A brief clinical tool to assess physical function: The mini physical performance test

Consuelo H. Wilkins[a,b,c,*], **Catherine M. Roe**[b,d], and **John C. Morris**[b,d,e,f]

[a] Department of Medicine, Division of Geriatrics and Nutritional Science, Washington University School of Medicine, 4488 Forest Park Ave, Suite 201, St. Louis, Missouri 63108, USA

[b] Alzheimer's Disease Research Center, Department of Neurology, Washington University School of Medicine, 4488 Forest Park Ave, Suite 101, St. Louis, Missouri 63108, USA

[c] Department of Psychiatry, Campus Box 8134, 660 South Euclid Ave, St. Louis, Missouri 63110, USA

[d] Department of Neurology, Washington University School of Medicine, 4488 Forest Park Ave, Suite 101, St. Louis, Missouri 63108, USA

[e] Department of Pathology and Immunology, Campus Box 8118, 660 South Euclid Ave, St. Louis, Missouri 63110, USA

[f] Physical Therapy, Washington University School of Medicine, Campus Box 8502, 4444 Forest Park Ave, Room 1101, St. Louis, Missouri 63108, USA

Abstract

The aim was to develop a brief physical performance assessment tool that can be reliably used to detect physical impairment in older adults with and without mild dementia. Scores on the 9-item physical performance test (PPT) from non-demented participants were used to develop and validate the 4-item mini-PPT. The validated mini-PPT was then used to predict total PPT score and functional physical status in participants with mild dementia. Receiver Operating Curve (ROC) analyses were used to generate a cut off score that classifies participants as functional vs. not functional. The setting was in the Alzheimer's Disease Research Center (Washington University). A total of 1,199 participants met inclusion criteria: 574 non demented participants, 436 with very mild dementia, measured by the clinical dementia rating (CDR) = 0.5 and 189 with mild dementia (CDR = 1). The mean age of the sample was 76.4 years, mean educational attainment was 14 years, 58% were women, and 11% were African American. A 4-item scale, the mini-PPT, was developed (based on the results of multiple regression analyses and clinical meaningfulness) that highly correlated with total PPT score (r = 0.917, p < 0.0001) in the non-demented sample The correlation of the mini-PPT with total PPT was 0.90 among those with very mild, and 0.91 among those with mild dementia. Using the ROCs, a cut off score of 12 correctly classified at least 85% of non demented and demented persons. The 4-item mini-PPT is highly correlated with the 9-item PPT in non demented and mildly demented persons. This brief tool may be useful in detecting early physical impairment in the clinical setting.

*Corresponding author (at address[a]): Phone: +(1-314)-286-2700, Fax: +(1-314)-286-2701, E-mail: cwilkins@dom.wustl.edu.

Conflict of interest statement: none.

1. Introduction

Declines in both physical and cognitive performance in older adults are commonly observed. However, until recently, impaired physical performance was thought to be associated with moderate to severe cognitive impairment (Laurin et al., 2001; Yaffe et al., 2001). Data now supports that physical impairment may actually precede the diagnosis of dementia of the Alzheimer's type (DAT) (Weuve et al., 2004; Podewils et al., 2005; Wang et al., 2006). Several studies report physical impairments, including gait apraxia, reduced strength and slower walking speed in persons, may antedate the diagnosis of DAT (Camicioli et al., 1998; Marquis et al., 2002; Waite et al., 2005).

As the focus of diagnosing DAT has shifted toward identifying affected persons in the earliest symptomatic stages, clinical tools that aid in the recognition of persons at high risk for developing DAT will be useful. The research tools often used to assess physical performance and gait are often time consuming and complex. Additionally, many tools used in older adults have not been validated in people with early dementia; therefore, the ability of the tools to distinguish physical impairments from cognitive ability to perform the tasks must be examined.

The PPT has been used successfully to investigate physical function in demented persons (Shah et al., 2004). Scores on the original PPT instrument, developed by Reuben and Siu (1990), correlate with degree of disability, loss of independence and mortality (Reuben et al., 1992). Although the PPT has been shown to detect early physical decline in community dwelling elderly (Brach et al., 2002), its use in clinical settings may be limited due to its length. The goal of this study was to develop a shortened tool (and create the mini-PPT) that could be used to assess physical functioning in persons with and without mild dementia.

2. Experimental materials and methods

2.1. Participants

All participants were older adults involved in studies of cognitive and functional aging at the Alzheimer's Disease Research Center (ADRC) at Washington University between January 1998 and August 2006. The ADRC recruits cognitively healthy and demented older adult participants from the greater metropolitan St. Louis area (population: 2.5 million). Inclusion criteria for this study were: age over 60 years, ambulatory, able to complete all assessments, and PPT score greater than 5. Persons with moderate or severe dementia and physical disability were excluded.

2.2. Clinical and cognitive assessments

The clinical evaluation included obtaining past medical, social and family history from a reliable informant, usually a spouse or adult child. Information regarding possible cognitive change was obtained by a clinician from semi structured interviews with the informant and separately with the participant. Included in the clinical assessment protocol were the mini mental state examination (MMSE) (Folstein et al., 1975), the short blessed test (SBT) (Katzman et al., 1983), an aphasia battery, a medication inventory, and a depressive features battery.

Using all information from the clinical assessment protocol but without reference to the participant's psychometric performance (see below), the clinician determined the CDR for the participant. The CDR determines the presence or absence of dementia and, when present, rates

its severity (Morris, 1993). The CDR rates cognitive performance in each of six categories: memory, orientation, judgment and problem solving, community affairs, home management and hobbies and personal care. A global CDR = 0 indicates no dementia; a global CDR = 0.5 indicates very mild dementia, and global CDRs of 1, 2 and 3 indicate mild, moderate and severe dementia, respectively. The CDR sum of the boxes score is the summation of the individual scores in each of the six CDR categories (i.e., boxes) and provides a more quantitative measure of cognitive impairment (Berg et al., 1992). Possible sum of boxes scores range from 0 (i.e., the individual scores are 0 in all six CDR categories) to 18 (i.e., the individual scores are 3 in all six CDR categories). A higher global CDR or larger sum of boxes score indicates greater dementia severity. Validity and inter-rater reliability for the CDR have been established (Burke et al., 1988; Morris et al., 1998). Individuals with dementia at the CDR = 0.5 or CDR = 1 stage were included in this study. Two to four weeks after the clinical assessment, participants completed a psychometric battery (Storandt and Hill, 1989); pyschometricians were not informed of the results of the clinical evaluation.

2.3. Physical assessment

Following the general physical and neurological examination, the PPT was administered by a trained research nurse. We modified the original PPT instrument (Reuben and Siu, 1990) by adding the chair rise and progressive Romberg-test of standing balance to simulate stair climbing tasks. Performance on these two tasks, which have been included in other physical performance batteries, (Guralnik et al., 1994; Wang et al., 2002) has been associated with self reported disability, nursing home placement and mortality (Guralnik et al., 1994). Specific tasks in our modified PPT are: (1) writing a sentence, (2) simulated eating (i.e., spooning beans into a container), (3) lifting a book, (4) simulated dressing (i.e., putting on and taking off a jacket), (5) picking up a penny from the floor, (6) turning in a complete circle (i.e., steadiness and continuity of steps), (7) walking 50 ft., (8) the chair rise (i.e., sitting in and rising from a chair five times), and (9) the progressive Rombergtest of standing balance (i.e., standing with feet in tandem, semi tandem and side by side positions). The PPT items will be referred to by the item numbers above in the text. The tasks in our PPT were scored on a 5-point scale: 0, 1, 2, 3 and 4. The total PPT score, a simple summation of the individual item scores, is a composite measure of physical function. The maximum (i.e., best) total score was 36, with a decreasing score indicating increasing disability.

2.4. Development and validation of the mini-PPT

Samples of non demented (CDR = 0) participants were used to develop and initially validate the mini-PPT. Non demented participants (n = 576) were randomly assigned to two groups: the development sample and the validation sample. For participants who took part in the PPT more than once over the study period, their earliest PPT score was used.

Within the development sample, Pearson product moment correlation coefficients were used to examine the correlation of scores on each PPT item with the total PPT score, as well as to examine the intercorrelations among the individual items. The RSQUARE procedure (PROC REG, SAS version 9.1.3 for Linux; SAS Institute, Cary, NC) was used to find the 20 models of PPT items, ten 3-item models and ten 4-item models that best predicted the total PPT score using multiple linear regression. After specifying the model size, the RSQUARE procedure computes all possible regressions (i.e., all possible combinations of items) containing that number of items, and the R^2 value for each regression, to allow comparison of the models' predictive ability. From the 20 models, four potential candidates for the mini-PPT were selected and were subjected to further testing within the validation sample. For each of the four models, a total score was calculated by taking the sum of the scores on the individual items. The correlation of each of the candidate mini-PPT scores with total PPT score were then found, and differences between the correlation coefficients were tested (Cohen and Cohen, 1983).

ROC analyses were conducted to test the ability of each of four candidate scales to discriminate between individuals who would and would not meet PPT criteria for "functional" (total PPT > 28), and differences between the ROC curves were tested using the method of DeLong et al. (1988). The correlation and ROC analyses were then repeated in the validation sample for each of the four candidate scales.

Based on the development and cross validation analyses with non demented participants, one of the four candidate scales was chosen for use as the mini-PPT, and the ability of mini-PPT score to predict total PPT score and functional physical status was then examined in two samples comprised of participants who were mildly demented. The first sample comprised participants who had global CDR scores of 0.5, indicating very mild dementia at the time the PPT was administered, and the second sample comprised participants with global CDR scores of 1, indicating mild dementia. The CDR = 1 group was restricted to those who had a sum of CDR box scores of 6 or less to exclude those with more substantial cognitive impairment.

To enhance the usefulness of the mini-PPT, a cutoff score that could be used in classifying physical performance as functional vs. not functional was developed. Based on inspection of the ROC curve in the non demented validation sample, potential cutoff values likely to yield high sensitivity and specificity values were found, and the sensitivity, specificity, positive predictive value, negative predictive value and percent correctly classified were calculated for each potential cutoff. These analyses were repeated in the very mildly and mildly demented samples to establish whether the same cutoff value should be used among persons with very mild or mild dementia.

3. Results

A total of 1,199 participants met inclusion criteria: 574 non-demented participants, 436 with very mild dementia (CDR = 0.5) and 189 with mild dementia (CDR = 1). The mean age of the sample was 76.4 years, mean educational attainment was 14 years, 58% were women, and 11% were African American. Of the non demented participants, one half (n = 287) were randomly assigned to the development sample, and the remainder (n = 287) to the validation sample. Demographic characteristics for each of the study samples are presented in Table 1.

Each individual PPT item was significantly correlated (p < 0.0001) with the total PPT score. Correlation coefficients ranged from 0.40 to 0.72, with items 5 (r = 0.72), 7 (r = 0.71), 8 (r = 0.72), and 9 (r = 0.63) correlating the highest with total PPT scores. None of the correlations for the other individual PPT items had a correlation coefficient above 0.51.

The ten models of 3 items and the ten models of 4 items that best predicted total PPT scores in the regression analyses had R^2 values that ranged from 0.77 to 0.79 and 0.84 to 0.86, respectively. Four of the 20 models were selected for further testing based on the extent to which the items were clinically meaningful (useful to assess physical function) and practical (easy to administer in an office). Additionally, each contained at least three of the four items with the highest Pearson product moment correlations with total PPT score: item 5, picking up a penny from the floor, item 7, timed 50 foot walk, item 8, the chair rise (i.e., sitting in and rising from a chair five times) and item 9, the progressive Romberg test of standing balance (i.e., standing with feet in tandem, semi tandem and side by side positions).

A score was found for each of the candidate mini-PPTs by summing the scores on the individual items. The correlation of each of these scores with the total PPT score, and the area under the ROC curve testing the ability of each of the candidate scores to discriminate participants classified by the PPT as "functional" from those who were not, was then calculated. As shown in Table 2, each of the candidate scales was highly correlated with the total PPT score, although the 4-item candidate scale correlated more with total PPT score than each of the 3-item models

(p < 0.0001). There were no significant differences between the correlations of the 3-item models with total PPT scores (p > 0.46). Likewise, the area under the ROC curve (AUC) was 0.90 or greater for each of the candidate scales (Table 2). Although these AUCs indicate that the ability of each of the scales to predict functional status was excellent, there was a statistically significant difference in discriminative ability across the three models (p < 0.0001), such that the AUC for the 4-item scale was significantly higher compared to each of the 3 item scales (p < 0.05), and for the candidate scale comprised of items 7, 8 and 9, compared to that made up of items 5, 7 and 9. The remaining two pairwise comparisons were not significant.

When the candidate scales were cross validated in the non demented validation sample, their correlations with total PPT score and ability to classify functional status were similar to the development sample (Table 2). Like the development sample, testing in the validation sample showed that the 4-item candidate scale correlated significantly better with total PPT score than each of the 3-item candidate scales (p < 0.0001). In predicting functional status, the 4-item scale also yielded higher AUCs compared to each of the 3-item candidate scales made up of items 7, 8 and 9 (p = 0.0072); items 5, 7 and 9 (p = 0.0044); and items 5, 8 and 9 (p = 0.0562).

Because of higher correlations with the total PPT, higher AUC values predicting functional status, and the negligible effect on time of administration of having one additional item, the 4-item scale was selected for use as the mini-PPT. We then tested the ability of the mini-PPT to predict total PPT scores and functional status in the mildly demented samples. The correlation of the mini-PPT with total PPT scores was 0.90 (p < 0.0001) and 0.91 (p < 0.0001) in those with very mild (CDR = 0.5) and mild (CDR = 1) dementia, respectively. The AUC (95% confidence interval = 95% CI) value was 0.93 (0.91–0.95) for the very mild dementia sample and 0.95 (0.93–0.98) for the mild dementia sample.

The ROC curve for the non demented validation sample (Figure 1) was examined to identify likely cutoff values on the mini-PPT that could be used to identify individuals who would score as functional on the original PPT. Based on inspection of Figure 1, as well as the ROC curves for the very mild dementia and mild dementia samples (not shown), cutoff scores of 11, 12 and 13 were thought to likely yield high values of both sensitivity and specificity in all samples, and were subjected to further testing. The prevalence values used in these analyses reflect the percentage of participants who scored 28 or above on the total PPT (i.e., functional) in each sample (non demented validation = 0.57; very mild dementia = 0.43; mild dementia = 0.28).

As shown in Table 3, use of a cutoff value of 12 or higher to classify a participant as having functional physical status resulted in the most favorable combinations of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) in each sample. Although cutoff values of 11 produce higher values for sensitivity and NPV, this increase is accompanied by decreased values of specificity and PPV. Likewise, cutoff values of 13 are associated with greater specificity and PPVs, but lower sensitivity and NPVs. Use of a cutoff of 12 or greater also results in the most cases being correctly classified (Table 3).

4. Discussion

In this study, a shortened physical performance tool, the mini-PPT, was developed and validated as an assessment tool to evaluate physical function in older adults with and without mild dementia. The mini-PPT is a brief physical assessment tool that can be administered in a clinical office without the need for additional props or equipment (only a penny and a chair are required).

The 4-item mini-PPT contains items that have been used in other studies assessing physical function in cognitive impaired populations, but unlike many other assessments, this tool allows clinicians and investigators to generate a score and to categorize patients or participants as

functional or impaired. The mini-PPT was able to correctly classify 85% or more of participants as functional compared to the lengthy and more burdensome PPT. The mini-PPT can be completed in less than 5 minutes.

The mini-PPT is similar to the performance based physical function test (PPF) (Wang et al., 2002) which has been used in cognitively impaired persons. Our mini-PPT includes a longer timed walk (50 feet vs. 10 feet) and uses picking a penny up from the floor (trunk flexion) as a measure of strength, while the PPF uses grip strength. Although grip strength has been used as a predictor of future cognitive decline in prior reports, most studies assessing grip strength use quartiles to classify participants and no cut off score is generated. Using a handheld dynamometer may also be impractical in the clinical setting. The mini-PPT assesses each of the 3 components of the short physical performance battery (SPPB) (Guralnik et al., 1994) but uses a longer timed walk and has the additional measure of trunk flexion. The longer walk and strength measure may allow earlier physical impairment to be detected.

While this study includes a large sample of older adults and a representative sample of African Americans, the participants are well educated and relatively healthy, so they may not be representative of all community-dwelling elderly. Additionally, the mini-PPT was developed using the 9-item PPT and was not compared to other performance measures. The PPT, however, has been compared to many other performance measures in prior studies and currently there is no universally accepted gold standard for assessing physical performance in older adults (Rozzini et al., 1993).

5. Conclusions

As data increasingly supports a relationship between physical function and development of dementia, the need to use reliable and valid assessment tools in this population becomes more important. This brief measure of physical function, the mini-PPT, was validated in non demented and mildly demented participants and correctly classified participants as functional or not using a simple summed score. The mini-PPT offers a simple, low cost tool to assess function in older adults that can be easily used in most clinical settings.

Acknowledgments

References

Berg L, Miller JP, Baty J, Rubin EH, Morris JC, Figiel G. Mild senile dementia of the Alzheimer type. 4 Evaluation of intervention. Ann Neurol 1992;31:242–249. [PubMed: 1637132]

Brach JS, Van Swearingen JM, Newman AB, Kriska AM. Identifying early decline of physical function in community-dwelling older women: performance-based and self-report measures. Phys Ther 2002;82:320–328. [PubMed: 11922849]

Burke WJ, Miller JP, Rubin EH, Morris JC, Coben LA, Duchek J, Wittels IG, Berg L. Reliability of the Washington University clinical dementia rating. Arch Neurol 1988;45:31–32. [PubMed: 3337672]

Camicioli R, Howieson D, Oken B, Sexton G, Kaye J. Motor slowing precedes cognitive impairment in the oldest old. Neurology 1998;50:1496–1498. [PubMed: 9596020]

Cohen, J.; Cohen, P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Vol. 2. Erlbaum; Hillsdale, NJ: 1983.

DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845. [PubMed: 3203132]

Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinicians. J Psychiatry Res 1975;12:189–198.

Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol A Biol Sci Med Sci 1994;49:M85–M94.

Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short orientation-memory-concentration test of cognitive impairment. Am J Psychiatry 1983;140:734–739. [PubMed: 6846631]

Laurin D, Verreault R, Lindsay J, MacPherson K, Rockwood K. Physical activity and risk of cognitive impairment and dementia in elderly persons. Arch Neurol 2001;58:498–504. [PubMed: 11255456]

Marquis S, Moore MM, Howieson DB, Sexton G, Payami H, Kaye JA, Camicioli R. Independent predictors of cognitive decline in healthy elderly persons. Arch Neurol 2002;59:601–606. [PubMed: 11939895]

Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. Neurology 1993;43:2412–2414. [PubMed: 8232972]

Morris JC, McKeel D, Fulling K, Torack RM, Berg L. Validation of clinical diagnostic criteria for Alzheimer's disease. Ann Neurol 1988;24:17–22. [PubMed: 3415196]

Podewils LJ, Guallar E, Kuller LH, Fried LP, Lopez OL, Carlson M, Lyketsos CG. Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study. Am J Epidemiol 2005;161:639–651. [PubMed: 15781953]

Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients. J Am Geriatr Soc 1990;38:1105–1112. [PubMed: 2229864]

Reuben DB, Siu AL, Kimpau S. The predictive validity of self-report and performance-based measures of function and health. J Gerontol Med Sci 1992;47:M106–M110.

Rozzini R, Frisoni GB, Bianchetti A, Zanetti O, Trabucchi M. Physical Performance Test and Activities of Daily Living scales in the assessment of health status in elderly people. J Am Geriatr Soc 1993;41:1109–1113. [PubMed: 8409158]

Shah KR, Carr D, Roe CM, Miller JP, Coats M, Morris JC. Impaired physical performance and the assessment of dementia of the Alzheimer type. Alzheimer's Dis Assoc Disord 2004;18:112–119.

Storandt M, Hill RD. Very mild senile dementia of the Alzheimer type. II Psychometric test performance. Arch Neurol 1989;46:383–386. [PubMed: 2705897]

Waite LM, Grayson DA, Piguet O, Creasey H, Bennett HP, Broe GA. Gait slowing as a predictor of incident dementia: 6-year longitudinal data from the Sydney Older Persons Study. J Neurol Sci 2005;229–230:89–93.

Wang L, Van Belle G, Kukull WB, Larson EB. Predictors of functional change: a longitudinal study of nondemented people aged 65 and older. J Am Geriatr Soc 2002;50:1525–1534. [PubMed: 12383150]

Wang L, Larson EB, Bowen JD, Van Belle G. Performance-based physical function and future dementia in older people. Arch Intern Med 2006;166:1115–1120. [PubMed: 16717174]

Weuve J, Kang JH, Manson JE, Breteler MM, Ware JH, Grodstein F. Physical activity, including walking, and cognitive function in older women. J Am Med Assoc 2004;292:1454–1461.

Yaffe K, Barnes D, Nevitt M, Lui LY, Covinsky K. A prospective study of physical activity and cognitive decline in elderly women: women who walk. Arch Intern Med 2001;161:1703–1708. [PubMed: 11485502]
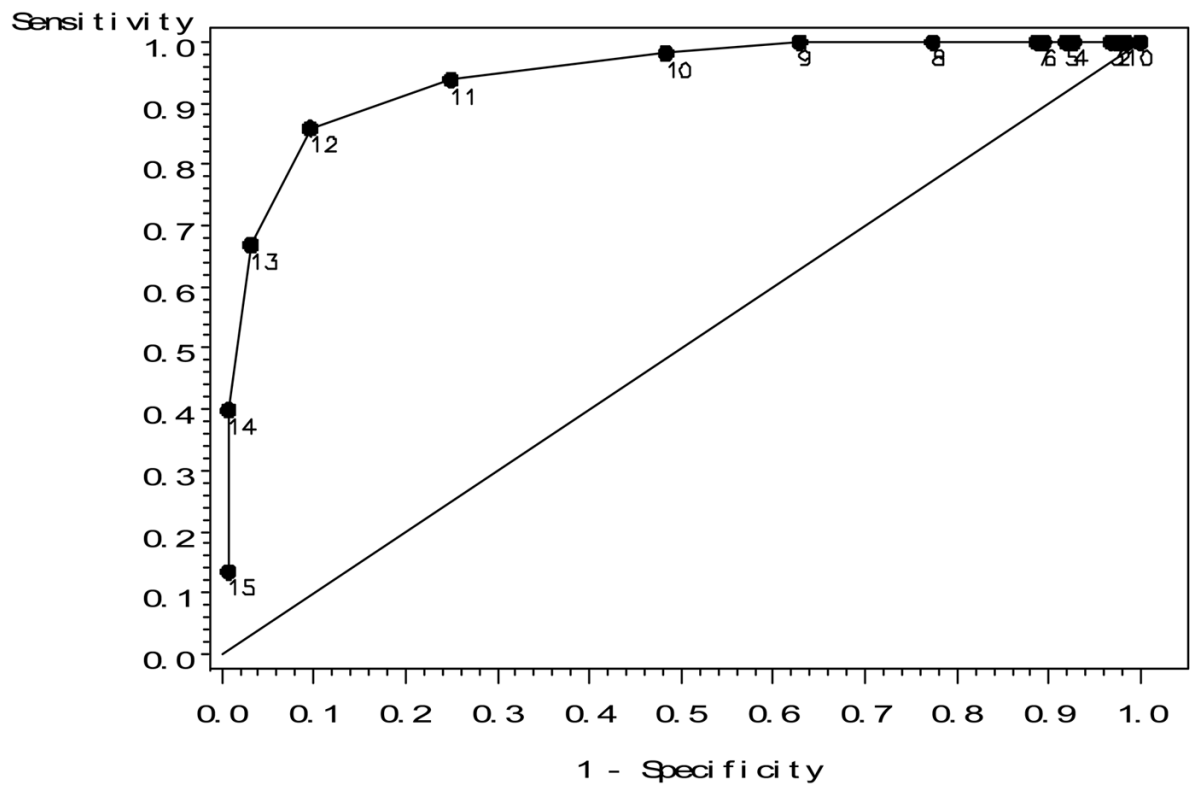
Figure 1.
ROC assessing the discriminative ability of the mini-PPT in predicting functional status calculated using the original PPT in the non-demented validation sample. Numeric labels indicate the mini-PPT score at that point on the ROC curve.

Table 1

Demographic characteristics of the study samples, mean ± S.D., or n(%)

	CDR = 0 Development	CDR = 0 Validation	CDR = 0.5	CDR = 1
Number	287	287	436	189
Women	159 (44.6)	177 (61.7)	241 (55.3)	118 (62.4)
Race				
African American	26 (9.1)	26 (9.1)	46 (10.6)	33(17.5)
White	256 (89.5)	258 (89.0)	387 (88.8)	155 (82.0)
Other races	4 (1.4)	3 (1.2)	3 (0.7)	1 (0.5)
Age, years	76.3 ± 8.4	75.8 ± 8.7	76.5 ± 7.3	77.4 ± 7.6
Education, years	14.7 ± 3.0	14.5 ± 3.2	13.8 ± 3.1	12.8 ± 3.3
Total PPT score	26.8 ± 5.0	27.3 ± 5.0	25.7 ± 5.3	23.6 ± 5.8
Functional (PPT > 28)	150 (52.3)	163 (56.8)	189 (43.4)	53 (28.0)
Robust (PPT > 32)	22 (7.7)	33 (11.5)	25 (5.7)	7 (3.7)

Table 2

Correlation with total PPT score, and AUC-ROC predicting functional status on the total PPT, for each candidate scale score

	CDR = 0, development sample (n = 287)			CDR = 0, validation sample (n = 287)		
PPT items in model	r with total PPT score	AUC	AUC 95% CI	r with total PPT score	AUC	AUC 95% CI
7,8,9	0.877[*]	0.927	0.900–0.955	0.869[*]	0.917	0.885–0.949
5,7,9	0.881[*]	0.900	0.866–0.933	0.855[*]	0.907	0.873–0.940
5,8,9	0.884[*]	0.924	0.895–0.953	0.884[*]	0.928	0.898–0.958
5,7,8,9	0.917[*]	0.941	0.916–0.966	0.907	0.940	0.913–0.966

[*]
p < 0.0001

Table 3

Statistics used in determining the optimal cutoff value on the mini-PPT for predicting functional status on the original PPT

Cutoff	CDR = 0, Validation			CDR = 0.5			CDR = 1		
	= 11	= 12	= 13	= 11	= 12	= 13	= 11	= 12	= 13
Sensitivity 95% CI	0.94 0.90–0.98	0.86 0.81–0.91	0.67 0.60–0.74	0.91 0.87–0.95	0.78 0.72–0.84	0.57 0.50–0.64	0.98 0.94–1.02	0.75 0.64–0.87	0.53 0.39–0.66
Specificity 95% CI	0.75 0.67–0.83	0.90 0.85–0.96	0.97 0.94–1.00	0.76 0.70–0.81	0.90 0.87–0.94	0.97 0.95–0.99	0.77 0.70–0.84	0.94 0.90–0.98	0.99 0.98–1.01
PPV	0.83 0.78–0.89	0.92 0.88–0.96	0.96 0.93–1.00	0.74 0.69–0.80	0.86 0.81–0.91	0.94 0.90–0.98	0.63 0.52–0.73	0.83 0.73–0.94	0.97 0.90–1.03
NPV	0.90 0.85–0.96	0.83 0.77–0.89	0.69 0.62–0.76	0.92 0.88–0.95	0.84 0.80–0.89	0.75 0.70–0.80	0.99 0.97–1.01	0.91 0.86–0.96	0.84 0.79–0.90
%CC	86	88	80	82	85	80	83	89	86

%CC = % correctly classified