Test-retest Reliability of the Physical Performance Test for Persons with Parkinson Disease

Karen A. Paschal, PT, DPT;[1] Amber R. Oswald, PT, DPT;[2] Robert W. Siegmund, PT, DPT;[2] Susan E. Siegmund, PT, DPT;[2] A. Joseph Threlkeld, PT, PhD[1]

[1] Department of Physical Therapy, School of Pharmacy and Health Professions, Creighton University, Omaha, NE
[2] Student in DPT program at Creighton University when study was conducted, Omaha, NE

ABSTRACT

Background and Purpose: Reliable measures are needed to document functional status and disease progression for people with Parkinson disease (PD). We, therefore, evaluated the reliability of the Physical Performance Test (PPT) for people with PD. **Methods**: Fourteen community-dwelling subjects with PD participated: 8 males, 6 females; modified Hoehn and Yahr Stages 2 and 2.5; mean age 62.4 years (±6.3). The test was administered twice, 1 week apart. The 7-item and 9-item summary scores of the PPT were each compared between sessions using repeated measures analysis of variance (ANOVA). The intraclass correlation coefficient (ICC) and method error (ME) were calculated to further assess reliability. **Results**: Between sessions, 7- and 9-item summed scores were not statistically different. The range of summed scores fell in the midst of the available score range for both the 7- and 9-item tests suggesting resistance to floor and ceiling effects. The ICCs showed good agreement (7-item = 0.818; 9-item = 0.895) indicating test reliability for this population. Based on the ME, an examiner can expect a 6% variation for the 7-item summary score and a 4% variation for the 9-item score summary between testing sessions. **Conclusions**: The 7- and 9-item PPTs were demonstrated to be reliable objective measures in individuals with PD. Simple props and brief administration time (10-15 minutes) make the test practical to use.

Key Words: Physical Performance Test, Parkinson disease, reliability

INTRODUCTION

Parkinson disease (PD) is a chronic, progressive disease of the central nervous system.[1] In the United States, the overall prevalence of PD is approximately 124 per 100,000 population.[2,3] The prevalence of PD increases dramatically with age. Data indicate that PD affects approximately 350,000 persons aged 65 years or older or 1,000 per 100,000 population.[2] Clinical presentation of PD may include slow or decreased movements (bradykinesia), an increase in muscle tone (rigidity), an inability to initiate movement (akinesia), postural instability, a festinating gait pattern, decreased arm swing during ambulation, and balance deficits.[1] The functional consequences of these impairments have been characterized using an array of standardized assessment tools, including the Physical Performance Test (PPT).[4-12]

The 9-item PPT is a multidimensional measure of function that can be used in clinical assessment. Activities of daily living used in the test include: sentence writing, simulated eating, turning 360°, putting on and removing a jacket, lifting and then placing a book on a shelf, picking up a penny from the floor, a 50-foot walk test, and 2 measures of stair climbing (time to ascend one flight; number of flights climbed up and down).[4] Scores for 7 of the 9 items are based on the time to complete the task. The 360° turn is scored on the continuity and steadiness of steps. The final item, stair-climbing, is scored on the number of flights completed. Scores of individual items are summed giving a possible range of 0-36. A higher summed score represents better performance (refer to Appendix A).

The PPT has been shown to be reliable and valid as a measure of physical functional capabilities,[13] risk of falls,[14] and dynamic balance[15] in older adults. Rozzini et al reported that the performance-based PPT predicted functional limitation before such limitations were detected using self-report scales.[16] In addition, the PPT has been used in a variety of functional assessment settings in determining effectiveness of exercise with older adults including: monitoring change in physical frailty,[17] prevention of functional and health-related declines in quality of life,[18] and increasing function following exercise training programs including strength training and treadmill walking.[19] The PPT has been used to determine the limitations due to Alzheimer disease by assessing physical function[20] and has been used as a functional evaluation tool in patients after transmetatarsal amputation.[21-23] The reliability of the 9-item PPT has not been determined for use with patients with PD.

Suteerawattananon and Protas reported the use of a limited form (7-item) of the PPT, as well as other timed functional tests, for 11 persons with PD.[24] In their study, the 7-item PPT was found to be reliable with a Spearman correlation coefficient of .920. However, their study did not include the 2 measures of stair-climbing, an important functional ability for community-

Address all correspondence to: Karen A. Paschal, Department of Physical Therapy, Creighton University, 2500 California Plaza, Omaha, NE 68178, Ph:402-280-5690, Fax: 402-280-5692 (kpaschal@creighton.edu).

dwelling adults and a gross measure of endurance, which is part of the 9-item PPT. Moreover, the results of Suteerawattananon and Protas were based on a small population with only a short (24 hour) retest period. Thus, the reliability of the 7-item PPT has not been confirmed. The following study was undertaken to determine the test-retest reliability of the 9-item PPT, to confirm the 7-item PPT test-retest reliability, and to expand the total database of persons with PD whose function has been studied using the PPT.

METHODS

Design

The study protocol was reviewed and approved by the Creighton University Human Subjects Institutional Review Board (IRB). Testing was conducted in the Creighton University Medical Center (CUMC) Department of Physical Therapy Biodynamics Laboratory. All personal information was coded to maintain confidentiality of subjects. This was a repeated measures assessment of between-day reliability.

Subjects

Subjects were recruited from a sample of convenience provided by the CUMC Department of Neurology. Inclusion criteria were the ability to walk with or without an assistive device, community-dwelling, and the ability to understand and follow simple directions. A CUMC neurologist cleared subjects for participation in the study.

Fourteen community dwelling subjects diagnosed with PD at modified Hoehn and Yahr stage 2-2.5 volunteered for this study. There were 8 males and 6 females with ages ranging from 49 to 69 years (mean = 62.4 years). Additional descriptive, anthropometric, and medical data for the subjects are provided in Table 1.

Materials and Procedures

Prior to testing, subjects provided informed consent, demographic information, and a brief medical history (Table 1). Subjects were tested on 2 occasions with scheduled appointments 1 week apart (± 24 hours). Testing times were scheduled to coincide with maximum pharmacological effectiveness, defined as approximately 2 hours after ingestion of the subject's usual dose of anti-Parkinson's medication. Vital signs were assessed prior to testing, prior to stair climbing, and immediately following stair climbing to ensure subject safety.

Three of 5 authors performed all of the testing. Prior to collecting data from any subjects, each examiner practiced the PPT examination sequence until the examiner achieved consistency under the supervision of a physical therapist with extensive experience in conducting the PPT on patients. For recording and safety purposes, the examiner was always assisted by one of the other authors when conducting subject testing. In all cases, test administration was completed in a 10 to 15 minute time frame.

Subjects were given as many as 2 opportunities to complete each item of the test; a second attempt was given when deemed appropriate by the examiner. For example, if a subject's wristwatch snagged on the cuff while donning a jacket (item 4, Appendix A) producing an atypical measurement, the subject was allowed to repeat the item. The opportunity to perform a task again was reserved strictly for instances when the subject was clearly not performing at his or her typical level. In order to reduce tester bias, initial subject scores were unavailable to examiners during the retest visit.

The standard PPT protocol was observed as written by Reuben and Siu[4] with the exception of clarifications noted in Appendix A. These changes were made to further standardize the administration of the PPT by more completely defining several of the tasks. For example, item number 3 of the PPT was defined as a sitting task. In all ambulation tasks, subjects were allowed to use their own assistive devices and were permitted to use one handrail on the steps.

Statistical Analysis

The 7-item summary score and the 9-item summary score were each compared between sessions using repeated measures ANOVA. Reliability was assessed using the type 2, 1 intraclass correlation coefficient (ICC) and the method error (ME). The ME statistic provides an estimate of the score variation between testing sessions (ME = standard deviation of the score differences between test 1 and test 2 divided by $\sqrt{2}$). To place the magnitude in perspective, the ME is converted to a percentage of the mean score differences and referred to as the coefficient of variation of the method error [CVme = ((2 ME / (mean of test 1 + mean of test 2)) x 100].[25] All statistics were calculated using a commercial statistical package (SPSS v.12.0, SPSS, Inc, Chicago, Ill).

RESULTS

Mean summed subject scores for both 7- and 9-item tests are presented in Table 2. There were no statistically significant differences between sessions for the 7-item summary score or the 9-item summary score (ANOVA p=0.257). The ICCs for the 7-item (0.818) and 9-item (0.895) summed scores indicated good reliability between testing sessions. For the 7- and 9-item summed scores, the method error was 1.2 and 1.1 respectively

Table 1. Subject Demographic Information

	Mean (sd)	Range
Age (yr)	62.4 (6.3)	49 - 69
Height (cm)	172.1 (8)	158 - 186
Weight (kg)	81 (14)	60.1 - 112.9
Modified Hoehn & Yahr (score)	2.2 (0.3)	2 - 2.5
Length of Diagnosis (yr)	6.4 (4.9)	1.6 - 17.2

(Table 3). The summary scores had variations of 6% and 4% CVme between sessions for the 7- and 9-item summary scores respectively.

DISCUSSION

Reliable objective measures are important in the study of function of persons with PD. In our study, both the 7- and 9-item PPT demonstrated good test-retest reliability when used for individuals with modified Hoehn and Yahr Stage 2 and 2.5 PD, making it an appropriate tool to observe and quantify physical function. Seven of the 11 subjects studied by Suteerawattananon and Protas had a modified Hoehn and Yahr level of Stage 3.[24] The combined results of our study with that of Suteerawattananon and Protas indicate that the 7-item PPT is reliable across PD severity levels of 2, 2.5, and 3.

Given the nature of PD, subjects experienced fluctuating symptoms which had the potential to change the outcome scores. Although the subjects were able to subjectively report changes in symptoms prior to session 2, these changes did not significantly alter the individual item scores between sessions and suggests that each item is reproducible. The CVme provides an estimate of the percentage score variation an examiner might anticipate assuming that the true value of the measured value has not changed between measurement sessions. These results (7-item = 6%; 9-item=4%) indicate that the PPT is relatively insensitive to the short-term symptom fluctuations typical of PD and captures broader functional measures. Suteerawattananon and Protas[24] retested subjects within 24 hours whereas our study provided 1 week (± 24 hours) between testing session. The probability of short-term symptomatic fluctuations would be greater with our longer period between testing sessions. Even so, the repeated PPT scores did not show significant fluctuation.

Table 2. Descriptive Statistics

Items	Test	Mean ± SD	Minimum-Maximum
1-7 (sum)	1	20.9 ± 2.1	16-24
	2	20.8 ± 2.2	17-24
1-9 (sum)	1	27.8 ± 2.5	23-31
	2	27.7 ± 2.7	21-31

Limitations of this study include the small sample size and the narrow range (2.0-2.5) of Hoehn and Yahr scores. Although the PPT was demonstrated to be a reliable objective measure in individuals with PD, the test has limitations as well. A learning curve can be expected with the use of the PPT. For example, any subject who misspelled 'whales' in item 1 on the first trial was able to spell the word correctly on the second trial even though he or she was not informed of the mistake.

The props needed to administer the PPT (Appendix A) are inexpensive and readily available. Minimal space in addition to a 25-foot walkway and a single flight of stairs is necessary. Total administration time of 10 to 15 minutes plus 2 to 3 minutes to score makes the PPT a practical instrument for clinical use.

Strong test-retest reliability along with mid-range scores suggest the PPT may be an appropriate tool to measure change in physical performance over time and may be resistant to floor or ceiling effects for people with modified Hoehn and Yahr Stage 2-2.5 PD. This indicates that the PPT may be a valid measure of current functional level in individuals with PD although direct assessment of validity was not included in our testing protocol. We also speculate that the PPT may be a useful tracking tool to monitor functional changes over time due to disease progression as well as the efficacy of rehabilitation interventions for persons with PD. Both validity and sensitivity remain to be studied.

CONCLUSIONS

This study indicates that the PPT can reliably measure the functional performance of people with PD. Using the PPT, specific functional deficits can be identified and a reliable baseline measure may be established for individuals prior to initiating interventions. Future studies might include testing the sensitivity of the instrument, as well as, random selection of subjects with a broader range of modified Hoehn and Yahr scores.

ACKNOWLEDGMENTS

We would like to acknowledge all of our participants for their time and effort. We extend our gratitude to the Creighton University Medical Center Neurology Department, specifically John Bertoni, MD, PhD and Pamela Sprenkle, PhD.

REFERENCES

1. Rao G, Fisch L, Srinivasan S, et. al. Does this patient have Parkinson's Disease? *JAMA*. 2003;289:347-353.
2. National Center for Health Statistics. NCHS Data on

Table 3. Method Error and Intra Class Coefficient (ICC)

Items	Method Error	Coefficient of Variation of the Method Error	ICC (95% CI)
1-7 (sum)	1.19	0.06	0.818 (0.434 – 0.942)
1-9 (sum)	1.14	0.04	0.895 (0.673 – 0.966)

Parkinson's Disease. January 8, 2004. Available at: http://www.cdc.gov/nchs/data/factsheets/Parkinsons.pdf. Accessed September 9, 2004.
3. U.S. Bureau of the Census, International Database, United States/2000. Available at http://www.census.gov/cgi-bin/ipc/idpyrs.pl?cty=US&out=s&ymax=300. Accessed September 9, 2004.
4. Reuben RB, Siu AL. An objective measure of physical function of elderly outpatients. *J Am Geriatr Soc*. 1990;38:1105-1112.
5. Hoehn M, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology*. 1967;17:427-442.
6. Fahn S, Elton, RL, UPDRS Program Members. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Goldstein M, Calne DB (Eds). *Recent Developments in Parkinson's Disease, Vol. 2*. Florham Park, NJ: Macmillan Healthcare Information; 1987:153-163.
9. Siderowf A, McDermott M, Kieburtz K, et al. Test-retest reliability of the Unified Parkinson's Disease Rating Scale in patients with early Parkinson's disease: results from a multicenter clinical trial. *Mov Disord*. 2002;17:758-763.
8. Biemans MAJ, Dekker J, VanderWoude LHV. The internal consistency and validity of the Self-assessment Parkinson's Disease Disability Scale. *Clin Rehabil*. 2001;15:221-228.
9. Hobson JP, Edwards NI, Meara RJ. The Parkinson's Disease Activities of Daily Living Scale: a new simple and brief subjective measure of disability in Parkinson's disease. *Clin Rehabil*. 2001;15:241-246.
10. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*. 1991;39:142-148.
11. Morris S, Morris ME, Iansek R. Reliability of measurements obtained with the Timed "Up & Go" Test in people with Parkinson disease. *Phys Ther*. 2001;81:810-818.
12. Behrman AL, Light KE, Flynn SM, Thigpen MT. Is the Functional Reach Test useful for identifying falls risk among individuals with Parkinson's disease? *Arch Phys Med Rehabil*. 2002;83:538-542.
13. King MB, Judge JO, Whipple R, Wolfson L. Reliability and responsiveness of two physical performance measures examined in the context of a functional training intervention. *Phys Ther*. 2000;80:8-16.
14. VanSwearingen JM, Paschal KA, Bonino P, Chen TW. Assessing recurrent fall risk of community-dwelling, frail older veterans using specific tests of mobility and the physical performance test of function. *J Gerontol*. 1998;53: M457-464.
15. Whitney S, Poole J, Cass S. A review of balance instruments for older adults. *Am J Occu Ther*. 1998;52:666-671.
16. Rozzini R, Frisoni GB, Ferrucci L, Barbisoni P, Bertozzi B, Trabucchi M. The effect of chronic diseases on physical function. Comparison between activities of daily living scales and the Physical Performance Test. *Age Ageing*. 1997;26:281-287.
17. Brown M, Sinacore DR, Ehsani AA, Binder EF, Holloszy JO, Kohrt WM. Low-intensity exercise as a modifier of physical frailty in older adults. *Arch Phys Med Rehabil*. 2000;81:960-965.
18. Reuben DB, Frank JC, Hirsch SH, McGuigan KA, Maly RC. A randomized clinical trial of outpatient comprehensive geriatric assessment coupled with an intervention to increase adherence to recommendations. *J Am Geriatr Soc*. 1999;47:269-276.
19. Peel C, Utsey C, MacGregor J. Exercise training for older adults with limitation in physical function. *J Aging Phys Act*. 1999;7:62-75.
20. Zanetti O, Frisoni GB, Rozzini L, Bianchetti A, Trabucchi M. Validity of direct assessment of functional status as a tool for measuring Alzheimer's disease severity. *Age Ageing*. 1998;27:615-622.
21. Mueller M, Salsich G, Strube M. Functional limitations in patients with diabetes and transmetatarsal amputations. *Phys Ther*. 1997;77:937-943.
22. Mueller M, Strube M. Therapeutic footwear: enhanced function in people with diabetes and transmetatarsal amputation. *Arch Phys Med Rehabil*. 1997;78:952-956.
23. Salsich G, Mueller M. Relationship between measures of function, strength and walking speed in patients with diabetes and transmetatarsal amputation. *Clin Rehabil*. 1997;11:60-70.
24. Suteerawattananon M, Protas EJ. Reliability of outcome measures in individuals with Parkinson's Disease. *Physiother Theory Pract*. 2000;16:211-218.
25. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. Norwalk, Conn: Appleton and Lange; 1993:525.

Appendix A. Physical Performance Test Protocol (Reuben and Siu, 1990)

(See Note 1)

1. Ask the subject, when given the command 'go,' to write the sentence 'whales live in the blue ocean.' Time from the word 'go' until the pen is lifted from the page at the end of the sentence. All words must be included and legible. Period need not be included for the task to be considered completed. *(See Note 2)*
2. Five kidney beans are placed in a bowl 5 inches from the edge of the desk in front of the patient. An empty coffee can is placed on the table at the patient's nondominant side. A teaspoon is placed in the patient's dominant hand. Ask the subject, on the command 'go,' to pick up the beans, one at a time, and place each in the coffee can. Time from the command 'go' until the last bean is heard hitting the bottom of the can. *(See Note 2)*
3. Place a Physician's Desk Reference or other heavy book on a table in front of the patient. Ask the patient, when given the command 'go,' to place the book on the shelf above the shoulder level. Time from the command 'go' to the time the book is resting on the shelf. *(See Note 2)*

4. If the subject has a jacket or cardigan sweater, ask him or her to remove it. If not, give the subject a lab coat. Ask the subject, on the command 'go,' to put the coat on completely such that it is straight on his or her shoulders and then remove the garment completely. Time from the command 'go' until the garment has been completely removed. *(See Note 3 and Note 4)*
5. Place a penny approximately 1 foot from the patient's foot on the dominant side. Ask the patient, on the command 'go,' to pick up the penny from the floor and stand up. Time from the command 'go' until the subject is standing erect with penny in hand. *(See Note 4)*
6. With subject in a corridor or in an open room, ask the subject to turn 360°. Evaluate using scale on PPT scoring sheet. *(See Notes 4-6)*
7 Bring subject to start on 50-foot walk test course (25 feet out and 25 feet back) and ask the subject, on the command 'go,' to walk to 25-foot mark and back. Time from the command 'go' until the starting line is crossed on the way back. *(See Note 5-6)*
8. Bring subject to foot of stairs (nine to 12 steps) and ask subject, on the command 'go,' to begin climbing stairs until he or she feels tired and wishes to stop. Before beginning this task, alert the subject to possibility of developing chest pain or shortness of breath and inform the patient to tell you if any of these symptoms occur. Escort the subject up the stairs. Time from the command 'go' until the subject's first foot reaches the top of the first flight of stairs. Record the number of flights (maximum is 4) climbed (up and down). *(See Notes 5-9)*

Note 1: Each subject was given up to 2 opportunities to complete each item if during the first attempt the subject was clearly not performing at their typical level. The best performance was recorded.

Note 2: We defined protocol items 1-3 as seated tasks. Additionally, the seated tasks were performed at a desk with a work surface height of 75 cm using a chair with a 42 cm seat height and a shelf positioned 43 cm above the desk surface within easy horizontal reach. The shelf was approximately shoulder height for most seated subjects thus the text had to be raised above shoulder height to be placed on the shelf.

Note 3: Our study provided one standard jacket which was used by all participants.

Note 4: Item 6 in our protocol was scored based on consensus of the tester and the assistant.

Note 5: A gait belt was placed on each subject before beginning all standing tasks to aid the examiner in preventing falls.

Note 6: Assistive devices were permitted for tasks 6 through 8.

Note 7: To reduce the influence of fatigue, each subject was transported via wheelchair to the stairwell where tasks 8 and 9 were performed.

Note 8: Our protocol defined a standard flight of stairs as 12 steps. Our staircase was equipped with handrails on both sides which each patient was free to use.

Note 9: In order to monitor for overexertion, each subjects' blood pressure and pulse was taken before items 8 and 9 were performed. After completion of items 8 and 9, the vital signs were periodically reassessed and subjects were not allowed to continue until the vital signs returned to pretesting levels.