HHS Public Access

Author manuscript

Physiother Res Int. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Physiother Res Int. 2016 March ; 21(1): 36–46. doi:10.1002/pri.1615.

Consistency in Administration and Response for the Backward Push and Release Test: A Clinical Assessment of Postural Responses

Beth A. Smith[1,*], **Patricia Carlson-Kuhta**[2], and **Fay B. Horak**[2,3]

[1]Division of Biokinesiology and Physical Therapy, University of Southern California, 1540 Alcazar Street, CHP 155, Los Angeles, CA 90089-9006, USA

[2]Balance Disorders Laboratory, Department of Neurology, Oregon Health & Science University, Portland, OR, USA

[3]Portland Veterans Affairs (VA) Medical Center, Portland, OR, USA

Abstract

Background and Purpose—The backward push and release test (PRT) is a standardized clinical test of postural responses elicited by perturbations. Our goal was to determine reliability of administration and response. This will inform clinical administration and determine whether to develop an instrumented version.

Methods—One examiner administered 10 backward PRT trials to adults with Parkinson disease (12), multiple sclerosis (14) and controls (12). We used three-dimensional motion analysis, force plates and instrumented gloves to measure administration and response. Administration variables were angle of posterior trunk lean and the distance of the centre of mass (CoM) behind the ankle. Postural response variables were latency of postural response from release to step initiation and first compensatory step length. Reliability was measured using the range of variables across trials, comparison of first and later trials, intraclass correlations (ICCs) to measure consistency and correlations between administration and response.

Results—There was inherent variability in administration, which affected postural response characteristics. Larger trunk angle and greater CoM–ankle distance were correlated with shorter postural response latencies and larger step lengths. Participant height also had an effect; taller participants had larger trunk angles prior to release resulting in longer latencies and larger step lengths. Using ICCs, consistency of trunk angle was likely acceptable and CoM–ankle distance was high. Consistency of latency was low, while step length was likely acceptable.

Discussion—Despite variability in administration and inconsistency in response, different postural response characteristics were detected between patients with different disease states.

[*]Correspondence: Beth A. Smith, PT, DPT, PhD, Assistant Professor of Research, Division of Biokinesiology and Physical Therapy, University of Southern California, 1540 Alcazar Street, CHP 155, Los Angeles, CA 90089-9006, USA. beth.smith@usc.edu.

Conflict of interest

OHSU and Dr. Horak have a significant financial interest in APDM, a company that may have a commercial interest in the results of this research and technology. This potential institutional and individual conflict has been reviewed and managed by OHSU.

Presentations: Abstracts of this work were presented at the 2012 Society for Neuroscience Annual Conference (New Orleans, LA).

Based on these results, we will create algorithms to instrument the PRT using inertial movement sensors to collect more sensitive measures of postural responses than observational clinical rating scales. Feedback for appropriate lean angle and calibration for participant height will improve consistency and usefulness of the instrumented PRT.

Keywords

movement; nervous system diseases; postural balance

Introduction

The backward push and release test (PRT) is a standardized clinical test of postural stepping responses. The subject leans backward into the examiner's supporting hands. A compensatory stepping response is elicited through a sudden release of trunk support when the body centre of mass (CoM) is held beyond its backward limits of stability (Jacobs *et al*., 2006; Horak *et al*., 2009). The PRT was created to be a more sensitive and consistent test of postural stability than the pull test (item 30 of the Unified Parkinson's Disease Rating Scale). The pull test also elicits a compensatory stepping response; however, the method of administration is different; the pull test elicits a step when the examiner pulls backward on the subjects' shoulders (Fahn and Elton, 1987; Jacobs *et al*., 2006). The PRT is a more sensitive and consistent test of postural stability than the pull test because examiners apply more consistent forces to subjects with the PRT (Jacobs *et al*., 2006).

Although the PRT could currently replace the pull test as a 'better' assessment of postural stability, both scales have limited sensitivity as they are scored on an observational rating scale based on the size and number of steps to equilibrium. Further, the PRT involves inherent variability in administration because the examiner must judge when the subject's CoM is behind their heels and because patients' willingness to lean back varies. We are proposing here that the PRT could potentially be used as a stand-alone assessment of postural stability; however, increased sensitivity and an understanding of inherent variability in PRT administration and response are needed first. To increase sensitivity, we are interested in creating an instrumented version of the PRT using inertial movement sensors to collect more sensitive measures than an observational rating scale. To understand inherent variability in PRT, we are interested in determining reliability of administration and response to the PRT.

Our questions here are focused on developing an instrumented version of and informing clinical administration of the PRT. How much variability is there in administration when performed repeatedly by the same examiner–subject pair? Does inherent variability in administration significantly affect postural responses? Is the first trial representative of the typical response, or should the average of three trials (or more) be used? These are important considerations for clinical use of the PRT as a stand-alone assessment of postural stability, and the detailed information collected on administration and response will provide guidance for creating algorithms for instrumentation.

Methods

To assess variability in administration and response, the same examiner administered 10 trials of the backward PRT to participants across an age range with-different balance abilities. We used three-dimensional motion analysis, two force plates and instrumented gloves to measure administration and response variables. Administration variables depended on how the examiner positioned the participant (angle of posterior trunk lean and distance of the centre of mass behind the ankle). Response variables reflected how the participant responded to the release of trunk support (latency from release to step initiation, step length, clinical scoring.) Variability was assessed by measuring the range of variables across trials and by using intraclass correlations (ICCs) to measure consistency of administration and response. We used correlations to test for relationships between administration and response.

Participants

Participants were 12 adults with Parkinson disease (PD), 14 with multiple sclerosis (MS) and 12 control participants (refer to Table 1). To represent clinical practice, participants were chosen to represent a range of age, disease severity and state (e.g. ‘on’ or ‘off’ dopamine replacement medications for participants with PD), as shown in Table 1. Our intention was to include a wide variety of responses, as would be encountered by a clinician administering the PRT. Our goal was to describe the broad, general associations between administration and response. We want the test to be sensitive to postural stability impairments regardless of the direction or presentation of the impairment (increased or decreased latency, longer or shorter steps). To further reflect clinical administration, no attempts were made to ‘blind’ the examiner to the participants’ condition. All participants were without conditions that would prevent standing independently for 30 minutes and had vision corrected to 20/40 or better. Participants were not taking medications known to affect balance or attention (other than medications for PD or MS). Participants were recruited through fliers in specialty medical clinics at Oregon Health & Science University and throughout the Portland, OR metropolitan area.

Procedures

Human subject rights were protected and the Oregon Health & Science University Institutional Review Board approved procedures. Participants came into the laboratory for 1 to 1.5 hours. We explained all procedures and answered any questions before they signed an informed consent document. For all participants, we administered a health survey and measured body segment lengths. For participants with PD, we administered part III (motor section) of the Unified Parkinson’s Disease Rating Scale (original version). For participants with MS, a neurologist provided the extended disability status score.

Participants wore shorts, a sleeveless top and their own securely fitting exercise sneakers. We used three-dimensional motion analysis with eight Falcon video cameras (Motion Analysis Corporation, Santa Rosa, CA, USA) to collect body position data at 60 Hz. Reflective markers were placed bilaterally at the fifth metatarsophalangeal joints, heels, lateral malleoli, lateral knee joint centres, greater trochanters, acromion processes, lateral epicondyles, wrist joint centres, mandibles and temples. The examiner wore gloves that had

pressure sensitive pads to measure release time. Glove and force plate data were collected at 200, 480 or 600 Hz with a custom QNX system (QNX Software Systems, Ottawa, Ontario, CA). Different sampling frequencies were used because of technical challenges in synchronizing all equipment; all data had a synchronized start time and were re-sampled to a common frequency before analysis. Participants also wore small inertial movement sensors on their shins, sternum and sacrum for instrumentation of the PRT; these data will not be discussed.

An experienced physical therapist administered 10 trials of the backward PRT to all participants per the instructions in the Balance Evaluation Systems Test (BESTest; Horak *et al*., 2009). Time between repetitions was 15–30 seconds, allowing the examiner and participant time to regain the starting position. The instructions are: "Stand in back and to the side of the patient with one hand on each scapula and ask them to lean backward. (Make sure there is room for them to step backward.) Require them to lean until their shoulders and hips are in back of their heels. Release your support when the subject is in place. Test must elicit a step." Participants stood with one foot on each force plate, positioned using the standardized foot-positioning wedge (trapezoid shape, 10 cm across back edge at heels, 15 cm across front edge towards toes, 20 cm along lateral edges) and tape outlines for consistency between trials. Participants wore a harness attached without tension to the ceiling. Standardized instructions were: "Stand with your arms down at your sides. Lean backward against my hands beyond your backward limits. When I let go, do whatever is necessary, including taking a step, to avoid a fall."(Horak *et al*., 2009). They experienced the lean (but not release) before beginning, to ensure they understood. The first release was recorded as trial 1. In many cases, verbal coaching from the examiner was required to encourage more leaning. In some cases, the examiner used one arm to support the participant and the other on their chest to 'place' them in position when they were resistant to leaning. The examiner judged amount of lean by feeling when the participants' weight passed behind their heels; there is a significant increase in the weight supported by the examiner at this point. They were held here for 3–4 seconds before release to prevent anticipation of an immediate release upon sufficient lean.

Data analysis

We used custom MATLab programs to calculate administration and response variables. CoM was calculated from the weighted sum of the individual body segments (Vaughan *et al*., 1992).

Administration variables were angle of posterior trunk lean and distance of the CoM behind the ankle. Angle of posterior trunk lean was calculated from motion analysis data. The angle was created in the sagittal plane by a marker on the ground and the heel and shoulder of the participant, and the angle at time of release was subtracted from the angle when the participant was standing upright at the start of each trial. The distance of the CoM behind the ankle was the difference between the estimated position of the CoM and the ankle marker in the anterior–posterior direction.

Postural response variables were latency from release to step initiation, step length and clinical scoring. Latency from release to step initiation was measured as the time elapsed

from when the examiner released the participant (measured by glove sensors) to when the foot left the ground to initiate the first compensatory recovery step (measured by the force place). The specific time the examiner released the participant was defined as the point when the glove pressure sensor signal started decreasing from full body support, as this is when the participant started falling. The specific time the foot left the ground to initiate the first recovery step was defined as the point when the vertical force from the force plate signal became 0, indicating the foot left the ground. Step length was the distance the toe marker travelled in the posterior direction from when the participant's foot left the ground to when it first contacted the ground. Clinical response ratings were scored using the BESTest: (3) recovers independently a single, large step; (2) more than one step used, but stable and recovers independently or one step with imbalance; (1) takes several steps to recover equilibrium, or needs minimum assistance; (0) no step, or would fall if not caught, or falls spontaneously (Horak *et al*., 2009).

Statistical analysis

We summarized the range of variables across 10 trials for each participant (Table 2). Of 380 collected trials, we had usable data from 360–371 trials. The number of usable trials varied by type of equipment recording the outcome, for example, sometimes motion analysis data were missing, but other data were recorded. Missing trials occurred randomly because of technical difficulties, and no participant was missing more than one trial for any variable. We used SPSS software (IBM, Chicago, IL) version 19 for statistical analysis (alpha = 0.05).

We tested for differences in administration and response variables by group using univariate analysis of variance (ANOVA) with Bonferroni correction for multiple comparisons. We expected that response would differ by group; however, administration should not. We used repeated measure ANOVA to test for differences in response variables between the first trial, the average of the first 3 trials and the average of all 10 trials.

To determine whether variability in administration systematically influenced response, we performed two tests. We used ICC to test the consistency of a single measurement among repeated trials given our single examiner/rater design. ICCs are recommended for ratio level data; we used a two-way mixed effects model with random administration effects and fixed response effects, reporting the ICCs and 95% confidence intervals [CIs] (Cusick *et al*., 2005; Spence Laschinger, 1992; Weir, 2005). We used one-tailed Pearson correlations (*r*) to test for relationships between administration and response.

Results

Clinical rating scale

Clinical rating scale scores included all possible 0–3 clinical ratings (Horak *et al*., 2009). Participant scores were consistent; most participants (31/38) scored the same clinical rating across all 10 trials. Five participants had ratings that differed by 1 point across 10 trials, while one participant showed a difference of 2 points and one showed a difference of 3 points (refer to Table 2).

Group differences

For administration, there was not a significant group difference in angle of posterior trunk lean (F[2,357] = 1.47, p = 0.23). There was a significant group difference in distance of the CoM behind the ankle (F[2,361] = 43.79, p <0.01). Participants with PD tended to bend their knees, so while the angle of the lean was not different, their CoM was further behind their ankles compared with the other groups. On average, the PD group COM was 38 mm further behind the ankle than the MS group (p <0.01) and 32 mm further than the control group (p <0.01), refer to Table 2).

For response, there was a significant group difference in latency from release to step initiation (F[2,362] = 38.41, p <0.01). Follow-up analysis revealed that participants with MS took longer to initiate their first recovery step. On average, the MS group took 126 ms longer than the control group (p <0.01) and 118 ms longer than the PD group (p <0.01, refer to Figure 1). There was also a significant group difference in the length of the first recovery step (F[2,371] = 41.74, p <0.01). Follow-up analysis revealed that participants with MS took larger and participants with PD took a smaller first recovery steps (refer to Figure 2). On average, the MS group took 5 cm larger steps than the control group (p <0.01) and 16 cm larger steps than the PD group (p <0.01). The PD group took 11 cm smaller steps than the control group (p <0.01, refer to Table 2).

First trial compared with the average of 3 and 10 trials

There was not a significant difference in latency from release to step initiation when comparing the first trial, the average of the first 3 trials and the average of all 10 trials (Huynh-Feldt F[1.9,66.3] = 0.74, p = 0.48). Average latency for the first trial was 297 ms, 283 ms for the average of 3 trials and 287 ms for the average of 10 trials.

There was a significant difference in step length of the first recovery step when comparing the first trial, the average of the first 3 trials and the average of all 10 trials (Huynh-Feldt F[1.6,55.3] = 3.83, p = 0.04). Average step length for the first trial was 44 cm, 42 cm for the average of 3 trials and 41 cm for the average of 10 trials, indicating that the first trial produced a larger first response step than later trials.

Consistency of administration and response

We present ICCs in Table 3. We included nine trials per participant. Interpretations of ICCs are the following: ~1 indicates ‘excellent’ reliability, >0.9 is ‘high’, 0.8–0.9 is ‘moderate’ and 0.7–0.8 is ‘likely acceptable’ (Vincent, 2004). Thus, for administration, consistency of angle of posterior trunk lean was likely acceptable, but consistency of distance of CoM behind the ankle was high. For response, consistency of latency from release to step initiation was low, while consistency of length of the first recovery step was likely acceptable. Whether inconsistent administration affected responses is addressed with correlation analyses.

While investigating low consistency of latency from release to step initiation, we noted that there were a number of trials with latencies less than 150 ms (refer to Figure 1). This is faster than possible for evoked stepping postural response latencies (Burleigh-Jacobs *et al.*,

1997; MacKinnon *et al*., 2007) and indicates that participants may have been anticipating the release, perhaps by some unintentional change in examiner pressure. We re-ran the ICC excluding values under 150 ms (from zero to four trials per participant) with the assumption that an instrumented version of the test would contain a warning to re-do these trials. We still found low consistency, as shown in Table 3.

Correlations between administration and response

We tested for correlations between administration variables (angle of posterior trunk lean, distance of the CoM behind the ankle) and response variables (latency from release to step initiation, step length). We included participant height as a variable that could affect both administration and response. We wanted to know if instrumented test algorithms should adjust individually for height.

The administration variables were correlated: Larger trunk angle was significantly correlated with greater CoM–ankle distance ($r = 0.14$, $p < 0.01$). We observed that heavier and/or taller participants have a CoM that is further behind their ankle at the same angle than lighter and/or shorter participants. For example, participants 8 and 9 have similar heights and trunk angles, however participant 8 is heavier and thus has a greater CoM–ankle distance. Participants 28, 29 and 30 also demonstrated this effect. Another variable is participant behaviour; we observed that some participants with PD would bend their knees while leaning back, which puts the CoM further behind the ankle than it would be with straight knees (refer to Figure 3).

Response variables were also correlated: Longer latency was significantly correlated with larger step length ($r = 0.14$, $p < 0.01$).

In regard to administration affecting response, larger trunk angle and greater CoM–ankle distance were significantly correlated with shorter latency ($r = 0.25$, $p < 0.01$ and $r = 0.14$, $p < 0.01$, respectively, refer to Figure 1) and larger step length ($r = 0.41$, $p < 0.01$ and $r = 0.22$, $p < 0.01$, respectively, refer to Figure 2). Height was significantly correlated with trunk angle ($r = 0.09$, $p = 0.04$), latency ($r = 0.13$, $p = 0.01$) and step length ($r = 0.19$, $p < 0.01$). Taller participants had a larger angle, longer latency and larger step length.

Discussion

Overall, we found three sources of variability in the PRT that have significant, but small, effects on response. First, there is an influence of participant height on administration and response. Taller participants had a larger trunk angle of administration and longer latency and larger step length responses. Although these correlations were statistically significant, they were small, indicating that the relationship exists but explains little of the overall variance in response. The second source of variability was inherent variability in the angle of administration (trunk angle). This ranged from 2 to 6 degrees of difference in angle across 10 trials within a subject. Larger trunk angle was significantly correlated with shorter latency and larger step length; but again, the correlations were statistically significant but small, indicating that the relationship exists but explains little of the overall variance in response. Variability in trunk angle is influenced by the behaviour of the examiner and the

participant. Finally, greater CoM–ankle distance was significantly correlated with shorter latency and larger step length, and again, the correlations were statistically significant but small. The CoM–ankle distance appears to be influenced by patient height, weight and behaviour (whether they bent their knees). Overall, many sources of variability in administration have small effects on response. Some sources of variability are constant (participant height), while others may vary trial to trial (trunk angle, participant behaviour).

Variability in PRT administration is present and has small, significant effects on postural responses. ICCs for consistency of angle of posterior trunk lean across 10 trials was likely acceptable while consistency of distance of CoM behind the ankle was high. This could reflect the examiner adjusting the trunk angle to maintain more consistent CoM–ankle distance, despite changes in participant behaviour. Postural response consistency of length of the first recovery step was likely acceptable, but consistency of postural response latency was low. Even when we removed unrealistic latency values less than 150 ms, consistency remained low. However, we did not find a significant difference in latency from release to step initiation when comparing the first trial, the average of the first 3 trials and the average of all 10 trials, so natural variability in postural response latencies may exceed test administration variability.

Despite variability in administration and inconsistency in response, groups with different disease states (MS, PD and control) responded significantly differently to the PRT. The group differences were consistent with known characteristics of their disease. Participants with MS took longer to initiate their first recovery step and took larger first recovery steps than the other groups, while participants with PD took smaller first recovery steps. Group differences in performance on this clinical balance assessment were robust; the PRT was able to detect differences in balance performance.

Although there is inherent variability in the PRT, across repeated trials, the examiner was able to consistently elicit a stepping response that reflected balance performance. Future research needs to assess sensitivity of the test to changes in balance performance over time or with intervention. We are exploring instrumentation using inertial movement sensors on the trunk and ankles to achieve more objective, sensitive measures than the rating scale. Based on the results here, we will focus on algorithms to calculate step latency, length and velocity of recovery step and number of steps and time to achieve equilibrium, including a warning about latencies that are ‘too short’.

Limitations

Although we observed group differences in PRT responses, these results need to be interpreted cautiously as the design of our study included heterogenous groups to produce a wide variety of responses. Although the heterogeneity limits conclusions about group performance, it allowed us to test the characteristics of the test across a wide variety of situations, as would be observed clinically. Additionally, the use of one examiner precludes us from reporting on inter-examiner variability. Finally, the use of a harness may limit external validity as most clinics do not have a harness, and its presence in our study could have affected participants’ willingness to lean or otherwise affected their performance.

Implications for physiotherapy practice

Because variability in administration had small but significant effects on response, it is important to address whether performing the PRT three times would provide a result that is more representative of true performance than administering only one trial. We found that the first trial was not significantly different from average of the first 3 or all 10 trials for latency, indicating that administering more than 1 trial is not going to provide a more accurate result. Step length was longer for the first trial; however, it was also affected by participant height and angle of lean. We will account for this in the instrumented version by providing feedback about lean angle in real time and normalizing results by participant height. It should be noted that we allowed the participant to experience the lean (knowing there would not be a release) before administering the test. In practice, if the participant did not understand the test directions or is hesitant to lean, the examiner should administer and use a second trial. Administering a second trial is not necessary when the PRT is performed according to the directions, but a second trial may be useful in cases where the administration directions are not successfully adhered to on the first trial.

Acknowledgments

We thank the participants, especially for being willing to perform many repetitions of the push and release test! We also thank Ryan Meyer and Arash Salarian for help with data collection and analysis.

Financial support: Dr. Smith's salary was supported sequentially by grants from the National Institutes of Health (NICHD F32 HD070796 to BAS), the Foundation for Physical Therapy (NIFTI to BAS) and National Institutes of Health grant K12-HD055929 (PI — Ottenbacher). Additional support came from the National Institutes of Health (NIA R37 A60006457 to FBH).

References

Burleigh-Jacobs A, Horak FB, Nutt JG, Obeso JA. Step initiation in Parkinson's disease: influence of levodopa and external sensory triggers. Movement Disorders. 1997; 12(2):206–215. [PubMed: 9087979]

Cusick A, Vasquez M, Knowles L, Wallen M. Effect of rater training on reliability of Melbourne Assessment of Unilateral Upper Limb Function scores. Developmental Medicine and Child Neurology. 2005; 47(01):39–47. [PubMed: 15686288]

Fahn, S.; Elton, R. Unified Parkinson's disease rating scale. In: Fahn, S.; Marsden, D.; Calne, D., editors. Recent Developments in Parkinson Diseases. Florham Park, NJ: London MacMillan; 1987. p. 153-163.

Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. Physical Therapy. 2009; 89(5):484–498. [PubMed: 19329772]

Jacobs JV, Horak FB, Van Tran K, Nutt JG. An alternative clinical postural stability test for patients with Parkinson's disease. Journal of Neurology. 2006; 253(11):1404–1413. [PubMed: 16788773]

MacKinnon CD, Bissig D, Chiusano J, Miller E, Rudnick L, Jager C, Zhang Y, Mille M, Rogers M. Preparation of anticipatory postural adjustments prior to stepping. Journal of Neurophysiology. 2007; 97(6):4368–4379. [PubMed: 17460098]

Spence Laschinger HK. Intraclass correlations as estimates of interrater reliability in nursing research. Western Journal of Nursing Research. 1992; 14(2):246–251. [PubMed: 1561790]

Vaughan, CL.; Davis, B.; O'Connor, J. Dynamics of human gait. Champaign, IL: Human Kinetics; 1992.

Vincent, W. *S*tatistics in Kinesiology. 2. Champaign, IL: Human Kinetics; 2004.

Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength and Conditioning Research. 2005; 19(1):231–240. [PubMed: 15705040]

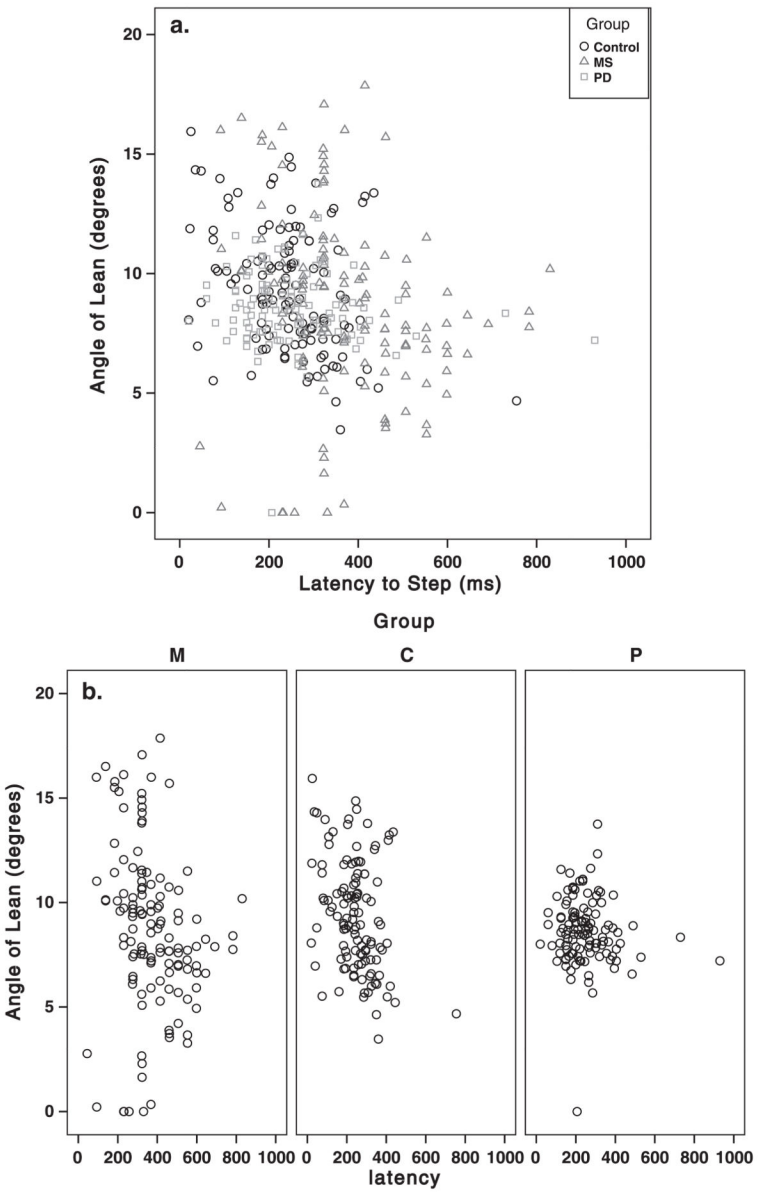

Figure 1.

Figure 1 shows that greater angle of trunk lean with administration was weakly correlated with a shorter latency to step initiation for all participants overall ($r = 0.25$, $p < 0.01$). Control, multiple sclerosis (MS) and Parkinson disease (PD) groups are represented by different shaped markers in Figure 1a and shown by group Figure 1b. Each marker indicates a separate trial. Overall, there was not a significant group difference in angle of lean. There was a significant group difference in the latency from release to step initiation ($F[2,362] = 38.41$, $p < 0.01$); participants with MS took longer to step than participants with PD or control participants

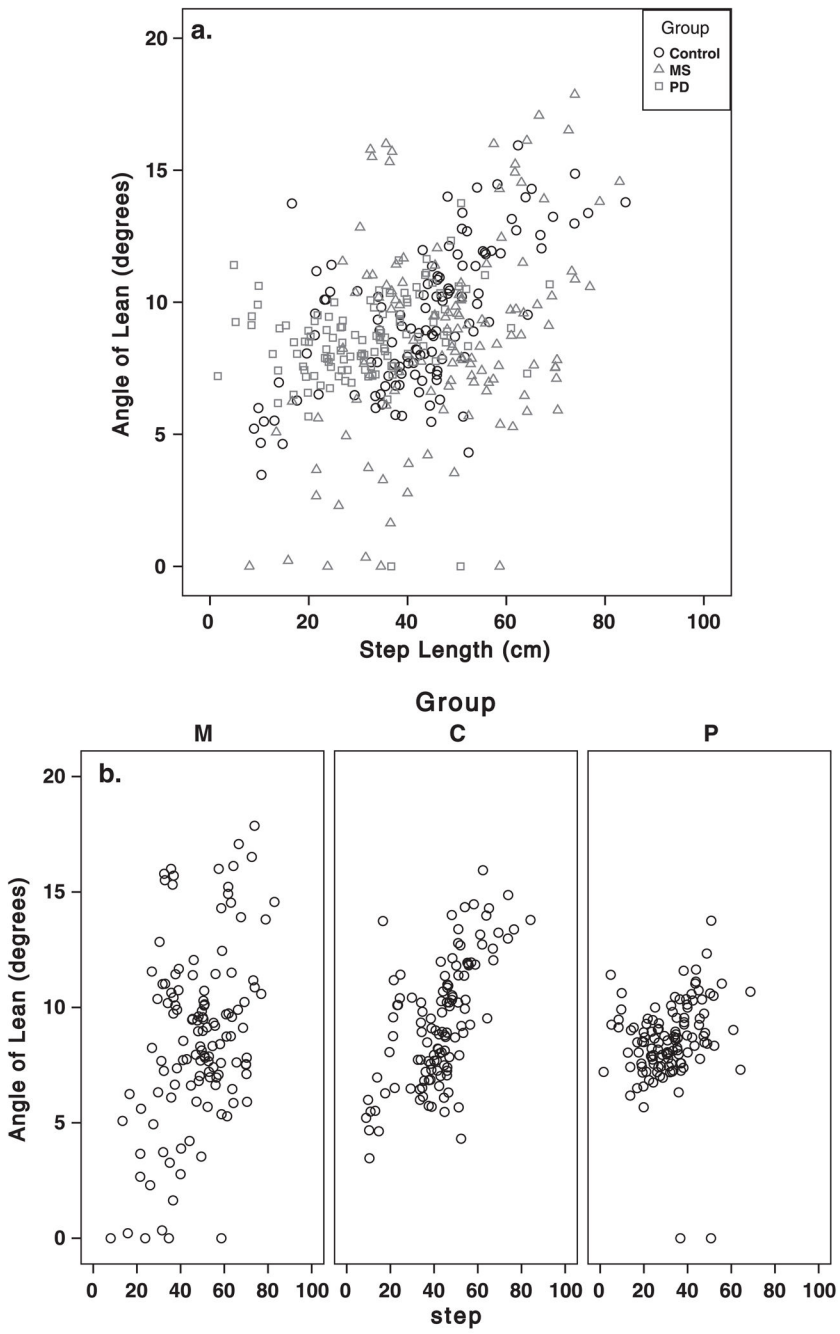

Figure 2.

The figure shows that greater angle of trunk lean with administration was weakly correlated with a larger step length of initial recovery step ($r = 0.41$, $p < 0.01$). Control, multiple sclerosis (MS) and Parkinson disease (PD) groups are represented by different shaped markers in Figure 2a and shown by group Figure 2b. Each marker indicates a separate trial. Overall, there was not a significant group difference in angle of lean. There was a significant group difference in the step length of initial recovery step ($F[2,371] = 41.74$, $p < 0.01$); participants with MS took larger steps while participants with PD took smaller steps

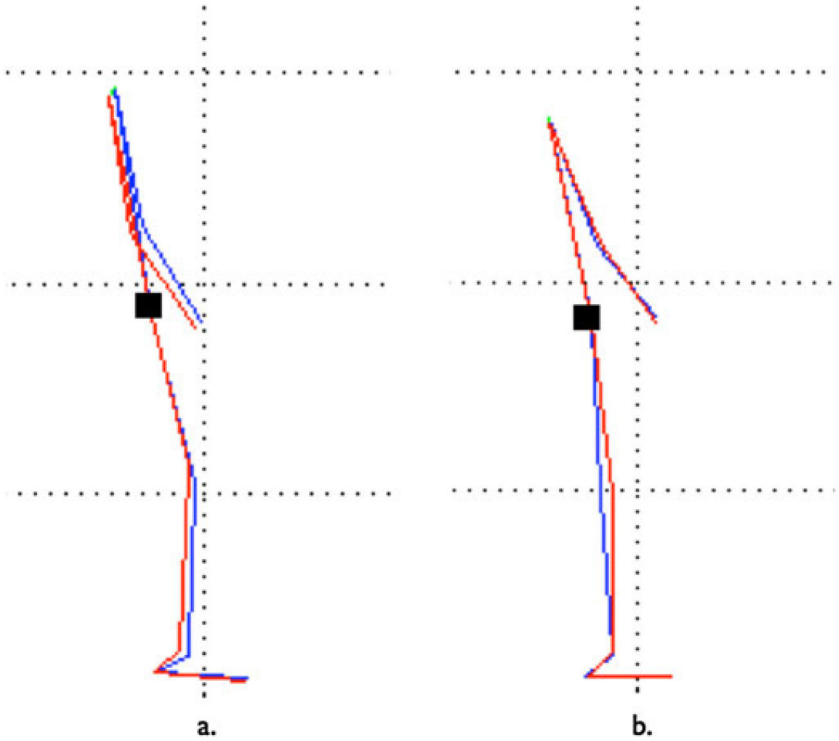

Figure 3.

Figure 3a shows a participant with Parkinson disease (PD) bending the knees during the lean, resulting in the centre of mass position (black square) being further behind the ankle than a comparable control participant in Figure 3b

Table 1

Participant characteristics

Participant	Gender	Age (years)	Height (cm)	Weight (kg)	Disease*	Disease Severity**
1	F	42	171.5	89.8	MS	4
2	F	25	162.6	101.2	MS	4.5
3	M	30	178	57	MS	4
4	F	30	165	60.8	MS	3.5
5	F	49	167.6	62.6	MS	2
6	F	28	166.5	110.7	MS	4
7	M	42	185.4	88.1	MS	4.5
8	F	28	176.5	108.9	MS	5
9	F	51	180.3	69.4	MS	4
10	F	48	162.6	64.9	MS	2.5
11	F	48	161.3	57.2	MS	4
12	F	55	166.4	120.2	MS	4
13	F	45	163.8	53.5	MS	4
14	F	45	165.1	119.3	MS	3
MS group M (SD)		40.4 (10.1)	169.5 (7.6)	83.1 (25.1)		3.8 (0.8)
15	F	67	162.6	68.9	PD-on	16
16	M	70	181	81.6	PD-off	15
17	M	73	165	68	PD-on	24
18	M	65	184.5	90.7	PD-off	30
19	M	65	175	92	PD-off	23
20	M	69	178	77.1	PD-off	37
21	M	62	180	120.2	PD-off	30
22	M	70	172.7	88.5	PD-on	30
23	M	65	185.4	89.3	PD-on	27
24	M	63	180.3	113	PD-on	29
25	F	60	160	73.5	PD-on	35
26	F	77	160	54.4	PD-on	16
PD group M (SD)		67.2 (4.9)	173.7 (9.5)	84.8 (18.6)		26 (7.3)

Participant	Gender	Age (years)	Height (cm)	Weight (kg)	Disease*	Disease Severity**
27	F	58	175	84.8	Control	
28	M	72	175	102.5	Control	
29	F	77	160	53	Control	
30	F	33	142	81.6	Control	
31	F	45	172.7	53.5	Control	
32	F	46	165.1	57	Control	
33	M	58	176.6	82.5	Control	
34	M	67	165	76.6	Control	
35	M	70	177.8	88.4	Control	
36	M	67	167	74.8	Control	
37	F	25	137	53.5	Control	
38	M	28	193	73.5	Control	
Control group *M (SD)*		53.8 (18.0)	167.2 (15.4)	73.4 (16.1)		

*Disease: MS = multiple sclerosis; PD-on = Parkinson disease, on dopamine replacement medication; PD-off = Parkinson disease, off dopamine replacement medication (last dose 12 hours before testing).

**Disease severity: for MS = extended disability status scale; for PD = motor score on Unified Parkinson's Disease Rating Scale, original version.

Physiother Res Int. Author manuscript; available in PMC 2016 September 01.

Table 2

Ten trial range of administration and response variables

		Administration variables		Performance variables		
Participant	**Disease***	**Trunk angle (degrees)**	**Centre of mass ankle (mm)**	**Latency (ms)**	**Step length (cm)**	**Clinical rating****
1	MS	5–11	77–86	277–460	61–74	3–3
2	MS	8–16	96–116	93–460	27–38	3–3
3	MS	8–12	100–108	198–418	48–69	2–2
4	MS	8–12	102–119	138–460	39–51	3–3
5	MS	8–14	101–108	275–598	46–68	3–3
6	MS	10–16	70–84	92–462	33–62	2–2
7	MS	12–18	37–54	138–553	59–83	3–3
8	MS	6–11	76–88	230–600	52–77	3–3
9	MS	7–9	25–43	415–783	27–58	2–2
10	MS	4–10	12–54	368–830	34–59	2–2
11	MS	8–10	10–18	220–368	31–70	2–3
12	MS	2–6	23–56	45–598	17–40	3–3
13	MS	4–8	14–37	368–783	37–53	3–3
14	MS	1–6	4–18	93–368	8–32	3–3
MS group M (SD)		9.0 (3.6)	62.6 (35.2)	369.6 (145.9)	48.3 (15.3)	2.7 (0.5)
15	PD-on	7–10	108–123	206–412	31–40	3–3
16	PD-off	8–12	23–47	60–250	33–51	3–3
17	PD-on	7–9	135–142	125–375	20–49	2–2
18	PD-off	10–14	104–113	160–390	39–69	2–3
19	PD-off	8–12	102–120	185–365	20–44	2–2
20	PD-off	7–11	126–152	125–425	2–25	0–0
21	PD-off	6–9	77–94	230–320	14–20	0–2
22	PD-on	6–9	75–93	105–385	19–64	3–3
23	PD-on	7–10	107–115	240–730	22–52	3–3
24	PD-on	7–10	83–93	20–270	31–38	3–3
25	PD-on	6–11	52–71	120–365	31–42	0–3
26	PD-on	7–9	124–136	80–930	17–27	1–2

		Administration variables		Performance variables		
Participant	**Disease***	**Trunk angle (degrees)**	**Centre of mass ankle (mm)**	**Latency (ms)**	**Step length (cm)**	**Clinical rating****
PD group M (SD)		8.7 (1.39)	100.7 (30.8)	251.1 (123.7)	31.8 (12.8)	2.3 (0.9)
27	Control	6–10	49–67	185–369	43–54	3–3
28	Control	3–7	104–126	40–755	9–22	3–3
29	Control	8–11	70–91	20–250	20–34	2–3
30	Control	9–12	2–8	125–355	39–51	2–3
31	Control	6–8	11–17	170–344	33–45	3–3
32	Control	6–10	34–47	152–379	29–45	3–3
33	Control	6–10	39–46	183–404	34–48	3–3
34	Control	5–10	93–105	160–352	35–64	3–3
35	Control	7–16	75–119	23–225	17–65	3–3
36	Control	8–12	87–101	200–290	38–57	3–3
37	Control	4–14	65–152	35–250	44–58	3–3
38	Control	9–15	91–113	90–435	56–84	3–3
Control group *M* (*SD*)		**9.3 (2.7)**	**68.5 (35.7)**	**243.4 (109.0)**	**42.8 (15.0)**	**2.9 (0.2)**

* Disease: MS = multiple sclerosis; PD-on = Parkinson disease, on dopamine replacement medication; PD-off = Parkinson disease, off dopamine replacement medication (last dose 12 hours before testing).

** Clinical rating: Clinical response ratings were scored using the rating scale from the BESTest: (3) recovers independently a single, large step; (2) more than one step used, but stable and recovers independently or one step with imbalance; (1) takes several steps to recover equilibrium, or needs minimum assistance; (0) no step, or would fall if not caught, or falls spontaneously. (Horak *et al*., 2009).

Table 3

Intraclass correlation coefficients (ICCs) and 95% confidence intervals (CIs)

	ICC	95% CI
Administration variables		
Trunk angle	0.71	0.61–0.81
Centre of mass ankle	0.95	0.92–0.97
Response variables		
Latency	0.48	0.35–0.62
Latencies ≥ 150 ms*	0.43	0.29–0.60
Step length	0.78	0.70–0.86

*We re-ran the ICC excluding latency values less than 150 ms, as an instrumented version of the test would contain a warning to re-do these trials as this is faster than possible for evoked stepping postural response latencies (Burleigh-Jacobs *et al*., 1997; MacKinnon *et al*., 2007). The ICC excluding latency values less than 150 ms was calculated from five values per participant.