J Neurol (2006) 253 : 1404–1413
DOI 10.1007/s00415-006-0224-x

ORIGINAL COMMUNICATION

Jesse V. Jacobs
Fay B. Horak
Van K. Tran
John G. Nutt

An alternative clinical postural stability test for patients with Parkinson's disease

Received: 17 October 2005
Received in revised form: 2 March 2006
Accepted: 10 March 2006
Published online: 20 June 2006

J.V. Jacobs, BA ·
F.B. Horak, PT. PhD (✉)
Neurological Sciences Institute
Oregon Health & Science University
505 NW 185th Avenue
Beaverton, OR 97006-3499, USA
Tel.: +1-503/418-2600
Fax: +1-503/418-2501
E-Mail: HorakF@ohsu.edu

V.K. Tran, MD · J.G. Nutt, MD ·
F.B. Horak, PT. PhD
Dept. of Neurology
Oregon Health & Science University
3181 S.W. Sam Jackson Park Road
Portland, Oregon 97239, USA

F.B. Horak, PT. PhD · J.G. Nutt, MD
Dept. of Physiology & Pharmacology
Oregon Health & Science University
3181 S.W. Sam Jackson Park Road
Portland, Oregon 97239, USA

J. G. Nutt, MD
Portland VAMC Parkinson Disease
Research
Education and Clinical Center
3710 SW Veterans Hospital Road
Portland, OR 97239, USA

Abstract We compared the sensitivity and consistency of a new Push and Release Test versus the Pull Test (item 30 of the Unified Parkinson's Disease Rating Scale; UPDRS) as clinical measures of postural stability. Subjects with Parkinson's disease and age-matched control subjects participated in 3 protocols investigating: (1) the sensitivity and specificity of the two tests related to the subjects' balance confidence, as measured by the Activities-specific Balance Confidence (ABC) scale, (2) the inter-rater reliability of the two tests, and (3) the consistency of the perturbation forces applied to the subjects by each balance test. As a test for concurrent validity, the balance tests were also compared with the subjects' retrospective reports of fall frequency. Compared with the Pull Test, the Push and Release Test was more sensitive to subjects with low balance confidence, but less specific for subjects with high balance confidence. The inter-rater correlations were higher with the Push and Release Test. Examiners applied more consistent perturbation forces to the subjects with the Push and Release Test than with the Pull Test. The Push and Release Test correlated better with self-reported falls. Therefore, the Push and Release Test provided a more sensitive and consistent test of postural stability than the Pull Test.

Key words Parkinson's disease · posture · balance · UPDRS · push and release test

Introduction

Evaluation of postural responses in standing balance is a key component to the neurological examination of patients with idiopathic Parkinson's disease (PD). The presence of abnormal postural responses separates milder PD (Hoehn & Yahr stages 1 and 2) from more severe and disabling PD (Hoehn & Yahr stages 3 to 5). Furthermore, impaired balance is often an indication for more aggressive therapy because of the risk of falls and their medical, psychological, and social consequences. Therefore, it is important for

JON 2224

clinicians to have accurate and sensitive tools to evaluate postural instability in PD.

The 'Retropulsion Test' or 'Pull Test' (Postural Stability Item #30 of the Unified Parkinson's Disease Rating Scale; UPDRS [5]) is a commonly used clinical test of postural stability for patients with PD. This test evaluates the ability of patients to recover from a backward pull on the shoulders. Unfortunately, the Pull Test is not sensitive to the early detection of fallers [2], and we find the Pull Test is often normal when patients report their balance to be abnormal. Although the Pull Test may be insensitive to a PD patient's fall history and low balance confidence because factors other than postural instability result in falls and low balance confidence, the Pull Test's insensitivity may also be due to limitations in the test itself. For example, when performing the Pull Test, it has been reported that (1) subjects can brace themselves to anticipate the perturbation [10], (2) examiners do not properly administer the test [10], and (3) the rating scale is insensitive [14]. Therefore, although some limitations of the Pull Test have been addressed [10, 14], an alternative test may better fill the need for a reliable clinical assessment of postural stability that is sensitive to a PD patient's fall history and low balance confidence.

To address the need for an improved clinical test of postural stability, we developed the Push and Release Test as an alternative to the Pull Test [8]. The Push and Release Test rates the postural response to a sudden release of a subject pressing backward on an examiner's hands placed on the subject's back. In addition to the new test maneuver, we developed a new rating scale that differentiates less severe abnormalities in balance than the rating scale used for the Pull Test in the UPDRS (i.e., differentiating mild impairment from moderate impairment by the number of steps needed to recover equilibrium). The new rating scale also provides 5 scores for rating a subject's response to the Push and Release Test, whereas the UPDRS rating scale only provides 3 scores that actually use the Pull Test because the final 2 scores represent a tendency for the subject to spontaneously drift out of balance or an inability to stand unassisted.

We had three hypotheses. First, we hypothesized that the Push and Release Test would correlate better with the subjects' perception of their balance, as determined by comparisons with the Activities-Specific Balance Confidence (ABC) Scale [13]. We chose the ABC Scale as our primary comparison because: (1) as yet, no "gold-standard" measure exists for testing postural stability, (2) the ABC scale provides correlations to balance impairment in PD and to the patients' mobility outside the clinic [1, 11, 13], and (3) a subject's balance confidence would likely coincide with his or her tendency to complain of a balance problem in the clinic.

Second, we hypothesized that the Push and Release Test would be more informative than the Pull Test because: (1) the forces applied to the subjects during the Push and Release Test would be less variable, (2) releasing subjects from a predetermined, unstable posture during the Push and Release Test removes the duration of the perturbation as a source of variability and prevents subjects from bracing in anticipation of the perturbation, and (3) the score values for the Push and Release rating scale explicitly differentiate mild impairment from moderate impairment by the actual number of steps taken in response to the perturbation.

Third, assuming postural instability represents a primary factor for inducing falls in PD patients, we hypothesized that the Push and Release Test would present a valid test of postural stability, as determined by correlations with self-reported falls.

In practice, the Pull Test and the Push and Release Test are administered several times in succession, allowing the examiner to determine the amount of force necessary to make the subjects take a corrective step. The repetition of the tests, however, simultaneously allows the subjects to adapt their responses. For this reason, we compared the first and third trial of each balance test to investigate which trial clinicians should use [2, 12].

Methods

We compared the Push and Release Test and the Pull Test:

1) *The Push and Release Test*. Subjects stood in a comfortable stance with their eyes open. Examiners stood behind the subjects, and subjects were instructed to do whatever necessary, including taking a step, to regain their balance. The subjects were then instructed to push backward against the palms of the examiners' hands placed on the subjects' scapulae while the examiners flexed their elbows to allow backward movement of the trunk while they supported the subjects' weight with their hands. The force of the subjects' push was not so strong as to cause their heels to come off the ground, and subjects were not permitted to passively lean back into the examiners' hands. When the subjects' shoulders and hips moved to a stable position just behind their heels, the examiners suddenly removed their hands, requiring the subjects to take a backward step to regain balance (Figure 1). After the subject reached the proper position, the time at which the examiner released his or her hands from the subject varied in order to ensure that the subject could not anticipate the release. We found that the position of the subjects' hips relative to their heels could be judged by both visual inspection and by paying attention to the force against the examiner's hands: once the proper position has been reached, the pressure on the examiner's hands stabilizes because the subject becomes dependent on the examiner for support against gravity. The subjects had to take a step for the test to be properly executed. The examiners also acted as attendants to prevent the subjects from falling if their responses were inadequate to retain

Fig. 1 (A) Photo depicting the Push and Release Test. Notice the subject's hips and shoulders are just behind the subject's heels as the subject actively pushes back into the examiner's hands without the subject's heels lifting off the ground

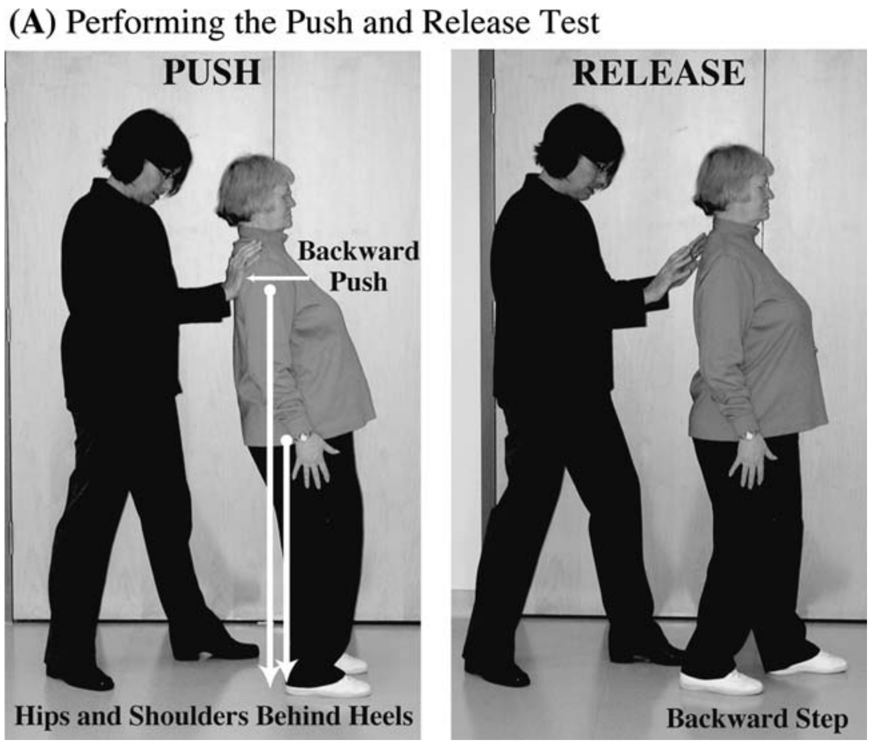

Table 1 Balance Test Rating Scales

Score	Push and Release Rating Criteria	UPDRS Rating Criteria
0	Recovers independently with 1 step of normal length and width	Normal
1	Two to three small steps backward, but recovers independently	Retropulsion, but recovers unaided
2	Four or more steps backward, but recovers independently	Absence of postural response, would fall if not caught by examiner
3	Steps but needs to be assisted to prevent a fall	Very unstable, tends to lose balance spontaneously
4	Falls without attempting a step or unable to stand without assistance	Unable to stand without assistance

balance. The subjects were rated on the 0- to 4-point Push and Release scale (Table 1). The values for the Push and Release Test's rating scale were defined by the actual number of steps a subject took in response to the release. A step was counted only if it was required for the subjects to maintain their balance. Steps taken to reorient the feet side by side were not counted.

2) *The Pull Test*. The subjects stood with their eyes open in a comfortable stance (but were instructed to change their stance to place the feet approximately shoulder length apart if they assumed an unusually narrow or wide stance), and the examiners stood behind the subjects. The examiners ensured that the subjects were standing upright and not leaning forward in anticipation of the pull. The subjects were prepared and instructed to do whatever necessary, including taking a step, to keep their balance. After giving several gentle pushes to the shoulders laterally, the examiners gave a sudden, brief, backward pull to the subjects' shoulders with sufficient force to cause the subjects to step to regain balance. The pull was considered inadequate if the subject did not step to regain balance in order to ensure low scores were not biased toward larger or stronger individuals. The examiners also served as attendants to catch the subjects if their responses were inadequate to maintain balance. The examiners scored responses on a 0- to 4-point scale, using both the scoring employed by the UPDRS for this test and the scoring for the new Push and Release Test (Table 1). When rating according to the UPDRS scale, retropulsion was defined as a response of more than two steps [10, 14].

We conducted three experiments to compare the Push and Release Test and the Pull Test. Experiment 1 examined the tests' relative sensitivity to balance confidence, measured with the ABC scale, and the tests' relative ability to identify patients with a history of falls. Experiment 2 examined the inter-rater reliability of the tests' scores among independent examinations of three raters (as opposed to the raters simultaneously scoring the same examination). Experiment 3 examined the consistency of the perturbation forces applied to the subjects for each balance test. All subjects gave informed consent to a protocol approved by the Institutional Review Board of Oregon Health & Science University (OHSU) in accordance with the Declaration of Helsinki.

Experiment one: test sensitivity, specificity, and validity

Subjects

Eighty-eight subjects with PD (27 females, 61 males) were tested during a visit to the Movement Disorders Clinic at OHSU. Subjects with PD were tested regardless of their medication dose cycle. Subjects were screened for the existence of other medical conditions that might affect balance. Consequently, of the 88 total subjects screened, five subjects were excluded from analysis because of other health problems (i.e., back injury, peripheral neuropathy, and arthritis) or due to atypical Parkinsonism. Fourteen other PD subjects were also excluded from analysis because, after giving informed consent, they chose not to complete the balance tests. Comparisons between the balance tests, therefore, include data from the remaining 48 male and 21 female subjects with PD. These subjects had a mean age (± the standard deviation) of 67 ± 12 years (range = 42–88 years), a 10 ± 6 -year duration of PD (range = 2–25 years), and a mean UPDRS motor score of 24 ± 12 points (range = 3–51 points). Of the 69 subjects that completed the balance tests, 63 subjects completed the ABC questionnaire.

To compare balance test responses among subjects with and without PD, 75 age-matched control subjects without PD were also tested. Subjects included spouses of the PD subjects during their visit to OHSU's clinic or subjects recruited at a local fitness center. Although the control subjects participated in an exercise program, we do not believe they were biased for better balance than the general population based on their performance on the balance tests (see results demonstrating that only 10 of the control subjects exhibited high levels of performance on all balance tests). All subjects were screened, but not excluded, for the existence of other medical conditions (e.g., hip and knee replacements, diabetes, stroke, laminectomy, heart surgery, hearing impairments, corrected vision impairments, or the use of multiple medications). We chose not to exclude the 18 control subjects who reported having medical conditions to preserve the environmental validity of the score distributions in an elderly population. Two control subjects were excluded from analysis, however, because (after giving informed consent) they chose not to complete the balance tests, and 4 additional control subjects were excluded because they were too young and did not match the age of our subjects with PD. Comparisons between the balance tests, therefore, include data from the remaining 35 male and 34 female control subjects without PD. These subjects had a mean age of 67 ± 10 years (range = 41–84 years). Of these 69 subjects, 61 subjects completed the ABC questionnaire.

Elderly subjects without PD may also have balance impairments, and the two balance tests may be differentially sensitive to these impairments. Therefore, to ensure that the Push and Release Test's hypothesized increase in sensitivity was not simply because control subjects were unable to demonstrate the Push and Release Test's definition of a normal, healthy response, 10 control subjects were classified as Elite subjects. Elite subjects were classified according to the following criteria: (1) subjects did not report any health complications, (2) subjects could stand on one leg for 30 seconds [6], (3) subjects could reach farther than 30 cm on a Functional Reach Test [4], (4) subjects had an ABC score higher than 90% confidence, and (5) subjects responded normally (score = 0) on every trial of the Pull Test.

Protocol

The Movement Disorder Clinic's medical staff acquired the PD subjects' demographic information and medical history. For control subjects, members of OHSU's Balance Disorders Laboratory acquired the subjects' demographic information and medical history. All subjects reported the number of falls they experienced over the last year. A neurologist performed the UPDRS test on subjects with PD. Subsequently, one of three physical therapists, a neurology research fellow, or a neuroscience graduate student administered the ABC scale questionnaire and rated the subjects' performance on three trials for each of the balance tests. The balance tests were blocked and examiners randomly chose the test order. Some of the examiners were not blind to the purpose of the study, potentially biasing our results. Such a bias is not likely, however, because the instructions, execution, and rating of the balance tests were clear, detailed, and explicit in order to minimize any subjective interpretations of the subjects' responses.

The ABC Scale was administered just prior to balance testing [13]. The rater administered the 16-item questionnaire by asking the subject, "On a scale of zero percent representing no confidence to 100 percent representing complete confidence, how confident would you be to...(ABC item stated)?" Although the subjects answered the ABC questionnaire under any medication state, during the questionnaire, the subjects were asked to rate their confidence as though they were performing the items of the questionnaire while in their average medication state (i.e., not during "OFF" periods or during periods of peak dosage). Each item's percentage was recorded, and the subject's average ABC-scale score was obtained by averaging the 16 sub-scores.

Analysis

Sensitivity related to balance confidence was defined as the percentage of subjects scored by the Push and Release Test or the Pull Test as having abnormal responses and reporting low balance confidence on the ABC scale to the total number of subjects reporting low balance confidence. Specificity was defined as the percentage of subjects with normal responses to the balance tests and reporting high balance confidence to the total number of subjects reporting high balance confidence. For either test's rating scale, a normal response was a zero-score and an abnormal response was any score above zero. We chose this response classification because the nominal scale erased any differences due to applying different rating scales to the Push and Release Test and the Pull Test. Low balance confidence was defined as an ABC-scale score of less than 80% confidence, and high balance confidence was defined as a score of 80% or greater confidence. The 80% cutoff score was chosen because lower cutoff scores degraded both of the balance tests' agreement with balance confidence, and a higher cutoff score risked including a healthy range of scores [13].

To determine whether subjects with good balance (defined by our Elite status) had a normal response to the Push and Release Test, a McNemar's test compared the Elite subjects' actual responses to the Push and Release Test with hypothetical responses that Elite subjects would always respond with a normal balance response.

Two-tailed Spearman's rho correlation coefficients compared each balance test with self-reported falls to establish concurrent validity. In addition, because retrospective fall reports may be unreliable [3], we attempted to minimize this problem by categorizing subjects as fallers (subjects who reported one or more falls in the past year) or non-fallers (subjects who reported that they had not fallen in the past year). Sensitivity related to falling was defined as the percentage of fallers with abnormal responses to the balance tests. Specificity was defined as the percentage of non-fallers with normal responses to the balance tests.

A Friedman's repeated measures ANOVA examined (1) differences among repeated trials of the Push and Release Test and the Pull Test, and (2) differences among the Push and Release Test scores and Pull Test scores. A Kruskal-Wallis Test examined differences in balance test scores or questionnaire scores among control subjects and subjects with PD. We reported averages as means ± the standard deviations (SD), and medians with the score range. Significance was defined by a p-value of less than 0.05.

Experiment Two: Inter-Rater Reliability Among Independent Examinations

Subjects

Three healthy subjects and eight subjects with PD were tested to examine the inter-rater reliability of the Push and Release Test and the Pull Test. Healthy subjects were two males and one female who were 62 ± 2 years of age (range = 61–64 years). Subjects with PD included seven males and one female who were 62 ± 7 years of age (range = 55–74 years).

Protocol

On the same day, two physical therapists and a neurology research fellow each separately tested and rated the 11 subjects on the Push and Release Test and the Pull Test. Subjects were tested by all three examiners, on average, within 31 ± 5 minutes. Subjects performed three trials of each balance test for each examiner. Tests were blocked and the examiners randomly chose the test order.

Analysis

We report inter-rater reliability with 2-way, mixed effect model, intra-class correlation coefficients (ICCs), with the absolute agreement definition (SPSS Inc., Chicago, Illinois).

Experiment Three: Consistency of Perturbation Forces Applied to the Subjects in the Push and Release Test and the Pull Test

Subjects

Eleven elderly women without PD (76 ± 4 years of age; range = 68–80) were tested to examine the consistency of the perturbation forces applied to the subjects in the Push and Release Test and the Pull Test. Subjects were screened, but not excluded, for health complications (e.g., knee and ankle replacements, diabetes, neuropathy, hearing impairments, or corrected vision impairments). We chose not to exclude subjects with health complications to ensure a wide distribution of balance test scores.

Protocol

A physical therapist and a neuroscience graduate student examined every subject 4 times with each balance test (for a total of 8 trials for each test per subject). Examiners performed the four trials of each balance test in sequence, although the test order and the order of examiners were randomized between subjects.

To measure the consistency of forces applied to a subject's torso for the Push and Release Test and the Pull Test, examiners wore sporting gloves fitted with twelve, 444 N-range Flexiforce pressure sensors on each glove (Tekscan Inc., South Boston, MA, USA). Paired sensors were placed on the distal pads of the index, middle, and ring fingers and at 3 locations around the palm of the glove (Figure 2). The signals from all sensors for both gloves were summed to determine the total force applied to the subjects.

Analysis

We analyzed (1) the peak total force applied during the Pull Test (Pull Force), (2) the peak total force applied just prior to the release of the Push and Release Test (Push Force), and (3) the duration of the pull during the Pull Test. We calculated the range of Pull Forces and Push Forces applied to each subject over four trials by each examiner. A repeated-measures ANOVA, including the 2-level factor of TEST and the 2-level factor of EXAMINER, determined differences among the range of Pull Forces and the range of Push Forces applied to each subject for each examiner. A paired t-test determined differences in pull duration between the two examiners.

(A) Glove with Pressure Sensors

Fig. 2 (A) Photo of the sporting glove fitted with 444 N-range Flexiforce pressure sensors (Tekscan Inc., South Boston, MA, USA) used to quantify the forces applied to subjects during the Push and Release Test and the Pull Test.

Results

Balance Test Sensitivity and Validity

The Push and Release Test elicited higher scores than the Pull Test (UPDRS scale), regardless of trial or group ($p < 0.01$, Table 2). Applying the Push and Release rating scale to the Pull Test also increased scores ($p < 0.01$, Table 2). Compared with control subjects, subjects with PD exhibited higher scores on both balance tests, and subjects with PD also reported lower ABC scores than control subjects ($p < 0.001$ for all comparisons, Figure 3A).

Compared with the Pull Test, the Push and Release Test was more sensitive to low balance confidence, but less specific to high balance confidence (Figure 3B). Further analysis revealed 15 control subjects

Table 2 Descriptive Statistics for each Balance Test by Group and Trial

Balance Test	First Trial Mean ± SD; Median (range)	Third Trial Mean ± SD; Median (range)
Subjects with PD (n = 69)		
Push & Release Test	2.10 ± 1.16; 2 (0–4)	1.66 ± 1.23; 2 (0–4)
Pull Test	1.24 ± 0.89; 1 (0–4)	0.85 ± 0.89; 0 (0–4)
Pull Test (P&R Scale)	1.76 ± 1.27; 1 (0–4)	1.09 ± 1.19; 1 (0–4)
Control Subjects (n = 69)		
Push & Release Test	0.63 ± 0.74; 0 (0–3)	0.50 ± 0.73; 0 (0–3)
Pull Test	0.42 ± 0.53; 0 (0–2)	0.17 ± 0.38; 0 (0–1)
Pull Test (P&R Scale)	0.65 ± 0.78; 0 (0–3)	0.33 ± 0.59; 0 (0–2)

had an abnormal Push and Release response despite normal Pull Test responses and high balance confidence. Compared with 33 control subjects who demonstrated high confidence and normal responses on both balance tests, these 15 control subjects with abnormal Push and Release responses also demonstrated shorter Functional Reach lengths (35 ± 8 cm compared with 40 ± 7 cm, $p < 0.05$) and tended to have shorter one-leg stance durations (13 ± 10 s compared with 19 ± 11 s, $p = 0.076$). The 10 Elite subjects always demonstrated a normal balance response in the Push and Release Test.

All of the control subjects reported that they had not fallen in the past year. For subjects with PD, 34 of 68 subjects reported that they had fallen in the past year. The Push and Release Test elicited higher correlations than the Pull Test with self-reported falls: Spearman's rho = 0.604 ($p < 0.000001$) on the first trial and 0.553 ($p < 0.00001$) on the third trial of the Push and Release Test, whereas Spearman's rho = 0.375 ($p < 0.01$) on the first trial and 0.451 ($p < 0.001$) on the third trial of the Pull Test. Applying the Push & Release rating scale to the Pull Test did not increase correlations with self-reported falls to the same degree as when performing the Push and Release Test: Spearman's rho = 0.443 on the first trial and 0.422 on the third trial of the Pull Test when rated by the Push and Release scale ($p < 0.001$). Compared with the Pull Test, the Push and Release Test was more sensitive for identifying fallers, but less specific for identifying non-fallers (Table 3).

Inter-Rater Reliability

When 3 examiners independently tested and rated subjects on 3 trials of each balance test, the Push & Release Test elicited higher ICCs than the Pull Test: The Push and Release Test's ICC = 0.84 on the first trial and 0.83 on the third trial, compared with the Pull Test's ICC = 0.45 on the first trial and 0.74 on the third trial. Applying the Push and Release rating scale to the Pull Test increased the ICC from 0.45 to 0.75 on the first trial, and increased the ICC from 0.74 to 0.84 on the third trial.

Consistency of Perturbation Forces

The range of Pull Forces applied to the subjects was greater than the range of Push Forces applied to the subjects ($p < 0.01$), and there was no significant difference in the range of forces applied by each examiner (Figure 4A and 4B). The duration of the pull during the Pull Test differed among examiners ($p < 0.0001$; Figure 4C).

Discussion

Impaired balance and the consequences of impaired balance are multi-factorial, and there is no single measure of balance that can serve as the "gold standard". We used the subjects' estimates of their balance, measured with the ABC scale, and their fall history as the best indices of the functional impact of impaired balance. Although this study would benefit from a prospective assessment of each balance test's ability to predict falls, the purpose of this study was to compare each test's ability to identify existing impairments of postural stability.

The Push and Release Test was consistently more sensitive to subjects with low balance confidence over repeated trials of the test, whereas the Pull Test's sensitivity decreased with trial repetition. The difference in Pull Test sensitivity with trial repetition may be because (1) subjects were able to prepare or brace themselves for a repeated trial of the Pull Test [10], but were unable to change their initial posture when repeating the Push and Release Test, and (2) examiners may have adjusted the strength of their pull during subsequent trials of the Pull Test, but the perturbation forces during the Push and Release Test were relatively consistent from trial to trial.

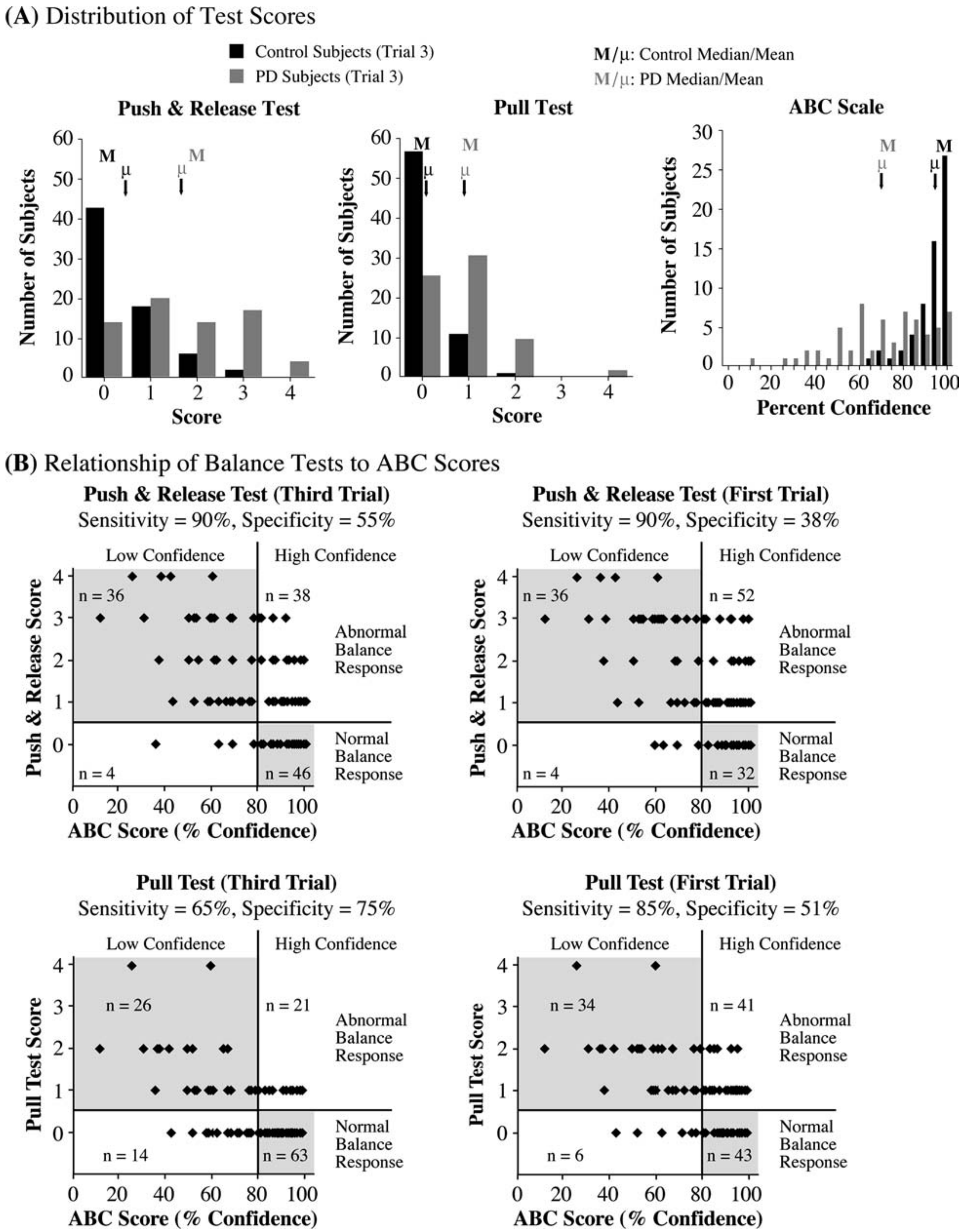
(A) Distribution of Test Scores
Control Subjects (Trial 3)
PD Subjects (Trial 3)
M/μ: Control Median/Mean
M/μ: PD Median/Mean
Push & Release Test
Pull Test
ABC Scale
Number of Subjects
Score
Percent Confidence
(B) Relationship of Balance Tests to ABC Scores
Push & Release Test (Third Trial)
Sensitivity = 90%, Specificity = 55%
Push & Release Test (First Trial)
Sensitivity = 90%, Specificity = 38%
Pull Test (Third Trial)
Sensitivity = 65%, Specificity = 75%
Pull Test (First Trial)
Sensitivity = 85%, Specificity = 51%
Low Confidence
High Confidence
Abnormal Balance Response
Normal Balance Response
Push & Release Score
Pull Test Score
ABC Score (% Confidence)
n = 36
n = 38
n = 4
n = 46
n = 52
n = 32
n = 26
n = 21
n = 14
n = 63
n = 34
n = 41
n = 6
n = 43

Table 3 The Sensitivity and Specificity of the Push and Release Test and the Pull Test to Falls

Trial	Fall History	Push & Release Test		Pull Test	
		Abnormal Response	Normal Response	Abnormal Response	Normal Response
Trial 3		Sensitivity 91%;	Specificity 32%	Sensitivity 76%;	Specificity 50%
	Fallers	31	3	26	8
	Non-Fallers	23	11	17	17
Trial 1		Sensitivity 100%;	Specificity 18%	Sensitivity 85%;	Specificity 26%
	Fallers	34	0	29	5
	Non-Fallers	28	6	25	9

Compared with the Pull Test, the Push and Release Test was less specific to the subjects with high balance confidence. The low specificity of the Push and Release Test may be because a subset of subjects were over-confident in their ability to perform balance tasks, and the Push and Release Test, unlike the Pull Test, was capable of identifying those subjects. Compared with highly-confident subjects who demonstrated normal responses to both the Pull Test and the Push and Release Test, highly-confident subjects with abnormal responses to the Push and Release Test but normal responses to the Pull Test demonstrated lower Functional Reach lengths and one-leg stance times, suggesting that the Push and Release Test identified overly-confident subjects with evidence of impaired balance that the Pull Test did not identify.

A strong association exists among falling and balance confidence [7, 9]. According to Friedman et al. [7], subjects who had low balance confidence at baseline were nearly twice as likely to fall within the next 20 months, and subjects who had fallen at baseline were nearly twice as likely to report low balance confidence 20 months later. Because every control subject in the current study reported not to have fallen in the past year, some of the control subjects may have remained confident in their balance despite diminishing balance competence. Thus, the low specificity of the Push and Release Test does not undermine its usefulness in the clinic, because the Push and Release Test may help identify patients with balance impairments before they experience a fall, thereby providing an opportunity to institute therapies before patients become injured or disabled by recurrent falls. A long-term prospective assessment, however, would be necessary to examine this hypothesis.

Compared with the Pull Test, the Push and Release Test provided a stronger correlation with falling. For both balance tests, however, the correlations were only moderate, probably because factors other than postural instability (such as freezing and dyskinesia) contribute to falls. In addition to a stronger correlation to the subjects' fall history, the Push and Release Test was more sensitive to subjects who had fallen but less specific for subjects who had not fallen. Due to the potential consequences of missing a patient that has postural instability, the Push and Release Test's low specificity to non-fallers with high balance confidence presents a relatively minor problem compared with the Pull Test's low sensitivity to subjects with low balance confidence, because the goal of the test is to identify postural instability early and in concurrence with patient complaints. Thus, sensitivity is preferred over specificity [2].

The Push and Release Test was more consistent across trials and raters than the Pull Test for examining postural stability. The Push and Release Test may have been more consistent because (1) the magnitude of the perturbation to the subjects was less variable, (2) examiners could take time to control the perturbation of the Push and Release Test, whereas the duration of the perturbation was a source of variability with the Pull Test, and (3) the functional definitions provided by the Push and Release rating scale improved inter-rater reliability.

Both a novel trial and a repeated trial of the Pull Test presented significant limitations: although the first trial of the Pull Test was more sensitive than the third trial to subjects with low balance confidence, and a novel trial was generally considered to be the most ecologically valid [2], the third trial of the Pull Test may actually be preferred because (1) the force of the perturbation elicited by the first trial of the Pull Test may not be appropriate, (2) associations with falls increased from the first to the third trial of the

Fig. 3 (A) Frequency histograms illustrating score differences among subjects with and without PD for the third trial of the Push and Release Test, the third trial of the Pull Test, and the ABC questionnaire. The horizontal position of the "M" represents the median score, and the horizontal position of the "μ" represents the mean score. **(B)** Scatter plots demonstrating the combination of specificity and sensitivity afforded by the first and third trial of the Push and Release Test and the Pull Test. The diamonds represent individual subjects. The vertical crossbars separate subjects who reported low balance confidence (ABC score < 80%) from those who reported high balance confidence (ABC score ≥ 80%). The horizontal crossbars separate subjects with normal responses to the balance tests (score = 0) from those with abnormal responses to the balance tests (score > 0). The diamonds in the shaded regions represent subjects whose balance confidence was correctly classified by their responses to the balance tests

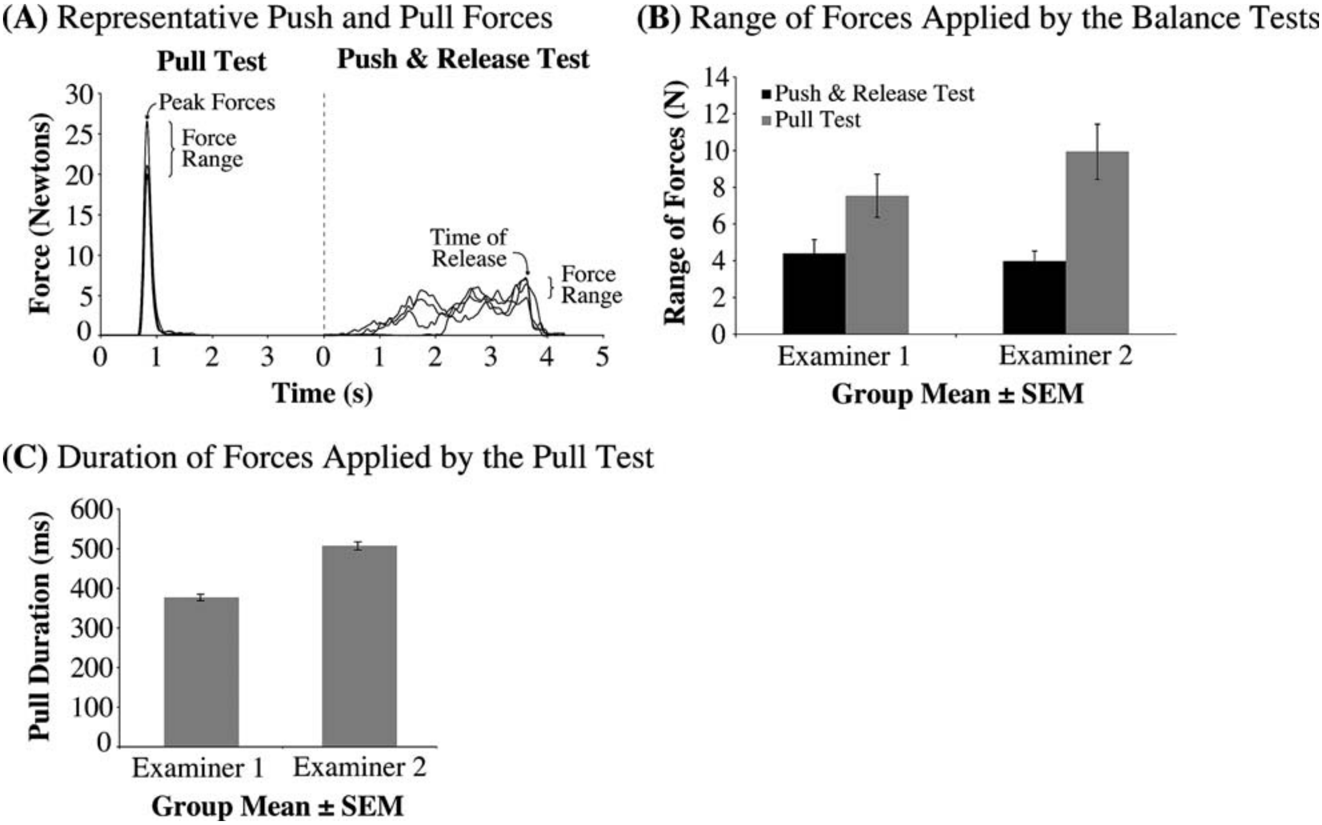

Fig. 4 **(A)** Representative traces of the total force applied to the gloves during the Pull Test and the Push and Release Test from 4 trials performed by one subject. **(B)** Bar graph demonstrating the difference between the range of the Pull Forces and the range of the Push Forces. **(C)** Bar graph demonstrating the difference among examiners in the duration of the pull during the Pull Test

Pull Test, and (3) inter-rater reliability improved from the first to the third trial of the Pull Test.

Performing a repeated trial of the Push and Release Test, however, provided a better assessment of postural stability than either trial of the Pull Test, regardless of rating scales, because (1) the Push and Release test provided a high level of sensitivity and consistency, (2) the Push and Release Test correlated best with falls, and (3) the Push and Release Test's low specificity to balance confidence may actually represent an ability to identify subjects with a minor balance impairment before the subjects recognize their deteriorating balance. Therefore, the third trial of the Push and Release Test provides a valid, more consistent, and more sensitive alternative to the Pull Test when assessing postural stability in subjects with PD.

Despite the advantages of the Push and Release Test, some limitations became evident. First, some subjects were hesitant to push back into the examiner's hands. Thus, it may be difficult to execute the test on all subjects. After completing this study and with further experience, however, we have found that subjects do not need to push actively into the examiner's hands when performing the test; a passive lean by the subjects into the examiner's hands is preferred. By eliminating the need for an active push, subjects may be less hesitant to perform the Push and Release Test, and a passive lean provides added reliability by eliminating variations in the test. As a second potential limitation, the Push and Release rating scale requires the examiner to differentiate whether a subject took a step to reorient the feet side by side or to maintain balance. While this instruction may introduce some subjectivity into the rating, the Push and Release Test's inter-rater reliability remained high, suggesting that separate examiners could consistently differentiate between the two reasons for stepping.

As a third limitation, our results showed that the Push and Release Test was less specific than the Pull Test for subjects with high balance confidence and no history of falls. While this lack of specificity may be due in part because the Push and Release Test identifies subjects (with or without PD) with burgeoning balance impairments, the test may be less informative when testing subjects with good balance. Because the

Push and Release Test provides high sensitivity and the Pull Test provides high specificity, the two tests may best be used together as complimentary assessments. Nevertheless, the results demonstrate that the third trial of the Push and Release Test provides a valid, more consistent, and more sensitive alternative to the Pull Test when assessing postural stability in subjects with PD.

Acknowledgement The authors thank Maryann Seeger PT and Kevi Ames PT for serving as raters, Kate Doebke RN and the OHSU Movement Disorders clinicians for referring subjects, and Linda Kleen RN and the YMCA for referring subjects. The authors also thank Patricia Carlson-Kuhta PhD, Leta Guptill, and Triana Nagel-Nelson for help with data collection, Andrew Owings for assembling and testing the pressure-sensitive gloves, and Dawn Peters for statistical assistance. The authors extend deep gratitude to our subjects for their participation. This project was supported by NIH grant AG-06457.

References

1. Adkin AL, Frank JS, Jog MS (2003) Fear of falling and postural control in Parkinson's disease. Mov Disord 18:496–502
2. Bloem BR, Grimbergen YA, Cramer M, Willemsen M, Zwinderman AH (2001) Prospective assessment of falls in Parkinson's disease. J Neurol 248:950–958
3. Cummings SR, Nevitt MC, Kidd S (1988) Forgetting falls. The limited accuracy of recall of falls in the elderly. J Am Geriatr Soc 36:613–616
4. Duncan PW, Weiner DK, Chandler J, Studenski S (1990) Functional reach: a new clinical measure of balance. J Gerontol 45:M192–197
5. Fahn S, Elton R (1987) Unified Parkinson's disease rating scale. In: Fahn S, Marsden D, Calne D (eds) Recent Developments in Parkinson Diseases. Macmillan, London, pp 153–163
6. Fregly AR, Graybiel A (1968) An ataxia test battery not requiring rails. Aerosp Med 39:277–282
7. Friedman SM, Munoz B, West SK, Rubin GS, Fried LP (2002) Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. J Am Geriatr Soc 50:1329–1335
8. Horak FB, Jacobs JV, Tran VK, Nutt JG (2004) The push and release test: An improved clinical postural stability test for patients with Parkinson's disease. Mov Disord 19:S170
9. Lajoie Y, Gallagher SP (2004) Predicting falls within the elderly community: comparison of postural sway, reaction time, the Berg balance scale and the Activities-specific Balance Confidence (ABC) scale for comparing fallers and non-fallers. Arch Gerontol Geriatr 38:11–26
10. Munhoz RP, Li JY, Kurtinecz M, Piboolnurak P, Constantino A, Fahn S, Lang AE (2004) Evaluation of the pull test technique in assessing postural instability in Parkinson's disease. Neurology 62:125–127
11. Myers AM, Powell LE, Maki BE, Holliday PJ, Brawley LR, Sherk W (1996) Psychological indicators of balance confidence: relationship to actual and perceived abilities. J Gerontol A Biol Sci Med Sci 51:M37–M43
12. Nutt JG, Hammerstadt JP, Gancher ST (1992) Parkinson's disease: 100 maxims. Edward Arnold, London
13. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) Scale (1995) J Gerontol A Biol Sci Med Sci 50A:M28–M34
14. Visser M, Marinus J, Bloem BR, Kisjes H, van den Berg BM, van Hilten JJ (2003) Clinical tests for the evaluation of postural instability in patients with Parkinson's disease. Arch Phys Med Rehabil 84:1669–1674