

Published in final edited form as:

Arch Neurol. 2011 May ; 68(5): 607–614. doi:10.1001/archneurol.2011.88.

Telephone-based identification of MCI and dementia in a multicultural cohort

Jennifer J. Manly, Ph.D.^{1,2,5}, Nicole Schupf, Ph.D.^{1,2,4}, Yaakov Stern, Ph.D.^{1,2,5,6}, Adam M. Brickman, Ph.D.^{1,2,5}, Ming-X. Tang, Ph.D.^{1,3}, and Richard Mayeux, M.D., M.Sc.^{1,2,4,5,6}

¹Gertrude H. Sergievsky Center, Columbia University Medical Center, New York, NY

²Taub Institute for Research on Alzheimer's Disease and The Aging Brain, Columbia University Medical Center, New York, NY

³Department of Biostatistics, School of Public Health, Columbia University Medical Center, New York, NY

⁴Department of Epidemiology, School of Public Health, Columbia University Medical Center, New York, NY

⁵Department of Neurology, Columbia University Medical Center, New York, NY

⁶Department of Psychiatry, Columbia University Medical Center, New York, NY

Abstract

Objective—Telephone-based interviews can be used for screening and to obtain key study outcomes when participants in longitudinal studies die or cannot be seen in person, but must be validated among ethnically and educationally diverse people.

Method—The sample consisted of 377 (31% non-Hispanic white, 35% non-Hispanic black, and 34% Caribbean Hispanic) older adults. The validation standard was diagnosis of dementia and mild cognitive impairment (MCI) based on in-person evaluation. The Telephone Interview for Cognitive Status (TICS) and the Dementia Questionnaire (DQ) were administered within the same assessment wave.

Results—The sample included 256 (68%) people with normal cognition, 68 (18%) with MCI, and 53 (14%) with dementia. Validity of the TICS was comparable among non-Hispanic whites, non-Hispanic blacks, and Hispanics, but the DQ had better discrimination of dementia from those without dementia and with MCI among Whites than other groups. Telephone measures discriminated best when used to differentiate demented from nondemented participants (sensitivity/specificity for the TICS = 88%/87%; DQ = 66%/89%) and when used to differentiate cognitively normal participants from those with cognitive impairment (i.e., MCI and demented combined; sensitivity/specificity for the TICS = 73%/77%; DQ = 49%/82%). When demographics and prior memory test performance were used to calculate pre-test probability, consideration of the telephone measures significantly improved diagnostic validity.

Conclusions—The TICS has high diagnostic validity for identification of dementia among ethnically diverse older adults, especially when supported by the DQ and prior visit data. However, telephone interview data were unable to reliably distinguish MCI from normal cognition.

INTRODUCTION

Telephone-based assessment of cognitive status and functional decline is an alternative to in-person assessment in longitudinal studies of cognitive function and dementia of older adults¹⁻⁹. A telephone interview has become the primary modality of cognitive data collection in several epidemiological studies¹⁰⁻¹³, and is now frequently used as screen for clinical trials requiring participants with cognitive impairment¹⁴⁻¹⁸.

A recent study at Mayo Clinic¹⁹ found that although the modified TICS had 83.3% sensitivity and 81.6% specificity for separating demented from non-demented participants, and 83.3% sensitivity and 78.3% specificity for separating cognitively impaired participants from cognitively normal participants, the measure could not reliably distinguish participants with MCI from those with normal cognition or MCI from dementia. One limitation of this study was that the participants were almost exclusively White and well educated. Therefore, a central goal of the current study was to determine the accuracy of a telephone interview in classifying these groups among an ethnically and educationally diverse community-based sample.

The primary reason for adding the telephone interview to the assessment battery in our study of was to be able to derive key diagnostic classifications, i.e., normal cognition, MCI, or dementia, from the telephone-based data even when participants are unable or unwilling to be seen in person at a follow-up visit. However, this would require the instruments to validly distinguish MCI from normal cognition and MCI from dementia. This has been a challenge in prior studies; although several research groups have documented high specificity for MCI using telephone-based measures^{14, 17, 19}, only one study¹⁸ showed high sensitivity for distinguishing MCI from normal cognition. Since participants in our study were seen in-person at a prior assessment wave, we sought to determine whether the distinction of MCI from other classifications would improve if prior visit data was used along with data from the telephone interview.

METHOD

The Columbia University Institutional Review Board approved this project. All individuals discussed the study with trained research staff and provided written informed consent.

Participants

The current sample included 377 English and Spanish-speaking participants in a longitudinal study of aging, cognitive function, and dementia among Medicare-eligible older adults residing in neighborhoods in Northern Manhattan, New York. The current sample was drawn from a cohort resulting from two recruitment efforts, the first in 1992 (n=2,125) and the other in 1999 (n=2,183). The sampling strategies and recruitment outcomes of these two cohorts are detailed in prior publications^{20, 21}. Re-evaluations occur during follow-up waves that are spaced approximately 18 to 30 months apart.

Racial and ethnic group definitions

Ethnic group was determined by self-report using the format of the 2000 US Census. All individuals were first asked to report their race (i.e., American Indian/Alaska Native, Asian, Native Hawaiian or other Pacific Islander, Black or African American, or White), then, in a second question, were asked whether they were Hispanic.

Language of Administration

Evaluations were conducted in either English or Spanish, based on the subject's opinion of which language would yield the best performance. Examiners were balanced bilinguals, who spoke both English and Spanish daily with friends, family, and colleagues.

Telephone Interview

The validation study for the telephone interview was initiated during the 2005 – 2007 assessment wave of the cohort. The telephone interview was conducted during the same assessment wave, but independently from, the in-person visit by trained interviewers. On average, calls occurred 7.3 months after the in-person visit, with a standard deviation of 10.9 months. One participant had only the TICS because the call was interrupted and the participant could not be re-contacted. Of the participants for whom the DQ interview was conducted, eight did not have the TICS because they were not well enough to come to the phone (n=4), the participant died soon after the in-person visit (n=2), or because the call was interrupted (n=2).

Telephone Interview for Cognitive Status (TICS)

The TICS was administered and scored according to published procedures⁴. The TICS is modeled after the Mini-Mental State Examination, producing scores ranging from 0 to 41. High test-retest reliability^{1, 7, 22} has been demonstrated in several studies. The published Spanish language adaptation of the TICS was used among Spanish-speaking participants². Total score was used in the analyses.

Dementia Questionnaire

The DQ is a semi-structured interview that includes yes or no questions assessing cognitive complaints in the domains of memory, confusion, and spatial orientation (8 questions) and language/verbal expression (3 items), as well as questions assessing problems with daily function (6 items). This questionnaire has established reliability and validity with high sensitivity and specificity for the detection of dementia and AD³. Information about cognitive complaints and functional abilities could be provided by either the participant or an informant, as long as they were knowledgeable about the functional status and medical history of the participant. The 17 questions described above were summed to create a score representing total burden of cognitive complaints and functional problems.

In-person Evaluation

Medical history and a neurological and physical examinations was performed at the initial visit and each follow-up. A medical burden score was calculated as a sum of multiple non-psychiatric medical conditions, and included hypertension, diabetes, heart disease, stroke, arthritis, COPD or other pulmonary conditions, thyroid disease, liver disease, renal insufficiency, peptic ulcer disease, peripheral vascular disease, cancer, Parkinson's disease, multiple sclerosis, and essential tremor. Current depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale²³. The Disability and Functional Limitations Scale^{24, 25} was used to assess instrumental activities of daily living via self and informant report, as well as perceived difficulty with memory.

Neuropsychological measures included the Buschke Selective Reminding Test (SRT;²⁶), matching and delayed recognition conditions of a multiple choice version of the Benton Visual Retention Test (BVRT;²⁷), the Rosen Drawing Test²⁸, a 15-item Boston Naming Test²⁹, the Controlled Oral Word Association Test³⁰, the Category Fluency Test³¹, the Color Trails Test³², and the Similarities subtest from the Wechsler Adult Intelligence Scale - Revised³³.

Consensus Diagnosis

After each clinical assessment, a group of physicians and neuropsychologists reviewed the functional, medical, neurological, psychiatric, and neuropsychological data (but were blinded to TICS and DQ data) and reached a consensus regarding the presence or absence of dementia using DSM-III-R criteria³⁴. For follow-up evaluations, this group was shielded from the prior consensus diagnoses. If dementia was diagnosed, the etiology was determined using published research criteria for probable and possible AD³⁵, vascular dementia³⁶, Lewy Body Dementia³⁷, and other dementias. Mild Cognitive Impairment was not diagnosed in the consensus conference, but was retrospectively applied, based on the neuropsychological, functional, and memory complaint measures described above using standard criteria³⁸ among participants not diagnosed with dementia at the consensus conference²⁰.

Statistical analyses

Characteristics of the three diagnostic groups were compared using Chi-square tests and ANOVA, and correlations between measures and demographic variables were calculated. Receiver operating characteristic (ROC) curves were drawn for each of the telephone measures administered (TICS total score and DQ summary score), using four planned comparisons of interest: 1) non-demented vs. demented, 2) cognitively normal vs. cognitively impaired (i.e., MCI and dementia), 3) normal vs. MCI, and 4) MCI vs. dementia. Areas under the curve were calculated and compared for all participants and separately for non-Hispanic whites, non-Hispanic blacks, and Hispanics. The cutoff yielding the best sensitivity and specificity for the overall sample was determined, and the sensitivity and specificity, negative and positive predictive values, and likelihood ratios were calculated³⁹ for each telephone instrument for prediction of each of these diagnostic criteria in the entire sample and within the three primary racial/ethnic groups.

Binary logistic-regression models were constructed to estimate the pretest probabilities of each of the four comparisons outlined above with the in-person diagnostic classification as the dependent variable and the following independent variables: age, sex, race/ethnicity, years of education, and the delayed word list recall score from the SRT taken from the prior assessment. Post-test probabilities for each participant were then estimated for two scenarios: 1) when both TICS and DQ were available by adding the total scores from both measures to the model, and 2) when the participant was dead or too ill to come to the telephone and the TICS was not available, by adding to the model the total score from the DQ only. Predicted values from the models were then used to generate receiver-operating-characteristic curves. Areas under these curves were compared, and the differences were calculated with 95 percent confidence intervals⁴⁰.

RESULTS

Sample Characteristics

Demographic characteristics and scores on key study measures of participants with normal cognition, MCI, and dementia are described in Table 1. The majority (87.3%) of Hispanic participants in this study were immigrants from the Caribbean, including the Dominican Republic (59.5%), Cuba (17.5%), and Puerto Rico (10.3%). The MCI group included 44 participants with MCI with memory impairment (64.7%) and 24 with MCI without memory impairment (35.3%). Of the 53 demented participants, most were diagnosed with probable (n=33) or possible (n=16) Alzheimer's disease, but the sample also included two people with Parkinson's Disease Dementia, a participant with Vascular Dementia, and one with a diagnosis of Lewy Body Disease. The TICS total score was significantly correlated with age ($r = -0.37$; $p < .001$), years of education ($r = 0.51$; $p < .001$), prior visit SRT total recall ($r =$

0.56; $p < .001$), prior visit SRT delayed recall score ($r = 0.48$; $p < .001$), and depressive symptoms ($r = -0.28$; $p < .001$). Men (mean = 28.1, SD = 0.9) obtained slightly higher scores on the TICS than did women (mean = 25.6, SD = 0.6; $F(1,369) = 6.1$; $p = .014$). Differences between scores of non-Hispanic whites (mean = 30.4, SD = 7.5), non-Hispanic blacks (mean = 27.4, SD = 7.0), and Hispanics (mean = 21.8, SD = 10.2) were all significant (omnibus $F(1,365) = 32.3$; $p < .001$; all pairwise comparisons $p < .05$). The DQ summary score was significantly related to age ($r = 0.17$; $p < .001$), years of education ($r = -0.3$; $p < .001$), prior visit SRT total recall ($r = -0.32$; $p < .001$), prior visit SRT delayed recall score ($r = -0.26$; $p < .001$), depressive symptoms ($r = 0.39$; $p < .001$), and medical burden score ($r = 0.21$; $p < .001$). There were no significant differences in DQ summary score between men (mean = 3.6, SD = 0.4) and women (mean = 4.1, SD = 0.2; $F(1,376) = 1.1$; $p = .291$). While whites (mean = 3.1, SD = 2.9) and blacks (mean = 3.2, SD = 2.9) did not differ from each other, Hispanics (mean = 5.3, SD = 4.5) reported more problems on the DQ than each of the other two groups ($F(1,372) = 15.8$; $p < .001$). TICS and DQ scores were significantly correlated to each other ($r = -0.59$; $p < .001$).

Receiver Operating Characteristic (ROC) Curve Analyses

Figure 1a shows the ROC curve for separation of demented versus non-demented (normal cognition and MCI) participants. The area under the curve (AUC), diagnostic characteristics, and optimal cutoffs derived from the ROC analyses for the TICS (Table 2) and the DQ (Table 3) are shown for each of the comparisons among all participants and then separately for non-Hispanic whites, non-Hispanic blacks, and Hispanics. The ability of the TICS to discriminate between demented and non-demented participants was comparable across racial/ethnic groups, but the DQ's discrimination was higher among non-Hispanic Whites than among racial/ethnic minorities. Figure 1b depicts the ROC curves when MCI participants were combined with the demented participants to form a cognitively impaired group, and compared with participants with normal cognition. The AUCs for the TICS and the DQ was comparable across ethnic groups for this comparison (Tables 2 and 3).

We sought to determine the diagnostic accuracy of the telephone-based measures when making more subtle distinctions between participants with normal cognition and MCI and between MCI and dementia. Figure 1c depicts the ROC curves when demented participants were eliminated from the analysis and participants with MCI and normal cognition were compared. The AUC for both measures for this comparison was relatively low (.71 for the TICS and .58 for the DQ), but discriminability was similar across ethnic groups (Tables 2 and 3). We then determined the ability of the instruments to distinguish people with MCI from those with dementia, when participants with normal cognition were omitted from the analysis (Figure 1d). For this comparison, the AUC was .91 for the TICS, and was comparable across ethnic groups. The AUC was .81 for the DQ in the whole sample, but for this comparison, the DQ had better discrimination among non-Hispanic Whites than among ethnic minorities. Examination of the Odds Ratios in Tables 2 and 3 reveals that both the TICS and the DQ performed best in distinguishing the nondemented group (normal cognition and MCI combined) from the demented participants, and in distinguishing people with MCI from people with dementia.

Deriving an optimal classification algorithm

Figure 2 depicts ROC curves for pretest and post-test probabilities as calculated in the logistic regression models. As shown in Table 4, adding information gathered from the TICS and DQ to the pre-test model significantly improved the diagnostic performance for all key clinical outcomes in the study. For example, the addition of the TICS and DQ to the pre-test prediction of dementia vs. no dementia improved the AUC by 6.5%. The best diagnostic performance was in distinguishing non-demented from demented participants (AUC = .96)

and MCI from demented participants (AUC = .95) using demographic information, prior SRT delayed memory score, and both TICS and DQ. Accurate identification of MCI among non-demented participants was poor overall, even when both TICS and DQ were available (AUC = .75). Predicted classification as normal cognition, MCI, and dementia, using the optimal cutoffs for the predicted values from the models separating demented from nondemented participants and normal cognition from cognitive impairment, was compared to the observed diagnoses. The cutoffs correctly identified 66.5% of participants with normal cognition, 55.4% of those with MCI, and 92.7% of those with dementia.

Models using only the DQ summary score showed improved classification over pre-test probabilities when the goal was to distinguish demented from nondemented people, and cognitively impaired from cognitively normal people. However addition of the DQ summary score did not improve diagnostic accuracy over pre-test probabilities when the goal was to identify MCI among non-demented participants or to identify dementia among cognitively impaired participants (Table 4).

DISCUSSION

The sensitivity and specificity of the TICS and DQ was variable and depended on the diagnostic groups serving as the standard for comparison. There were no consistent racial/ethnic differences in the ability of the TICS to discriminate diagnostic classifications. However, in distinguishing demented people from nondemented (normal cognition and MCI combined), and people with dementia from people with MCI, the DQ performed better among non-Hispanic whites than among non-Hispanic blacks and Hispanics.

Used alone, the TICS had high sensitivity for distinguishing demented people from nondemented (normal cognition and MCI combined), and excellent specificity when distinguishing people with dementia from people with MCI. The DQ had lower sensitivity and higher specificity than the TICS for all comparisons, but was most valid when distinguishing demented people from those with MCI. The superior specificity of the DQ to the TICS is expected given the original purpose of developing the instruments: the DQ was designed to pick up on changes in memory and function that are specific to dementia and are not seen in normal aging or MCI.

Comparing likelihood ratios with the recent Mayo clinic study by Knopman et al.¹⁹, our use of the TICS overall and within each racial/ethnic group yielded superior performance to the TICS-M when distinguishing demented from nondemented participants (MCI and normal cognition combined), and MCI from dementia. Identification of cognitive impairment (MCI + dementia) from normal cognition was comparable to the Mayo TICS-M among non-Hispanic whites, non-Hispanic blacks, and Hispanics. Although similar among Hispanics, the Mayo study showed better accuracy for the separation of MCI from normal cognition than among the whites and blacks in the current study. Both studies had lower diagnostic validity for identification of MCI versus normal cognition than was reported by Cook et al. in a study of mostly White, community dwelling nondemented older adults¹⁸.

Our standard diagnostic algorithms for dementia and MCI require an in-person visit; therefore, diagnoses could not be derived for participants not seen due to death, moving out of the area, or refusal. The current study revealed that if a participant and/or informant can be reached by phone, presence of dementia can be estimated with moderate to high validity among non-Hispanic whites, non-Hispanic blacks, and Hispanics. Our analyses indicate that the diagnostic utility of telephone instruments will increase or decrease in response to variations in prevalence of cognitive impairment and dementia in the population, and thus will vary in cohorts that differ by age, racial, ethnic group, education, and other

demographic variables. Indeed, we found that adding age, sex, education, race/ethnicity, and prior memory scores to the data supplied by the TICS and DQ significantly improved the diagnostic accuracy of the telephone interview data. Even when the TICS was not used in the model, the nondemented/demented classification was highly accurate as predicted by demographics, DQ data and prior visit data.

It was hoped that the availability of the DQ, which taps into cognitive complaints and functional status, and prior visit memory test performance, when added to the direct cognitive assessment provided by the TICS, would improve the identification of MCI among nondemented participants. However none of the models tested were able to reliably differentiate MCI from normal cognition – this was true across all racial/ethnic groups. Addition of a delayed word list recall to the TICS may marginally improve identification of MCI, but prior research suggests that even with this added component, the classification rate may remain too low to advocate for the use of the TICS-M as a stand-alone measure for identification of MCI¹⁹.

Acknowledgments

This study was supported by National Institute on Aging grants P01-AG07232 (R. Mayeux) and R01-AG16206 (J. Manly). The authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Brandt J, Spencer M, Folstein M. The telephone interview for cognitive status. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*. 1988; 1:111–117.
2. Desmond DW, Tatemichi TK, Hanzawa L. The Telephone Interview for Cognitive Status (TICS): Reliability and validity in a stroke sample. *International Journal of Geriatric Psychiatry*. 1994; 9(10):803–807.
3. Kawas C, et al. A validation study of the Dementia Questionnaire. *Archives of Neurology*. 1994; 51:901–906. [PubMed: 8080390]
4. Brandt, J.; Folstein, MF. Telephone Interview for Cognitive Status, Professional Manual. Psychological Assessment Resources, Inc.; Odessa, FL: 2003.
5. Gallo J, Breitner J. Alzheimer's disease in the NAS-NRC Registry of aging twin veterans, IV. Performance characteristics of a two-stage telephone screening procedure for Alzheimer's dementia. *Psychol Med*. 1995; 25:1211–1219. [PubMed: 8637951]
6. Plassman B, et al. Properties of the telephone Interview for Cognitive Status: application in epidemiological and longitudinal studies. *Neuropsychiatry Neuropsychol Behav Neurol*. 1994; 7:235–241.
7. Welsh K, Breitner J, Magruder-Habib K. Detection of dementia in the elderly using telephone screening of cognitive status. *Neuropsychiatry Neuropsychol Behav Neurol*. 1993; 6:103–110.
8. De Jager C, Budge M, Clarke R. Utility of TICS-m for the assessment of cognitive function in older adults. *Int J Geriatr Psychiatry*. 2003; 18:318–24. [PubMed: 12673608]
9. Ellis R, et al. Diagnostic validity of the dementia questionnaire for Alzheimer disease. *Arch Neurol*. 1998; 55:360–5. [PubMed: 9520010]
10. Langa K, et al. Cognitive health among older adults in the United States and in England. *BMC Geriatrics*. 2009; 9(1):23. [PubMed: 19555494]
11. Herzog AR, Wallace RB. Measures of Cognitive Functioning in the AHEAD Study. *J Gerontol B Psychol Sci Soc Sci*. 1997; 52B:37–48. Special_Issue. [PubMed: 9215356]
12. Howard VJ, et al. The Reasons for Geographic and Racial Differences in Stroke Study: Objectives and Design. *Neuroepidemiology*. 2005; 25(3):135–143. [PubMed: 15990444]
13. Grodstein F, et al. Postmenopausal hormone therapy and cognitive function in healthy elderly women. *Journal of the American Geriatrics Society*. 2000; 48(7):746–752. [PubMed: 10894312]

14. Crooks V, et al. Validation of multi-stage telephone-based identification of cognitive impairment and dementia. *BMC Neurology*. 2005; 5(1):8. [PubMed: 15829005]
15. Lipton RB, et al. Screening for Dementia by Telephone Using the Memory Impairment Screen. *Journal of the American Geriatrics Society*. 2003; 51:1382–1390. [PubMed: 14511157]
16. Lautenschlager NT, et al. Effect of Physical Activity on Cognitive Function in Older Adults at Risk for Alzheimer Disease: A Randomized Trial. *JAMA*. 2008; 300(9):1027–1037. [PubMed: 18768414]
17. Lines CR, et al. Telephone screening for amnesic mild cognitive impairment. *Neurology*. 2003; 60(2):261–266. [PubMed: 12552041]
18. Cook SE, Marsiske M, McCoy KJM. The Use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the Detection of Amnesic Mild Cognitive Impairment. *J Geriatr Psychiatry Neurol*. 2009; 22(2):103–109. [PubMed: 19417219]
19. Knopman DS, et al. Validation of the Telephone Interview for Cognitive Status-modified in Subjects with Normal Cognition, Mild Cognitive Impairment, or Dementia. *Neuroepidemiology*. 34(1):34–42. [PubMed: 19893327]
20. Manly JJ, et al. Implementing Diagnostic Criteria and Estimating Frequency of Mild Cognitive Impairment in an Urban Community. *Archives of Neurology*. 2005; 62:1739–1746. [PubMed: 16286549]
21. Tang MX, et al. Incidence of Alzheimer's disease in African-Americans, Caribbean Hispanics and Caucasians in northern Manhattan. *Neurology*. 2001; 56:49–56. [PubMed: 11148235]
22. Plassman BL, et al. Properties of the telephone interview for cognitive status. Application in epidemiological and longitudinal studies. *Neuropsychiat Neuropsychol Beh Neurol*. 1994; (7): 235–241.
23. Kohout FJ, et al. Two Shorter Forms of the CES-D Depression Symptoms Index. *Journal of Aging and Health*. 1993; 5:179–193. [PubMed: 10125443]
24. Gurland B, et al. The SHORT-CARE: an efficient instrument for the assessment of depression, dementia, and disability. *Journals of Gerontology*. 1984; 39:166–169. [PubMed: 6699370]
25. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychology*. 1968; 114:797–811.
26. Buschke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology*. 1974; 24:1019–1025. [PubMed: 4473151]
27. Benton AL. *Revised Visual Retention Test* (4th ed.). 1974
28. Rosen, W. *The Rosen Drawing Test*. Veterans Administration Medical Center; Bronx, NY: 1981.
29. Kaplan, E.; Goodglass, H.; Weintraub, S. *Boston Naming Test*. Lea & Febiger; Philadelphia, PA: 1983.
30. Benton AL, Hamsher K. *Multilingual Aphasia Examination*. 1983
31. Goodglass, H.; Kaplan, D. *The assessment of aphasia and related disorders*. 2nd ed.. Lea and Febiger; Philadelphia: 1983.
32. D'Elia, LF., et al. *Color Trails Test Professional Manual*. Psychological Assessment Resources; Odessa, FL: 1994.
33. Wechsler, D. *Wechsler Adult Intelligence Scale-Revised*. The Psychological Corporation; New York, NY: 1981.
34. American Psychiatric, A.. *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*. American Psychiatric Press Inc.; Washington, DC: 1987.
35. McKhann G, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*. 1984; 34:939–944. [PubMed: 6610841]
36. Roman GC, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*. 1993; 43:250–260. [PubMed: 8094895]
37. McKeith IG, Perry EK, Perry RH. Report of the second dementia with Lewy body international workshop: diagnosis and treatment. *Consortium on Dementia with Lewy Bodies. Neurology*. 1999; 53:902–905. [PubMed: 10496243]

38. Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*. 2004; 256:183–194. [PubMed: 15324362]
39. Sackett, DL., et al. *Evidence-based medicine. How to practise and teach EBM*. 2nd edition. Churchill Livingstone; Edinburgh: 2000.
40. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988; 240:1285–1293. [PubMed: 3287615]

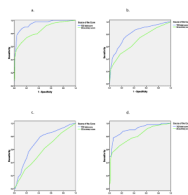


Figure 1. ROC curves for the TICS and the DQ for diagnosis of a) dementia vs. non-demented (normal cognition and MCI combined); b) cognitive impairment (dementia and MCI combined) vs. normal cognition; c) MCI vs. normal cognition, with demented participants eliminated from the analysis; and d) dementia vs. MCI, with participants with normal cognition removed from the analysis.

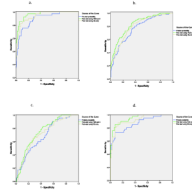


Figure 2.

ROC curves for the optimal prediction model for a) dementia vs. non-demented (normal cognition and MCI combined); b) cognitive impairment (dementia and MCI combined) vs. normal cognition; c) MCI vs. normal cognition, with demented participants eliminated from the analysis; and d) dementia vs. MCI, with participants with normal cognition removed from the analysis. Pre-test probabilities for each model are predicted values from binary logistic-regression models using age, sex, race/ethnicity, years of education, and the prior assessment delayed word list recall score from the SRT as predictors of the four target diagnostic states.

Table 1

Demographic Characteristics and Test scores of the Validation Sample

	Normal Cognition	MCI	Demented	Overall F or χ
N	256	68	53	
Age	80.3 (5.8)	82.2 (6.3)	85.9 (5.3)	20.4***
Years of Education	11.2 (4.6)	10.4 (4)	6.4 (4.9)	24.2***
N,% Female	66.0%	75.0%	69.8%	2.1
Race/Ethnicity				
N,% non-Hispanic white	33.2%	32.4%	15.1%	7.7*
N,% Hispanic	30.5%	23.5%	62.3%	22.5***
N,% non-Hispanic black	35.2%	44.1%	20.8%	7.5*
N, % other race	1.2%	0%	1.9%	1.8
Last SRT total recall	39.6 (9.9)	31.9 (8.3)	22.1 (7.8)	74.6***
Last SRT delayed recall	5.5 (2.7)	4.1 (2.2)	1.6 (1.9)	51.0***
CES-D score	1.3 (1.8)	1.2 (1.6)	2.3 (2.4)	6.1**
Medical burden score	3 (1.5)	3 (1.7)	3.3 (1.4)	0.7
Time between telephone interview and in-person visit	-6.4 (10.2)	-5.8 (7.5)	-13.3 (15.7)	9.4***
TICS total score	29.5 (6.2)	26.0 (5.0)	10.9 (10.2)	161.0***
DQ memory complaints score	1.3 (1.5)	2.1 (1.7)	4.3 (2.6)	66.1***
DQ language problems score	0.8 (0.9)	0.8 (1)	1.4 (1)	10.5***
DQ functional complaints score	0.9 (1.1)	0.8 (1.2)	3.2 (1.9)	78.2***
DQ summary score	2.9 (2.7)	3.7 (3.0)	8.8 (4.6)	78.6***

Notes:

*
p < .05**
p < .01***
p < .001

Table 2
Diagnostic characteristics of the TICS in identification dementia, cognitive impairment, and MCI.

Group A	Group B	Cut-Score	AUC	Sen.	Spec.	PPV	NPV	HR	OR	LR+	LR-
nondemented	demented	≤ 22	0.94	0.88	0.87	0.51	0.98	0.87	47.4	6.7	0.14
Non-Hispanic white			0.93	0.71	0.96	0.56	0.98	0.95	64.4	19.1	0.30
Non-Hispanic black			0.94	0.90	0.86	0.35	0.99	0.86	53.5	6.2	0.12
Hispanic			0.93	0.90	0.77	0.57	0.96	0.80	31.6	4.0	0.13
normal cognition	cognitive impairment	≤ 26	0.81	0.73	0.77	0.59	0.86	0.75	8.7	3.1	0.36
Non-Hispanic white			0.77	0.38	0.94	0.69	0.82	0.80	9.8	6.4	0.66
Non-Hispanic black			0.81	0.73	0.75	0.57	0.86	0.74	7.9	2.9	0.37
Hispanic			0.86	0.94	0.58	0.58	0.94	0.72	20.2	2.2	0.11
normal cognition	MCI	≤ 29	0.71	0.79	0.58	0.34	0.91	0.62	5.2	1.9	0.36
Non-Hispanic white			0.72	0.59	0.78	0.41	0.88	0.74	5.0	2.6	0.53
Non-Hispanic black			0.76	0.87	0.58	0.41	0.93	0.65	9.0	2.1	0.23
Hispanic			0.72	0.94	0.36	0.23	0.96	0.46	8.3	1.5	0.18
MCI	dementia	≤ 19	0.91	0.78	0.94	0.90	0.85	0.87	55.3	13.2	0.24
Non-Hispanic white			0.86	0.71	1.00	1.00	0.92	0.93	-	-	0.29
Non-Hispanic black			0.89	0.70	0.93	0.78	0.90	0.88	32.7	10.5	0.32
Hispanic			0.89	0.81	0.88	0.93	0.70	0.83	29.2	6.5	0.22

Abbreviations: AUC, area under the curve; Sen, sensitivity or the proportion in Group B correctly identified in each comparison using the cutoff; Spec, specificity or the proportion in Group A correctly identified using the cutoff; PPV, positive predictive value or the probability that people with scores below the cutoff have the Group B diagnosis; NPV, negative predictive value or the probability that people with scores above cutoff do not have Group B diagnosis; HR, hazard ratio; OR, odds ratio; LR+, likelihood ratio for a positive result or how many times greater the probability of a score below the cutoff is among people in Group B than among Group A; LR-, likelihood ratio for a negative result or how much the odds of having the Group B diagnosis decrease when the test score is above the cutoff.

Table 3

Diagnostic characteristics of the DQ in identification of dementia, cognitive impairment, and MCI.

Group A	Group B	Cut-Score	AUC	Sen.	Spec.	PPV	NPV	HR	OR	LR+	LR-
nondemented	demented	≤ 7	0.84	0.66	0.89	0.49	0.94	0.85	15.0	5.8	0.38
Non-Hispanic white			0.98	0.75	0.97	0.67	0.98	0.96	104.0	26.8	0.26
Non-Hispanic black			0.76	0.55	0.88	0.30	0.95	0.85	9.1	4.7	0.51
Hispanic			0.79	0.67	0.78	0.52	0.87	0.75	7.3	3.1	0.42
normal cognition	cognitive impairment	≤ 6	0.70	0.49	0.82	0.57	0.77	0.72	4.4	2.8	0.62
Non-Hispanic white			0.71	0.40	0.88	0.55	0.81	0.76	5.0	3.4	0.68
Non-Hispanic black			0.65	0.41	0.87	0.59	0.76	0.73	4.6	3.1	0.68
Hispanic			0.71	0.59	0.71	0.57	0.73	0.67	3.6	2.1	0.57
normal cognition	MCI	≤ 6	0.58	0.29	0.82	0.31	0.81	0.71	1.9	1.7	0.86
Non-Hispanic white			0.62	0.18	0.88	0.29	0.81	0.74	1.7	1.5	0.93
Non-Hispanic black			0.61	0.33	0.87	0.45	0.80	0.73	3.3	2.5	0.77
Hispanic			0.53	0.38	0.71	0.21	0.85	0.66	1.5	1.3	0.88
MCI	dementia	≤ 9	0.81	0.55	0.96	0.91	0.73	0.78	26.2	12.4	0.47
Non-Hispanic white			0.97	0.63	1.00	1.00	0.88	0.90			0.38
Non-Hispanic black			0.70	0.36	0.97	0.80	0.81	0.80	16.6	10.9	0.66
Hispanic			0.75	0.58	0.88	0.90	0.50	0.67	9.5	4.6	0.48

Abbreviations: AUC, area under the curve; Sen, sensitivity or the proportion in Group B correctly identified in each comparison using the cutoff; Spec, specificity or the proportion in Group A correctly identified using the cutoff; PPV, positive predictive value or the probability that people with scores below the cutoff have the Group B diagnosis; NPV, negative predictive value or the probability that people with scores above cutoff do not have Group B diagnosis; HR, hazard ratio; OR, odds ratio; LR+, likelihood ratio for a positive result or how many times greater the probability of a score below the cutoff is among people in Group B than among Group A; LR-, likelihood ratio for a negative result or how much the odds of having the Group B diagnosis decrease when the test score is above the cutoff.

Table 4

Pre- and post-test probabilities for key diagnostic outcomes.

	Non-demented vs. demented			Cognitively normal vs. cognitively impaired				
	AUC	Vs. pre-test model	99% CI	p	AUC	Vs. pre-test model	99% CI	p
Pre-test model	.899				.765			
Post-test model: TICS and DQ	.964	.065	.026 - .104	< .001	.829	.064	.03 - .098	< .001
Post-test model: DQ only	.940	.042	.04 - .08	0.03	.797	.032	.01 - .062	0.042
Cognitively normal vs. MCI								
	Cognitively normal vs. MCI			MCI vs. demented				
	AUC	Vs. pre-test model	99% CI	p	AUC	Vs. pre-test model	99% CI	p
Pre-test model	.692				.869			
Post-test model: TICS and DQ	.750	.058	.017 - .099	0.005	.945	.075	.014 - .136	0.015
Post-test model: DQ only	.718	.026	.009 - .062	0.148	.917	.047	.007 - .102	0.089

Note: Pretest model included age, sex, education, race/ethnicity, SRT delay as predictors of the diagnostic outcome