

PAPER

Multiple balance tests improve the assessment of postural stability in subjects with Parkinson's disease

J V Jacobs, F B Horak, V K Tran, J G Nutt

J Neurol Neurosurg Psychiatry 2006;77:322–326. doi: 10.1136/jnnp.2005.068742



Appendices A and B are available at <http://jnnp.com/supplemental>.

See end of article for authors' affiliations

Correspondence to:
Dr Fay B Horak,
Neurological Sciences
Institute of Oregon Health
and Science University,
OHSU West Campus –
Bldg 1, 505 NW 185th
Avenue, Beaverton, OR
97006-3499, USA;
HorakF@ohsu.edu

Received 31 March 2005
Revised version received
27 September 2005
Accepted 12 October 2005

Objectives: Clinicians often base the implementation of therapies on the presence of postural instability in subjects with Parkinson's disease (PD). These decisions are frequently based on the pull test from the Unified Parkinson's Disease Rating Scale (UPDRS). We sought to determine whether combining the pull test, the one-leg stance test, the functional reach test, and UPDRS items 27–29 (arise from chair, posture, and gait) predicts balance confidence and falling better than any test alone.

Methods: The study included 67 subjects with PD. Subjects performed the one-leg stance test, the functional reach test, and the UPDRS motor exam. Subjects also responded to the Activities-specific Balance Confidence (ABC) scale and reported how many times they fell during the previous year. Regression models determined the combination of tests that optimally predicted mean ABC scores or categorised fall frequency.

Results: When all tests were included in a stepwise linear regression, only gait (UPDRS item 29), the pull test (UPDRS item 30), and the one-leg stance test, in combination, represented significant predictor variables for mean ABC scores ($r^2=0.51$). A multinomial logistic regression model including the one-leg stance test and gait represented the model with the fewest significant predictor variables that correctly identified the most subjects as fallers or non-fallers (85% of subjects were correctly identified).

Conclusions: Multiple balance tests (including the one-leg stance test, and the gait and pull test items of the UPDRS) that assess different types of postural stress provide an optimal assessment of postural stability in subjects with PD.

The Unified Parkinson's Disease Rating Scale (UPDRS) is commonly used for clinical assessment of disease severity in subjects with Parkinson's disease (PD).¹ The presence of postural instability separates mild PD (Hoehn and Yahr stages 1 and 2) from moderate and severe PD (Hoehn and Yahr stages 3–5); quality of life inexorably deteriorates for patients with moderate or severe PD.² Despite the severe consequences associated with falling and instability, the UPDRS only includes one test specifically focused on postural stability, and the differentiation of Hoehn and Yahr stages 1 and 2 from stages 3–5 is based only on this test, the pull test, item no. 30 of the UPDRS.¹ The pull test, however, does not adequately discriminate PD subjects from healthy subjects or predict falls in PD subjects (although these shortcomings may depend on the severity of disease in the subjects tested, the subjects' medication dose cycle, or the manner of test execution and scoring).^{3–6}

Although the pull test is the only UPDRS item specifically focused on postural stability, it is not the only item of the UPDRS motor exam related to balance impairment. According to Adkin *et al*, a composite score of postural instability and gait difficulty (the PIGD score) which includes the four UPDRS items of arising from chair (item 27), posture (item 28), gait (item 29), and postural stability (item 30, the pull test) correlates with the Activities-specific Balance Confidence scale (the ABC scale, a questionnaire assessing balance confidence during activities of daily living) and with the sway of the centre of pressure during various tasks of standing balance.⁷ Thus, although the pull test may not represent a comprehensive evaluation of postural stability on its own, when combined with these three other UPDRS items

related to PIGD, clinicians may gain additional insight into a patient's postural stability and its relationship to balance confidence and falls.

Despite containing multiple items related to PIGD, the UPDRS may not provide a comprehensive assessment of balance impairment. Additional simple tests of postural stability could improve assessment of balance impairment in subjects with PD, because multiple tests could assess different types of postural stress contributing to instability and falls.^{3–8} Subjects with PD fall in many different circumstances, such as during base-of-support transitions of walking or turning, when standing at the limits of stability during a reach, or after an external perturbation, such as a push, slip, or trip.³ Different mechanisms of neural control may underlie postural stability within each of these different contexts. Therefore, to adequately determine a PD subject's postural stability, the examiner may need to use multiple tests of postural stability in order to adequately assess each type of postural control. We chose the one-leg stance test and the functional reach test as additions to the pull test because, together, the three balance tests assess different aspects of postural stress: (i) the functional reach test examines a subject's perceived limits of stability, (ii) the one-leg stance test examines a subject's ability to maintain equilibrium during a transition to a small base of support, and (iii) the pull test examines postural control in response to an external

Abbreviations: ABC scale, Activities-specific Balance Confidence scale; PD, Parkinson's disease; PIGD, postural instability and gait difficulty; SD, standard deviation; UPDRS, Unified Parkinson's Disease Rating Scale

perturbation.^{1 9 10} Combining these tests, therefore, may furnish a more comprehensive evaluation of postural stability in subjects with PD. To test the hypothesis that multiple clinical tests of postural stability provide a more accurate assessment of balance impairment in PD than any single test, we calculated regression models that included the subjects' responses from the one-leg stance test and the functional reach test, in addition to items 27–30 of the UPDRS, to determine which combination of these tests best predicts balance confidence and falls in patients with PD.

METHODS

Subjects

Sixty seven subjects with idiopathic PD (20 females and 47 males) and 65 age matched control subjects without PD (31 females and 34 males) gave informed consent and participated in the protocol approved by the Institutional Review Board of Oregon Health and Science University (OHSU), Portland, OR, USA. PD subjects were recruited as successive attendees to OHSU's Movement Disorders Clinic according to the following inclusion criteria: (i) they exhibited at least two of the three symptoms of rest tremor, rigidity, and bradykinesia, (ii) their neurological history suggested no alternative diagnosis, (iii) they did not exhibit marked musculoskeletal impairments unrelated to PD, and (iv) they were able to stand unaided to perform the balance tests. Based on these criteria, subjects with atypical parkinsonism or subjects with more severe imbalance were excluded from the study. The 67 control subjects were recruited from a local fitness centre if they matched the ages of the PD subjects and had no history of neurological impairment. For this study, the control group serves only to provide a reference for comparison when analysing the performance of the PD subjects on the balance tests and the ABC questionnaire. The PD subjects had a mean \pm standard deviation (SD) age of 67 ± 12 years (range 42–88 years) and had had PD for 10 ± 6 years (range 2–25 years); the control subjects were 67 ± 10 years of age (range 41–84 years).

The PD subjects were tested regardless of where they were in their medication dose cycle. Twenty PD subjects reported themselves to be in the "on" medication state, 10 PD subjects felt "in between", eight PD subjects reported themselves to be in the "off" medication state, and 28 PD subjects did not report their medication state. We collapsed the data of the PD subjects across medication states because both ANOVA and Kruskal-Wallis comparisons determined that there were no significant differences in the balance test or questionnaire scores between subjects in different medication states ($p > 0.09$ for all comparisons).

Protocol

For the PD subjects, a neurologist performed the UPDRS motor exam. For all subjects, a physical therapist or a neurologist rated the subjects' performance on three trials of each balance test, administered the ABC questionnaire, and asked the subjects how many times they had fallen in the previous year. We timed the one-leg stance test until subjects reached a maximum of 30 s. ABC scores represent the subjects' average per cent confidence to perform activities of daily living.¹¹ Because retrospective fall reports may be unreliable, we categorised subjects as non-fallers (subjects who reported zero falls in the past year), one-time fallers (subjects who reported experiencing one fall in the past year), and recurrent fallers (subjects who reported more than one fall in the past year).¹² Please refer to appendix A of the supplemental online material (available at <http://jnnp.com/supplemental>) for details regarding the execution of each balance test and the ABC questionnaire.

Analysis

To compare differences in performance between the PD subjects and the control subjects, we compared the two groups' one-leg stance times, functional reach lengths, mean ABC scores, and pull test scores using Mann-Whitney U tests. We chose a non-parametric test, because a Shapiro-Wilks' test for normality determined that the subjects' scores were not normally distributed.

To determine whether items 27–30 of the UPDRS, the one-leg stance test, and the functional reach test represented independent predictors of balance confidence, we performed linear regression models that included mean ABC scores as the dependent variable, and each of the six tests, by themselves, as the independent variable. We also performed a stepwise linear regression that included items 27–30 of the UPDRS as the independent variables and mean ABC scores as the dependent variable to determine which combination of UPDRS items best predicts ABC scores (stepwise regression no. 1). We performed a second stepwise linear regression that, in addition to items 27–30 of the UPDRS, included the one-leg stance test and the functional reach test as the independent variables to determine whether the addition of other balance tests to these UPDRS items significantly improved the regression models' predictions of ABC scores (stepwise regression no. 2).

Multinomial logistic regressions, with categorised fall frequency as the dependent variable, determined how well combinations of items 27–30 of the UPDRS, the one-leg stance test, and the functional reach test predicted the subjects' fall history. We performed several multinomial logistic regressions that included different combinations of independent variables. We only included independent variables that represented significant predictors of categorised fall history when only one variable was included in the model. We then tested different combinations of these independent variables to determine an optimal model: (i) that correctly identified the highest percentage of subjects as non-fallers, one-time fallers, or recurrent fallers with the fewest number of independent variables, and (ii) in which all the independent variables represented significant predictors of fall history, given the inclusion of the other independent variables.

RESULTS

The third trial of the one-leg stance test, the pull test, and the functional reach test was used because it predicted ABC scores and classified the subjects' fall history with higher levels of significance than the first or second trials of the balance tests. For a detailed comparison of the balance tests' first and third trial levels of significance, please refer to appendix B of the supplemental online material (available at <http://jnnp.com/supplemental>).

When comparing the balance test and questionnaire scores of PD subjects with control subjects, the PD subjects exhibited significantly higher pull test scores (median score 1 for the PD subjects and 0 for the control subjects; $p < 0.001$), shorter functional reach lengths (length 28 ± 9 cm (mean \pm SD) for the PD subjects and 38 ± 8 cm for the control subjects; $p < 0.001$), and lower mean ABC scores (per cent confidence $72 \pm 21\%$ for the PD subjects and $92 \pm 13\%$ for the control subjects; $p < 0.001$), but similar one-leg stance times (16 ± 12 s for the PD subjects and 16 ± 11 s for the control subjects; $p = 0.93$). Thirty four PD subjects were characterised as non-fallers, 13 PD subjects as one-time fallers, and 19 PD subjects as recurrent fallers. All control subjects were characterised as non-fallers.

In regression models including only one independent variable, every balance test and UPDRS-PIGD item independently represented significant predictors of ABC scores

(please refer to “One-variable regressions” of table 1), but Pearson r^2 values between actual ABC scores and predicted ABC scores ranged from only 0.24 to 0.37. In the stepwise regression models, multiple tests significantly contributed to the prediction of ABC scores, and Pearson r^2 values increased to 0.51 when comparing actual ABC scores with predicted ABC scores (table 1). The stepwise regression model which included items 27–30 of the UPDRS demonstrated that, without adding the one-leg stance test or the functional reach test to the model, the gait item, the posture item, and the pull test of the UPDRS represented significant predictor variables for ABC scores (please refer to “Stepwise regression no. 1” of table 1). When adding the one-leg stance test and the functional reach test to the model, however, the gait item, the pull test, and the one-leg stance test, not the posture item, represented significant predictor variables for ABC scores (please refer to “Stepwise regression no. 2” of table 1). Figure 1 illustrates the regression model’s improvement in predicting ABC scores by combining the gait item, the pull test, and the one-leg stance test compared to using each test alone. Stepwise regression no. 2 calculated the predicted ABC score from the equation: $ABC\ score = 81.02 + (0.429 * one\ leg\ stance\ time) - (8.57 * gait\ score) - (9.51 * pull\ test\ score)$.

In multinomial logistic regressions including only one independent variable, four tests represented significant predictors of the subjects’ fall history: the one-leg stance test, the pull test, the gait item, and the arise from chair item (table 2). A regression model including both the one-leg stance test and the gait item of the UPDRS accurately predicted the categorised fall history of the largest percentage of subjects, and both items represented significant predictor variables for identifying non-fallers, one-time fallers, and

Table 1 The significance of improvement provided by each measure to the prediction of ABC scores

| | Significance of improvement to the regression models | | |
|---------------------------------|--|---------------------------|---------------------------|
| | One-variable regressions | Stepwise regression no. 1 | Stepwise regression no. 2 |
| UPDRS item 30: the pull test | <0.0001 | <0.001 | <0.05 |
| UPDRS item 29: gait | <0.0001 | <0.05 | <0.005 |
| UPDRS item 28: posture | <0.0001 | <0.05 | 0.12 |
| UPDRS item 27: arise from chair | <0.0005 | 0.60 | 0.68 |
| One-leg stance | <0.0001 | Not included | <0.05 |
| Functional reach | <0.0001 | Not included | 0.38 |

recurrent fallers (table 2). The posture item of the UPDRS and the functional reach test were not significant predictors of the subjects’ fall history (table 2).

Because the one-leg stance test and the gait item of the UPDRS represented significant predictors of ABC scores and provided the most accurate identification of the subjects’ fall history, we determined cut-off scores for these tests that provide the highest combination of sensitivity and specificity for identifying subjects with or without a history of one or more falls. For the gait item of the UPDRS, a cut-off score of two points provided the highest combination of sensitivity and specificity for a history of one or more falls: only 28% of

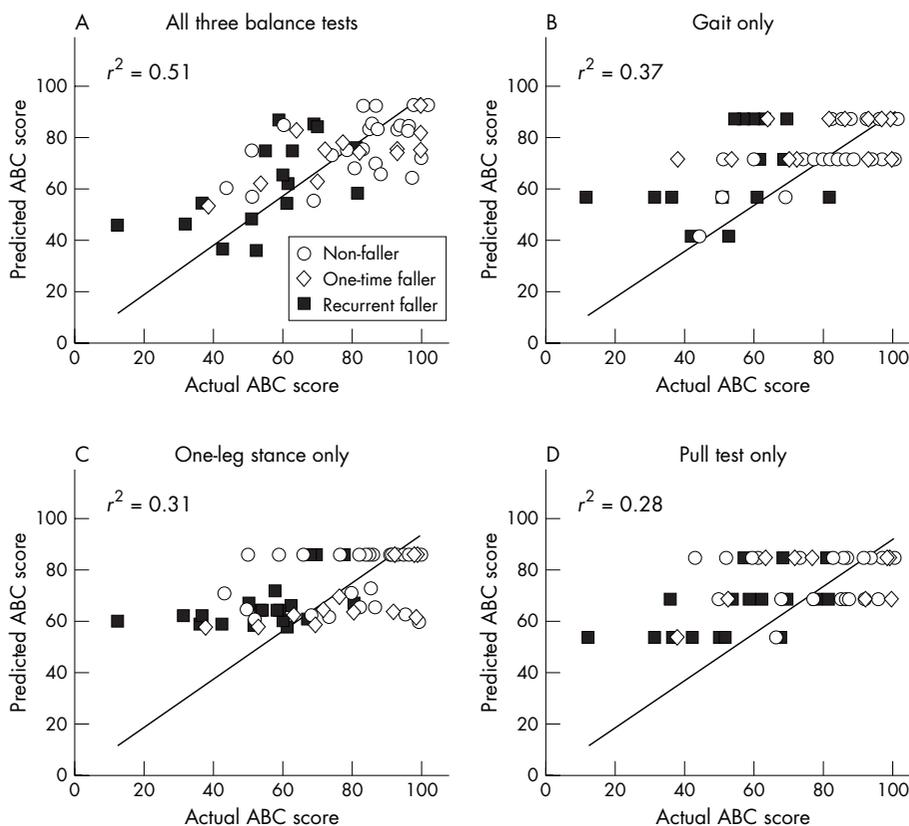


Figure 1 Scatter plots comparing the actual mean ABC scores with predicted scores from linear regression models including (A) the pull test, the one-leg stance test, and the gait item, (B) only the gait item, (C) only the one-leg stance test, and (D) only the pull test. Axes represent the subjects’ average per cent confidence to complete the 16 tasks listed in the ABC scale’s questionnaire, and the diagonal line represents a perfect prediction of ABC scores.

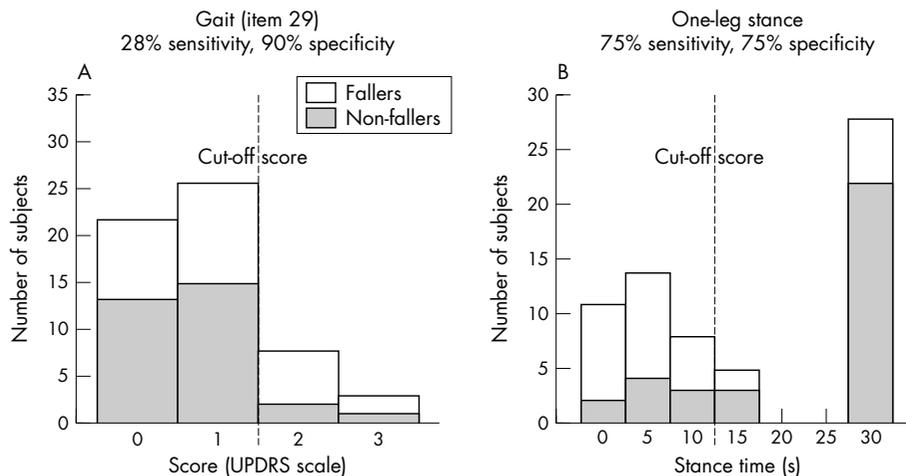


Figure 2 Frequency histograms illustrating the distribution of scores on (A) the gait item and (B) the one-leg stance test. Grey bars indicate subjects who reported experiencing no falls in the past year (non-fallers) and white bars indicate subjects who reported experiencing one or more falls in the past year (fallers). The dashed vertical line indicates the cut-off score that provided the highest combined sensitivity for identifying fallers and specificity for identifying non-fallers.

subjects with a history of falling exhibited a gait score of two or more points (low sensitivity), and 90% of the non-fallers exhibited a gait score of less than two points (high specificity; fig 2A).

For the one-leg stance test, a cut-off time of 10 s provided the highest combination of sensitivity and specificity for a history of one or more falls: 75% of subjects with a history of falling exhibited one-leg stance times of 10 s or less (high sensitivity), and 74% of the non-fallers exhibited one-leg stance times greater than 10 s (high specificity; fig 2B). Therefore, logistic regression models correctly predicted the highest number of subjects as fallers or non-fallers when combining a test that was highly sensitive for identifying fallers (that is, the one-leg stance test) with a test that was highly specific for identifying non-fallers (that is, the gait item).

DISCUSSION

A combination of balance tests and UPDRS items that relate to PIGD predicted balance confidence in PD subjects, as measured by ABC scores, better than any single balance test or UPDRS-PIGD item. In the regression model that included items 27–30 of the UPDRS, the pull test, gait item, and posture item significantly improved the prediction of ABC scores, suggesting that each of these items assesses different aspects of balance control relating to a PD subject's balance confidence. When the one-leg stance test and the functional reach test were added to the regression model, the one-leg stance test replaced the posture item as a significant contributor to the prediction of ABC scores. Both combinations, however, provided similar correlations among actual and predicted ABC scores. Therefore, it may only be mildly useful to add the one-leg stance test to the UPDRS motor exam in order to predict balance confidence.

The ability to discriminate subjects with high or low balance confidence is important for several reasons: (i) balance confidence correlates with balance impairment in PD,⁷ (ii) physically active subjects report higher ABC scores than those who are less active,^{11 13 14} and (iii) compared with subjects who report high balance confidence on the ABC scale, subjects with low balance confidence are nearly twice as likely to fall within the next 2 years.¹⁵ Therefore, there may be a benefit in administering a combination of balance tests and UPDRS-PIGD items (with, perhaps, the ABC questionnaire), because, together, they may provide a more accurate assessment of how balance impairments due to PD affect a subject's choice of daily activities and provide insight about that subject's risk for future falls.

Assessing multiple balance tests and UPDRS items also improved predictions of the subjects' fall history. A combination of the one-leg stance test and the gait item of the UPDRS provided the optimal regression model for identifying subjects as fallers or non-fallers by accurately identifying the fall history of the most subjects with the fewest significant predictor variables. As significant predictors within the same model, the one-leg stance test and the gait item of the UPDRS must have assessed different balance impairments that contributed to the falls of different subsets of PD subjects, thereby identifying the fall history of more subjects together than when considered separately. In support of this notion, the one-leg stance test was the most sensitive predictor for identifying fallers, whereas the gait

Table 2 The percentage of subjects correctly identified as fallers or non-fallers by multinomial logistic regressions

| Balance tests included in the regression (p value of each test's contribution to the model) | Per cent of subjects correctly identified | | |
|---|---|------------------|-------------------|
| | Non-fallers | One-time fallers | Recurrent fallers |
| Balance tests in combination | | | |
| One-leg stance test and gait (<0.005, <0.05) | 97% | 75% | 69% |
| One-leg stance test and arise from chair (<0.05, 0.062) | 90% | 68% | 56% |
| One-leg stance test and pull test (<0.05, 0.105) | 94% | 46% | 63% |
| Arise from chair and gait (<0.005, <0.05) | 81% | 67% | 50% |
| Arise from chair and pull test (<0.05, 0.109) | 84% | 67% | 25% |
| Pull test and gait (0.055, 0.097) | 94% | 8% | 50% |
| Balance tests alone | | | |
| Arise from chair (<0.005) | 74% | 68% | 38% |
| One-leg stance test (<0.01) | 94% | 39% | 37% |
| Pull test (<0.01) | 97% | 0% | 37% |
| Gait (<0.05) | 90% | 0% | 50% |
| Posture (0.057) | 100% | 0% | 0% |
| Functional reach test (0.313) | 91% | 23% | 68% |

item of the UPDRS was more specific for identifying non-fallers. Therefore, combining a balance test that was sensitive to subjects who had fallen with a balance test that was specific to subjects who had not fallen may have improved the models' predictions of fall history by providing complementary, multi-factorial assessments of the subjects' postural instability.

For the one-leg stance test, a cut-off time of 10 s provided the best combination of sensitivity and specificity for fall history in the PD subjects, consistent with a previous report by Smithson *et al* who reported that PD subjects with a history of falling, on average, exhibited one-leg stance times of under 10 s, and PD subjects without a history of falling, on average, exhibited one-leg stance times of about 15 s.⁵ In addition to the one-leg stance test's ability to accurately predict the PD subjects' fall history, the one-leg stance test and the arise from chair item were also the most informative tests for predicting subjects with a history of only one fall. Because the one-leg stance test and the arise from chair item both evaluate a subject's ability to make a transition from a large base of support to a small base of support, this result suggests that an impaired ability to shift to a smaller base of support coincides with the onset of falls in PD subjects. Thus, evaluating PD patients during base-of-support transitions may allow clinicians to identify subjects at risk for falls earlier than with the pull test, thereby providing an opportunity to institute therapies before the patients become injured or disabled by recurrent falls. A long term, prospective assessment of falls, however, would be necessary to establish the one-leg stance or arising from a chair as tasks that predict future falls in PD subjects.

Although we did not record the circumstances surrounding the subjects' falls, we suspect that the context of the subjects' falls affected the relative strength of each balance test's ability to predict the subjects' fall history. PD subjects fall most when turning, when getting out of a bed or chair, or when walking, followed by falls due to external perturbations (for example, from a slip or trip).³ Therefore, the one-leg stance test, the gait item, and the pull test were likely informative of the subjects' fall history because they assess balance during tasks similar to those which are the most troublesome for PD subjects.

The results demonstrate that many UPDRS items and balance tests are capable of relating to a subject's balance confidence and fall history, regardless of the subject's medication state at the time of testing. The balance tests, however, may be more informative when testing subjects in a certain medication state, and some tests may be more sensitive than others to changes in a subject's medication state. Therefore, because we did not control for the subjects' medication dose cycle, the results reported in this study may represent underestimations of the balance tests' ability to predict balance confidence and fall history.

Compared with individual balance tests or individual UPDRS items related to PIGD, multiple balance tests and UPDRS-PIGD items improved the assessment of balance confidence and fall history. The one-leg stance test, the pull test, and the gait item of the UPDRS, together, represented the most informative measures of balance confidence and fall

history, perhaps because (i) the tests assess different aspects of postural stress related to falling, and (ii) the tests provide additional sensitivity and specificity for postural instability. Therefore, we recommend that clinicians evaluate postural stability in PD patients using the one-leg stance test, the pull test, and the gait item of the UPDRS motor exam.

ACKNOWLEDGEMENTS

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, and Dawn Peters for statistical consultation.

Authors' affiliations

J V Jacobs, F B Horak, Neurological Sciences Institute, Oregon Health and Science University, Portland, OR, USA

V K Tran, J G Nutt, Department of Neurology, Oregon Health and Science University, Portland, OR, USA

This project was supported by NIH grant AG-06457

Competing interests: none declared

Ethics approval: subjects gave informed consent, and the protocol was approved by the Institutional Review Board of Oregon Health and Science University, Portland, OR

REFERENCES

- 1 **Fahn S**, Elton R. Unified Parkinson's disease rating scale. In: Fahn S, Marsden D, Calne D, eds. *Recent developments in Parkinson diseases*. London: Macmillan, 1987:153–63.
- 2 **Keranen T**, Kaakkola S, Sotaniemi K, *et al*. Economic burden and quality of life impairment increase with severity of PD. *Parkinsonism Relat Disord* 2003;**9**:163–8.
- 3 **Bloem BR**, Grimbergen YA, Cramer M, *et al*. Prospective assessment of falls in Parkinson's disease. *J Neural* 2001;**248**:950–8.
- 4 **Visser M**, Marinus J, Bloem BR, *et al*. Clinical tests for the evaluation of postural instability in patients with Parkinson's disease. *Arch Phys Med Rehabil* 2003;**84**:1669–74.
- 5 **Smithson F**, Morris ME, Iansek R. Performance on clinical tests of balance in Parkinson's disease. *Phys Ther* 1998;**78**:577–92.
- 6 **Munhoz RP**, Li JY, Kurtinecz M, *et al*. Evaluation of the pull test technique in assessing postural instability in Parkinson's disease. *Neurology* 2004;**62**:125–7.
- 7 **Adkin AL**, Frank JS, Jog MS. Fear of falling and postural control in Parkinson's disease. *Mov Disord* 2003;**18**:496–502.
- 8 **Behrman AL**, Light KE, Flynn SM, *et al*. Is the Functional Reach Test useful for identifying falls risk among individuals with Parkinson's disease? *Arch Phys Med Rehabil* 2002;**83**:538–42.
- 9 **Duncan PW**, Weiner DK, Chandler J, *et al*. Functional reach: a new clinical measure of balance. *J Gerontol* 1990;**45**:M192–7.
- 10 **Fregly AR**, Graybiel A. An ataxia test battery not requiring rails. *Aerosp Med* 1968;**39**:277–82.
- 11 **Powell LE**, Myers AM. The Activities-specific Balance Confidence (ABC) Scale. *J Gerontol A Biol Sci Med Sci* 1995;**50A**:M28–34.
- 12 **Cummings SR**, Nevitt MC, Kidd S. Forgetting falls. The limited accuracy of recall of falls in the elderly. *J Am Geriatr Soc* 1988;**36**:613–16.
- 13 **Myers AM**, Powell LE, Maki BE, *et al*. Psychological indicators of balance confidence: relationship to actual and perceived abilities. *J Gerontol A Biol Sci Med Sci* 1996;**51**:M37–43.
- 14 **Myers AM**, Fletcher PC, Myers AH, *et al*. Discriminative and evaluative properties of the Activities-specific Balance Confidence (ABC) Scale. *J Gerontol A Biol Sci Med Sci* 1998;**53**:M287–94.
- 15 **Friedman SM**, Munoz B, West SK, *et al*. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. *J Am Geriatr Soc* 2002;**50**:1329–35.