Dementia = (MC)²: A Four-Item Screening Test for Mild Cognitive Impairment and Dementia

Vanja C. Douglas, MD,
UCSF Department of Neurology, Neurohospitalist Program, 505 Parnassus Ave, M798, Box 0114, San Francisco, CA 94143-0114, Tel: 415-514-1575, Fax: 415-476-3428

John Neuhaus, PhD,
UCSF Department of Neurology, San Francisco, USA

Julene Johnson, PhD,
UCSF Department of Neurology, San Francisco, USA

Caroline Racine, PhD,
UCSF Department of Neurology, San Francisco, USA

Bruce L. Miller, MD, and
UCSF Department of Neurology, San Francisco, USA

S. Andrew Josephson, MD
UCSF Department of Neurology, San Francisco, USA

Vanja C. Douglas: Vanja.douglas@ucsf.edu

Abstract

To develop a practical informant-based screening tool that reliably identifies patients with mild cognitive impairment (MCI) and dementia, we analyzed data from a sample of patients and normal controls seen in a memory clinic. All subjects were evaluated with the Clinical Dementia Rating scale (CDR). Individual CDR responses were dichotomized and entered into a forward stepwise multivariable logistic regression model. Four independent predictors of MCI and dementia thus identified were combined into a prediction rule that was validated in a separate cohort drawn from the same clinic. Using a cut point of 2 or more positive responses to the four questions, the final prediction rule had sensitivity of 95% (95% CI 92 – 97%) for MCI or dementia, and a specificity of 91% (95% CI 86 – 95%). When applied to the validation cohort, the sensitivity for MCI or dementia was 96% (95% CI 94 – 98%), and the specificity was 96% (95% CI 92 – 98%). Using both cohorts, the positive likelihood ratio for MCI or dementia was 15.6 (95% CI 14.0 – 17.3) and the negative likelihood ratio 0.05 (95% CI 0.04 – 0.07). This tool has the potential to identify patients who warrant further cognitive evaluation in busy outpatient or emergency department settings.

Search terms

Dementia; Mild Cognitive Impairment; Screening; Diagnosis

INTRODUCTION

A rapid screening test for mild cognitive impairment (MCI) or early dementia would be very useful in a variety of clinical settings. For example, such a screening test could help busy primary care physicians decide whether or not to pursue a more in-depth cognitive evaluation in older adults. In the emergency department, it could help physicians triage admitted older adult patients to a geriatric care unit or to specialty consultation.

There are two general methods to screen for MCI and dementia: patient performance-based testing and informant interview. Multiple performance-based tests exist, the most widely used of which is the Mini-Mental State Examination.[1] These tests are impractical for a variety of reasons: they are time-consuming, they have a ceiling effect making them insensitive for MCI and early dementia, and they may be affected by an individual's level of education.[2–7] Informant interview bypasses the problem of a ceiling effect, is very sensitive to subtle changes such as MCI, and is independent of inter-individual characteristics such as level of education. The current gold standard informant-based assessment of dementia is the Clinical Dementia Rating scale (CDR), which is a semi-structured interview validated in large multicenter clinical studies.[8–10] However, the CDR, like most informant interviews, requires trained practitioners and at least 30 minutes to administer.[11]

The ideal practical screening test for busy practitioners should be easily administered and scored in addition to being accurate for detecting both MCI and dementia. Several informant-based screening tests shorter than the CDR exist but none meet these criteria. An eight-item questionnaire, the AD8, reliably identifies people with early dementia but has not been validated for MCI, and a still shorter screening test would be better and more easily applied broadly by a diverse set of busy clinicians.[12] Another informant interview consisting of five questions about memory and activities of daily living showed good accuracy for detecting MCI when measured against the CDR as a gold standard, but it has yet to be validated in a prospective cohort and has a complicated scoring system that does not lend itself to easy memorization by clinicians.[13]

The purpose of this study was to develop a simple and quick screening tool for identifying MCI and dementia in older adults. We analyzed responses to individual questions from the CDR among patients and controls to determine which were most predictive of MCI and dementia. We then combined these questions into a prediction rule that was validated in a separate cohort.

METHODS

Subjects and Data Collection

We identified all patients diagnosed with MCI, frontotemporal lobar degeneration (FTLD), vascular dementia (VaD), possible or probable Alzheimer disease (AD) and dementia with Lewy bodies (DLB) at a tertiary care memory clinic between 1999 and 2009. Those who were evaluated with the CDR were included in the analysis. Diagnosis was determined by a neurologist specialized in neurodegenerative disease after a comprehensive neurological and neuropsychiatric evaluation and was based on established criteria. MCI was diagnosed based on Petersen criteria.[14] Controls were recruited from the community; all controls evaluated with the CDR during the time period of the study were included in the analysis. Demographic and clinical data were collected at the time of each evaluation. This study was approved by the UCSF Committee on Human Research.

All included patients and controls were evaluated with the CDR. At our institution the CDR is administered by a trained nurse or social worker as a structured informant interview based on the Washington University, St. Louis worksheets.[8] It consists of 46 questions that assess a range of cognitive and social functions including memory, orientation, judgment and problem solving, community affairs, household chores and hobbies, and personal care that allow calculation of the CDR score. Most questions allow only a "yes/no" response, such as "Does he/she have a problem with his/her memory or thinking?" Questions that provide more than "yes/no" response options were dichotomized into normal/not normal by following the rule that the best answer was categorized as the response expected for a patient with normal cognition and all other responses categorized as not normal. For example,

response options for the question "How often does he/she know the exact month?" include "usually," "sometimes," "rarely," and "don't know." Using the rule, "usually" was categorized as normal, "sometimes" and "rarely" as not normal, and "don't know" as missing. Open-ended uestions not amenable to dichotomization were excluded from analysis (Supplementary Table 1).

Statistics

We created our model using a derivation cohort consisting of patients and controls evaluated prior to 2006; it was validated in the patients and controls seen in 2006 and later. The derivation cohort consisted of 104 patients with MCI, 309 with dementia, and 180 controls; the validation cohort consisted of 114 patients with MCI, 309 with dementia, and 267 controls. Baseline demographic characteristics between patients with MCI, those with dementia, and controls were compared using the chi square test for categorical variables and analysis of variance for continuous variables.

Of the 37 CDR questions amenable to dichotomization, 33 were entered into a forward stepwise logistic regression model testing their ability to predict MCI and dementia in the derivation cohort. Items were entered in the order they appear on the CDR questionnaire. Four CDR questions were left out of the model a priori because they were predicated on the patient working, driving, or living in a nursing home and therefore generated a large number of missing responses. Predictor variables were selected for inclusion in the model if $p < 0.1$. The model found five questions independently predictive of MCI and dementia. All possible combinations of four of these five questions were used to create five separate four-question prediction rules. A single point was assigned for each question endorsed by the informant as below normal function; thus the maximum score was four. Receiver operating characteristic (ROC) curves were created comparing the performance of each rule against the clinical diagnosis for each subject. The model that maximized sensitivity and specificity at a single cut-point was chosen and tested in the validation cohort. Confidence intervals for sensitivity and specificity as well as likelihood ratios were calculated according to published methods. [15] All statistical tests were performed using STATA 10.0.[16]

RESULTS

Patients and controls were similar in terms of age, gender, race, and educational background, although the patients were slightly older (Table 1). A range of dementia subtypes were represented, including 208 (67%) patients with AD, 44 (14%) with FTLD, 40 (13%) with DLB, and 17 (6%) with VaD. Patients with MCI almost universally had CDR scores of 0.5 (97%). Most patients with dementia were mildly impaired, with total CDR scores of 0.5 in 31% of cases and 1 in 45%. Higher scores of 2 or 3 were seen in 22% and 2% of cases respectively.

Five questions from the CDR were found to be independently predictive of MCI and dementia in the multivariable logistic regression model (Table 2). This enabled the creation of five separate prediction rules using all possible combinations of four independently predictive questions. There was no significant difference in the area under the ROC curves (AUC) for the four rules with the largest AUC ($p = 0.25$; Figure 1A). Including all five questions in the prediction rule did not result in a better model ($p = 0.21$ for difference in AUC). A cut-point using a score of 2 or more led to the best test characteristics for each rule (Table 3). Because the AUC was similar for each, the model that maximized both sensitivity and specificity was chosen as the final prediction rule.

The final prediction rule included two questions about memory and thinking, one regarding the ability to handle complicated financial transactions, and one concerning the ability to

understand situations and explanations (Table 4). Using a score of 2 or more as the cut-point for classifying a subject as having MCI or dementia, this rule had a sensitivity of 95% (95% CI 92 – 97%) and specificity of 91% (95% CI 86 – 95%), with a positive likelihood ratio of 10.8 (95% CI 9.2 – 12.8) and a negative likelihood ratio of 0.06 (95% CI 0.04 – 0.09). It was slightly more sensitive for detecting dementia (97%, 95% CI 95 – 99%) than MCI (87%, 95% CI 79 – 93%) (Table 5). The area under the ROC curve was 0.97 (95% CI 0.96 – 0.99) (Figure 1B).

This prediction rule was validated in a separate cohort. The validation cohort was comparable to the derivation cohort in terms of number of patients (n=423) and controls (n=267), as well as their gender (49% vs. 48% male; p = 0.68) and level of education (mean = 16 ± 6 vs. 16 ± 7 years; p = 0.25). Subjects in the validation cohort were slightly younger (67 ± 10 vs. 70 ± 11 years; p < 0.001) and there was slightly more racial diversity (84% vs. 88% white; p = 0.04). The number of patients with AD (205), DLB (24), and VaD (10) were similar (p > 0.05 for each) but there were more with FTLD (70; p = 0.01). The distribution of CDR scores was similar to the derivation cohort, with 33% of dementia patients having a score of 0.5, 44% a score of 1, 16% a score of 2, and 5% a score of 3 (p = 0.07 for any difference). Almost all patients with MCI had a score of 0.5 (93%).

The prediction rule performed well in the validation cohort. The sensitivity for detecting MCI or dementia using a cut-point of ≥ 2 was 96% (95% CI 94 – 98%) and the specificity 96% (95% CI 92 – 98%). In the validation cohort, the rule was sensitive for both MCI (94%, 95% CI 88 – 98%) and dementia (97%, 95% CI 94 – 98%). When restricted to patients with mild impairment (CDR scores of 0, 0.5, or 1), the rule performed similarly, detecting MCI or dementia with a sensitivity of 95% (95% CI 92 – 97%) and a specificity of 96% (95% CI 92 – 98%). The mean $D=(MC)^2$ score was higher for demented patients than for those with MCI (Supplementary Table 2). Combining all patients and controls from both the derivation and validation cohorts, the overall positive likelihood ratio for MCI or dementia was 15.6 (95% CI 14.0 – 17.3) and the negative likelihood ratio was 0.05 (95% CI 0.04 – 0.07).

DISCUSSION

Using a large cohort of patients seen in a tertiary memory clinic, we identified four items from the CDR, the current gold-standard informant-based tool for diagnosing dementia, that were most independently predictive of a diagnosis of MCI or dementia and combined these into a screening tool that is easy to remember, easy to administer, and easy to score. The combination of these four items was both sensitive and specific for detecting MCI and dementia in older adults using the gold standard of a diagnosis made by a behavioral neurologist.

The likelihood ratios are robust in both directions. Likelihood ratios are common in medical parlance and allow the users of diagnostic tests to translate the results into clinical meaning for individual patients.[17] All test results depend on the pretest probability of the disease in the patient being tested. For screening tests, the pretest probability is simply the prevalence of the disease (which often depends on a few risk factors such as age). Likelihood ratios allow one to convert pretest probability into post-test probability. For example, the pretest probability (prevalence) of dementia in a patient aged 71–79 in the United States is 5%.[18] If a patient in this age range tested positive using our screening tool, the likelihood of that patient having MCI or dementia would be increased to 45%. One arrives at this figure by converting the pretest probability into pretest odds (5 ÷ 95 = 0.053), multiplying the pretest odds by the positive likelihood ratio to get the post-test odds (0.053 × 15.6 = 0.82), and then converting the post-test odds to post-test probability (0.82 ÷ (0.82 + 1) = 0.45 = 45%). Conversely if the patient tested negative, the likelihood of MCI or dementia would fall to

0.3%. This would allow a clinician to stratify patients into those who need further evaluation and those who do not.

While informant-based tools can be impractical in some settings, such as when a patient is not accompanied by a caregiver or partner, they have the advantage of lacking the ceiling effect that is common to many performance-based tests for cognitive dysfunction and are therefore much more sensitive for the detection of MCI and early dementia. They are also not influenced by inter-individual characteristics such as level of education. Because early detection is the goal of screening, an informant-based tool such as that presented here is likely to be more powerful than performance-based tests that require a greater level of impairment before they identify an abnormality. It is also possible that a combination of a brief performance-based measurement and an informant interview may prove more effective than ether tool alone.[19]

This study has several strengths. They include the large cohort in which the screening tool was generated, the variety of dementia subtypes represented, the inclusion of a large number of subjects with MCI and mild dementia, and the ability to validate the findings in a separate, equally robust, cohort. Additionally the screening tool was tested against the current clinical diagnostic gold standard: a thorough multidisciplinary evaluation by a neurologist and neuropsychologist with neuroimaging and laboratory data. The screening tool presented here has similarities to the AD8, a validated eight-item informant based screening tool designed to detect very mild dementia.[12] Two of the items in the AD8 were also identified in $D=(MC)^2$: consistent problems with memory and thinking and difficulty handling complicated financial affairs. Although sensitivity may be sacrificed with $D=(MC)^2$ compared to the longer AD8, its brevity and ease of use may make it more practical in the clinical setting.

Although the test characteristics of this screening tool are promising, several limitations are important to recognize. First, the tool was not developed in the setting in which it would likely be applied. Subjects were all patients referred to a memory clinic; in general such patients or at least their family members have identified a problem leading to referral and therefore the likelihood they would endorse items on the CDR is high. It is possible that when applied in the community among patients whose cognitive impairment is not yet recognized, the questions might be endorsed at a lower rate leading to decreased sensitivity. The population available for study was relatively homogenous, with only a small percentage of Asians and blacks. Our results did not change when the analysis was restricted to whites (data not shown), but there were not enough subjects to compare the results across racial and ethnic groups. Therefore our results may not be generalizable to other ethnicities. Future studies testing the validity of this screening tool should include a more diverse population.

Secondly, the lack of prospective validation is a limitation. The four questions were culled from over 30 CDR questions that were all asked during the course of a structured interview; it is not clear that these four questions would be answered in the same way if asked in isolation. It is also possible that specificity is overestimated because the screening tool was derived from the CDR, which also factors into the ultimate clinical diagnosis. Therefore, subjects who endorse items on the CDR are more likely to be diagnosed with MCI or dementia. This element of circularity could only be eliminated with a prospective validation study in which the four questions identified here are asked independently of the gold standard in the context of an evaluation in a primary care of other non-specialized setting where the tool is likely to be of most practical use. The magnitude of this potential limitation is unclear because the diagnosis of MCI or dementia against which the screening tool was compared was made by a behavioral neurologist after consideration of a large amount of information of which the CDR score is only a minor part: a thorough interview of the patient

and collateral source by a physician, a neurologic examination, a complete battery of neuropsychiatric testing, brain imaging, and laboratory analysis. In addition, the diagnosis of MCI also requires poor performance on at least one delayed memory test, so an abnormal CDR (including a score of 0.5) is indicative of impaired function, but does not guarantee a diagnosis of MCI. Prospective validation is a critical step in the development of any screening tool and this must be undertaken before tool presented here can be used clinically. Finally, some items on the CDR could not be included because they were not amenable to dichotomization; it is possible information that could have led to a more robust model was lost as a result.

The four-item informant-based screening tool for MCI and dementia presented here would be simple to administer and easy to remember and score, which would make it practical for everyday use. It is not meant to provide a diagnosis of dementia but merely to identify those at-risk patients who would benefit from referral for a more thorough evaluation. Its performance must be evaluated further in an independent and prospective fashion prior to being adopted for widespread use.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. Nov; 1975 12(3):189–198. [PubMed: 1202204]

2. Kalbe E, Calabrese P, Schwalen S, et al. The Rapid Dementia Screening Test (RDST): a new economical tool for detecting possible patients with dementia. Dement Geriatr Cogn Disord. 2003; 16(4):193–199. [PubMed: 14512713]

3. Kalbe E, Kessler J, Calabrese P, et al. DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. Int J Geriatr Psychiatry. Feb; 2004 19(2):136–143. [PubMed: 14758579]

4. Leopold NA, Borson AJ. An alphabetical 'WORLD'. A new version of an old test. Neurology. Dec; 1997 49(6):1521–1524. [PubMed: 9409339]

5. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am Geriatr Soc. Sep; 1992 40(9):922–935. [PubMed: 1512391]

6. Brown J, Pengas G, Dawson K, et al. Self administered cognitive screening test (TYM) for detection of Alzheimer's disease: cross sectional study. BMJ. 2009; 338:b2030. [PubMed: 19509424]

7. Wilber ST, Carpenter CR, Hustey FM. The Six-Item Screener to detect cognitive impairment in older emergency department patients. Acad Emerg Med. Jul; 2008 15(7):613–616. [PubMed: 18691212]

8. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. Nov; 1993 43(11):2412–2414. [PubMed: 8232972]

9. Morris JC, Ernesto C, Schafer K, et al. Clinical dementia rating training and reliability in multicenter studies: the Alzheimer's Disease Cooperative Study experience. Neurology. Jun; 1997 48(6):1508–1510. [PubMed: 9191756]

10. Schafer KA, Tractenberg RE, Sano M, et al. Reliability of monitoring the clinical dementia rating in multicenter clinical trials. Alzheimer Dis Assoc Disord. Oct-Dec; 2004 18(4):219–222. [PubMed: 15592134]

11. Burke WJ, Miller JP, Rubin EH, et al. Reliability of the Washington University Clinical Dementia Rating. Arch Neurol. Jan; 1988 45(1):31–32. [PubMed: 3337672]

12. Galvin JE, Roe CM, Xiong C, et al. Validity and reliability of the AD8 informant interview in dementia. Neurology. Dec 12; 2006 67(11):1942–1948. [PubMed: 17159098]

13. Li M, Ng TP, Kua EH, et al. Brief informant screening test for mild cognitive impairment and early Alzheimer's disease. Dement Geriatr Cogn Disord. 2006; 21(5–6):392–402. [PubMed: 16645272]

14. Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol. Mar; 1999 56(3):303–308. [PubMed: 10190820]

15. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol. 1991; 44(8):763–770. [PubMed: 1941027]

16. StataCorp. Stata Statistical Software: Release 10. College Station: TSL; 2007.

17. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. Jama. Mar 2; 1994 271(9):703–707. [PubMed: 8309035]

18. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. Neuroepidemiology. 2007; 29(1–2):125–132. [PubMed: 17975326]

19. Galvin JE, Roe CM, Morris JC. Evaluation of cognitive impairment in older adults: combining brief informant and performance measures. Arch Neurol. May; 2007 64(5):718–724. [PubMed: 17502471]
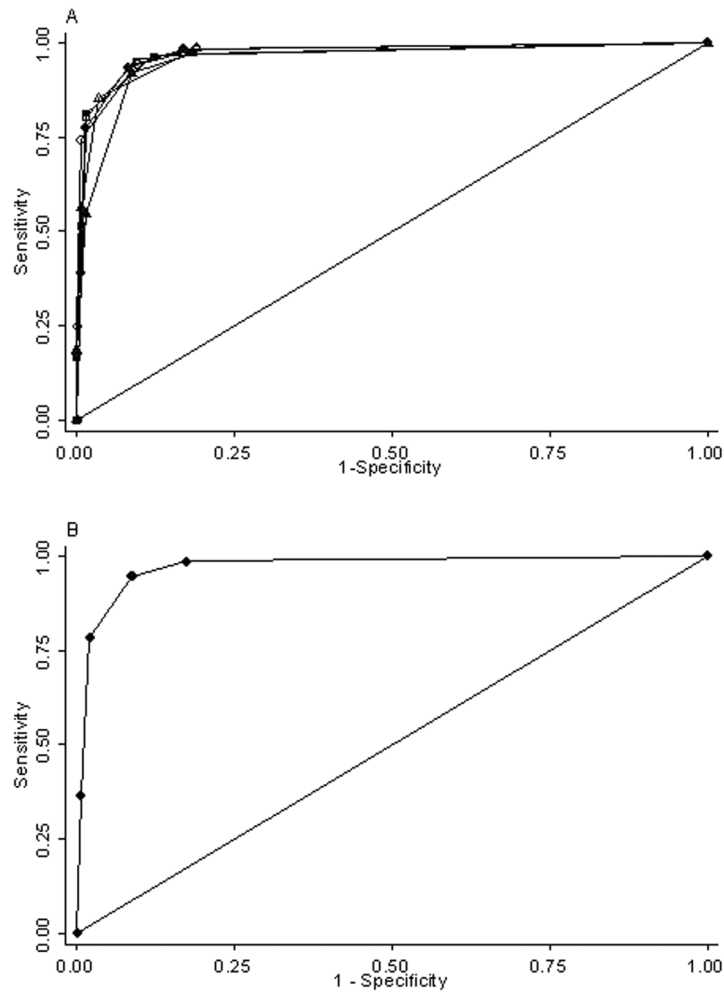
Figure 1. ROC Curves for Potential Screening Tools and for D=(MC)^2
A. Receiver operating characteristic (ROC) curves for each of five screening tools using the clinical dementia rating questions independently predictive of mild cognitive impairment or dementia. B. ROC curve for the model with the highest sensitivity and specificity (AUC = 0.97).

Table 1

Baseline Characteristics Among Patients and Controls

	MCI n=104	Dementia n=309	Controls n=180	p
Age (mean ± SD)	73 ± 9	72 ± 10	65 ± 13	<0.001
Gender (% male)	42	50	50	0.18
Race (%)				
White	89	85	94	0.25
Asian	4	6	3	
Black	3	4	2	
Other	4	5	1	
Years of Education (mean ± SD)	17 ± 9	15 ± 8	17 ± 3	0.07

MCI: mild cognitive impairment

Table 2

Independent Predictors of MCI or Dementia From the CDR

	Odds Ratio (95% CI)		p
1. Problem with memory or thinking	36	(6.5 – 199)	<0.001
2. If yes, then a consistent problem	3.4	(0.9 – 12)	0.062
3. Subject appears ill to casual observer	16	(1.6 – 166)	0.019
4. Unable to handle complicated financial transactions	30	(1.3 – 682)	0.035
5. Unable to understand situations or explanations	7.8	(0.7 – 87)	0.096

MCI: mild cognitive impairment; CDR: clinical dementia rating scale

Table 3

Test Characteristics of Five Prediction Rules Using a Cut-point of ≥ 2

	Sensitivity for MCI or Dementia	Specificity	LR+	LR−	Area Under ROC Curve[†]
A. Questions 1, 2, 4, 5[*]	94.6%	91.3%	10.8	0.06	0.97
B. Questions 1, 2, 3, 4	94.3%	90.5%	10.0	0.06	0.97
C. Questions 1, 2, 3, 5	92.1%	91.2%	10.5	0.09	0.96
D. Questions 1, 3, 4, 5	85.2%	96.6%	25.2	0.15	0.97
E. Questions 2, 3, 4, 5	81.4%	98.7%	60.6	0.19	0.97

[*] See Table 2 for definitions of each question. The rule incorporating questions 1, 2, 4, and 5 maximized sensitivity and specificity.

[†] p = 0.25 for difference in area under the curve between A, B, D, and E.

MCI: mild cognitive impairment; CDR: clinical dementia rating scale; LR: likelihood ratio; ROC: receiver operating characteristic

Table 4

Dementia = $(MC)^2$: Scoring of Prediction Rule

		0 points	1 point each
Memory	Does he/she have a problem with his/her memory or thinking?	No	Yes
Consistency	If yes, is this a consistent problem (as opposed to inconsistent)?	No	Yes
Money	Rate his/her ability to handle complicated financial or business transactions (e.g., balance checkbook, pay bills)	No Loss	Some Loss or Severe Loss
Comprehension	Can he/she understand situations or explanations?	Usually	Sometimes or Rarely

Total (a score of 2 or greater is predictive of MCI or dementia): 0 – 4

Table 5

Performance of Dementia = (MC)^2 for Detecting MCI and Dementia by Total Score

Score	Sensitivity for MCI	Sensitivity for Dementia	Combined Sensitivity	Specificity	LR+ (MCI or Dementia)	LR− (MCI or Dementia)
≥ 1	98%	99%	99%	83%	5.7	0.02
≥ 2	87%	97%	95%	91%	10.8	0.06
≥ 3	49%	89%	78%	98%	38.8	0.22
4	12%	46%	37%	99%	54.4	0.64

MCI: mild cognitive impairment; LR: likelihood ratio