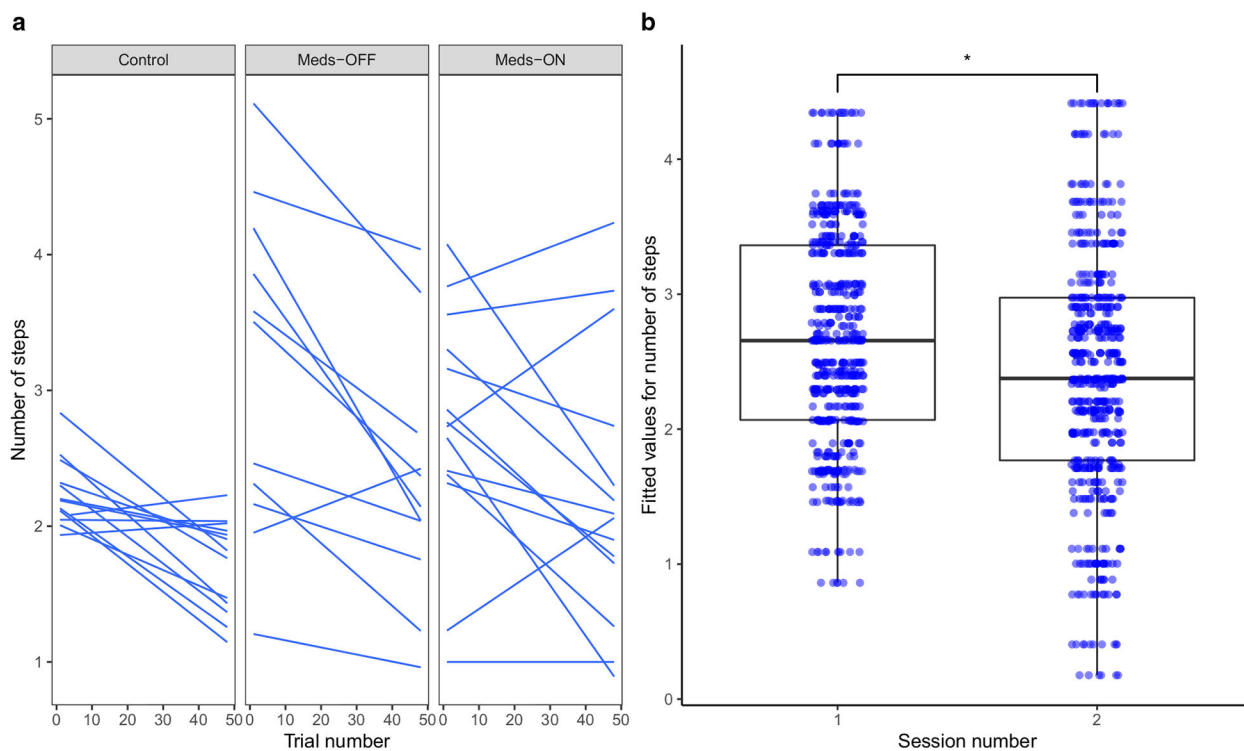
24. Hue O, Simoneau M, Marcotte J et al. (2007) Body weight is a strong predictor of postural stability. *Gait Posture* 26:32–38. 10.1016/j.gaitpost.2006.07.005 [PubMed: 16931018]
25. Błaszczyk JW, Cie linska- wider J, Plewa M et al. (2009) Effects of excessive body weight on postural control. *J Biomech* 42:1295–1300. 10.1016/j.jbiomech.2009.03.006 [PubMed: 19386313]
26. Daly S, Hanson JT, Mavanji V et al. (2022) Using kinematics to re-define the pull test as a quantitative biomarker of the postural response in normal pressure hydrocephalus patients. *Exp Brain Res* 240:791–802. 10.1007/s00221-021-06292-5 [PubMed: 35041069]
27. Lu C, Amundsen-Huffmaster SL, Louie KH et al. (2021) Comparison of forward and backward postural perturbations in mild-to-moderate Parkinson’s disease. *Gait Posture* 84:205–208. 10.1016/j.gaitpost.2020.12.012 [PubMed: 33360643]

**Fig. 1.**

**a** Number of steps taken and **b** first step duration during the clinical and treadmill perturbation tasks for each group. Open circles represent group average values while the vertical lines represent the standard error of the group. Orange is for testing under 0% body weight support (BWS) and blue is for 20% BWS



**Fig. 2.** The relationship between number of steps and first step duration during clinical (top row) and treadmill (bottom row) perturbations. Open circles represent individual average values (orange = 0% BWS, blue = 20% BWS). Regression lines for all subjects in each group are represented as solid lines (0% BWS) and dash lines (20% BWS)



**Fig. 3.**

**a** The relationship between number of steps and trial number. Each blue line represents a fitted regression line for each individual. **b** Fitted values of number of steps for each visit in PD. The values were estimated from the mixed model including factors of age, sex, pull type, body weight support, medication state and test order. Each blue circle represents individual trial. Even taking into account medication state (ON/OFF) and other variables, number of steps were significantly lower on the second session (asterisk denotes the significant main effect of session number  $p < 0.001$ )

**Table 1**

Summary of participants demographics

	Group	
	PD	Control
Sex (Male/Female)	11/3	8/5
Age (years) *	68.1 ± 8.6	69.4 ± 7.7
Disease duration (years) *	11.4 ± 4.2	N/A
MDS-UPDRS motor score OFF-Meds *#	36.8 ± 25.3	N/A
MDS-UPDRS motor score ON-Meds *#	21.3 ± 15.3	N/A
Levodopa equivalent dosage (mg) *	880.7 ± 447.1	N/A

N/A not applicable

\* Values are presented as mean ± SD

# MDS-UPDRS stands for Movement Disorder Society-Sponsored Unified Parkinson's Disease Rating Scale Revision