

# Detecting Cognitive Impairment in Primary Care: Performance Assessment of Three Screening Instruments

Douglas L. Trenkle<sup>a,\*</sup>, William R. Shankle<sup>b,c</sup> and Stanley P. Azen<sup>d</sup>

<sup>a</sup>Maine Coast Memorial Hospital, Ellsworth, ME, USA

<sup>b</sup>Medical Care Corporation, Irvine, CA, USA

<sup>c</sup>Department of Cognitive Sciences, University of California, Irvine, CA, USA

<sup>d</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California, CA, USA

**Abstract.** Early detection of Alzheimer's disease and related disorders (ADRD) is important, especially in primary care settings. We compared performances of two common screening tests, the Mini-Mental State Exam (MMSE) and Clock Drawing Test (CDT), with that of the MCI Screen (MCIS) in 254 patients over 65. None had previous diagnosis of ADRD, and 81% were asymptomatic by Functional Assessment Staging Test (FAST) (FAST = 1). 215 patients completed all screening tests – 141 had  $\geq 1$  abnormal result, 121/141 completed standardized diagnostic assessment, and the remaining 74/215 (34%) screened entirely normally and weren't further evaluated. Potential bias due to unevaluated cases was statistically adjusted. Among diagnosed cases: AD = 43%, cerebrovascular disease = 36%, other causes = 21%. Bias-adjusted MCI prevalence for FAST stages 1 and 1–3 were 13.9–20.3% and 23.0–28.3%. Bias-adjusted results for the CDT, MMSE and MCIS were: clinical diagnosis validity (kappa statistic) = {−0.02 ( $p = 0.61$ ), 0.06 ( $p = 0.23$ ), 0.92 ( $p < 0.0001$ )}; sensitivity = {59%, 71%, 94%}; specificity = {39%, 36%, 97%}; overall accuracy = {54%, 62%, 96%}; positive predictive value = {16%, 17%, 86%}; and negative predictive value = {83%, 87%, 96%}.

The MMSE and CDT were not valid for early detection, while the MCIS had high validity and accuracy in the primary care cohort.

**Keywords:** Normal aging, mild cognitive impairment, Alzheimer's disease, prevalence, sensitivity, accuracy, early detection, primary care, receiver operating characteristic, validity

## INTRODUCTION

Postmortem studies of persons with either normal aging or mild cognitive impairment (MCI), particularly the amnesic type [4,10,13,17,31,47,53,62,67,68], have demonstrated the presence of several types of neuropathology – most often in the medial temporal lobe [49] but also described in the visual association cortex [37] – including neurofibrillary changes of

Alzheimer's disease (AD), Lewy bodies, cerebral infarcts, and hippocampal sclerosis [7,49]. These studies support a transition of AD and related disorders (ADRD) from normal aging to MCI to dementia. Detecting AD earlier means treatment can begin when there is less neuropathology [36], offering greater potential to improve symptoms and/or delay disease progression [11,19,34,55,59,70,74]. Earlier AD detection also allows non-pharmacological interventions that can delay institutionalization [39,56], plus give patients and families adequate time to make future plans.

Currently,  $> 2/3$  of ADRD patients in primary care settings are moderate-to-severely demented when first diagnosed [7,23,63,69], which corresponds to 9–11

\*Corresponding author: Douglas L. Trenkle, D.O., Maine Coast Memorial Hospital, 50 Union Street, Ellsworth, ME 04605, USA. Tel.: +1 207 664 5722; Fax: +1 207 664 5490; E-mail: dtrenkle@mainehospital.org.

years after symptom onset [51]. Such delayed detection is partly due to lack of sensitive and specific screening tools that can be easily incorporated into a physician's practice [2].

The most commonly used screening tools are the Mini Mental Status Exam (MMSE) and the Clock Drawing Test (CDT). The MMSE tests mainly left hemisphere function, including repetition, naming, orientation, attention, recall and comprehension. The CDT measures attentiveness, visual-spatial ability and constructional praxis but does not specifically measure episodic memory. Although these tests help monitor dementia, their sensitivities for detecting MCI have not been well established in primary care settings.

The MCI Screen (MCIS) is an examiner-administered test of executive function, working and episodic memory, and language, derived from the National Institute of Aging CERAD (Consortium to Establish a Registry for AD) battery. Overall accuracy in differentiating normal aging from MCI and mild dementia is 97.3% and 99%, respectively [61].

This article reports the screening performance of the CDT, MMSE and MCIS in a primary care practice cohort, staged by the Functional Assessment Staging Test (FAST) [58]. Patients had not been previously detected to have MCI or dementia, and were part of Maine's Hancock County Aging Project, established in 2004.

## METHODS

### *Study design and cohort*

Participants of this study were chosen from a cohort of 254 patients seen at a primary care practice in Hancock, Maine during 2004–2006. Eligibility criteria included a) no prior diagnosis or detection of cognitive impairment or dementia; and b)  $\geq 65$  years old or subjective decline in cognition or cognitively related function. In compliance with the Helsinki guidelines for human subject research, all patients were legally competent to provide informed consent, and were informed of the following: a) study's purpose, aims, potential risks and benefits; b) the confidential manner in which the data would be collected and handled to protect privacy; c) they could abstain or withdraw at any time from the study without affecting their physician-patient relationship; and d) they would not be identified by any published results.

Of 254 patients, 39 did not complete all screening tests (*incompletely screened*) – 14 (36%) had no subjective or objective difficulties (FAST stage 1); 8 (21%) had subjective complaints (FAST stage 2); 2 (5%) had objective complex functional impairment (FAST stage 3); 2 (5%) had mild dementia (FAST stage 4); and 13 (33%) were not FAST staged. The remaining 215 patients were staged using the FAST, and completed the CDT, MMSE, and MCIS screening tests (*completely screened*).

### *Screening instruments*

The MMSE is the most widely used examiner-administered test to detect dementia in primary care [21]. The MMSE has a 30 point scale (30 = perfect score), and has published data for normal aging, MCI and dementia. Among published studies of MCI vs. normal aging, an MMSE score of  $> 28$  for normals gave the highest sensitivity when the specificity was fixed at 90% or higher [5,24].

The CDT, another widely used examiner-administered test, has a 15 point scale (15 = perfect score) [65,66]. The CDT measures abilities that are largely untested by the MMSE. A CDT score of  $> 12$  for normals gave the highest sensitivity when the specificity was fixed at 90% or higher [18].

The MCIS, an online based examiner-administered test, was developed based on the 10-word recall test from National Institute of Aging AD/DR CERAD neuropsychological test battery [72,73] combined with sophisticated scoring algorithms that analyze the pattern of words recalled by each subject. The MCIS consists of three 10-word learning trials followed by an immediate recall task at each trial, self-evaluation of subject's ability to recall, distraction task using triadic comparison of animal names, delayed free- and cued-recall tasks and free untrained recall of animal names from the triadic comparison task. It takes about 10 minutes to administer, and receiver operating characteristic (ROC) analysis has shown that the MCIS has overall classification accuracy of 97.3% and 99 % in differentiating normals from MCI and from mild dementia, respectively [61]. See Appendix A for further details.

### *Staging of functional and AD severity: The FAST*

The FAST is a well validated, brief AD staging tool that reflects progressive, cognitively dependent, functional decline [3,51,58]. Patients with amnesic or non-amnesic definitions of MCI can be functionally sub-

classified as either FAST stages 1, 2 or 3. FAST stages 1-3 are characterized as follows: FAST 1: no subjective decline and no objective impairment in cognitively related function; FAST 2: subjective decline in cognitively related function without no objective evidence; and FAST 3: objective evidence of decline in cognitively related function (e.g., performance of complex tasks at work, home or socially) but no decline in instrumental activities of daily living (e.g. handling finances, shopping, preparing meals). FAST stages 4, 5, 6 and 7 correspond to mild, moderate, moderately severe and severe dementia. Normally aging individuals are functionally classified as FAST stage 1 or 2 depending upon the absence or presence of subjective functional decline plus the absence of objective, cognitively related functional decline. Normally aging individuals should also have no progressive cause of cognitive impairment. The mean duration of FAST stage 1 is not well characterized; that of FAST stage 2 is 15 years [52]; and that of FAST stage 3 is 7 years [51]. This functional subdivision of MCI using FAST staging helps characterize MCI more precisely.

#### Screening test validation: Clinical diagnosis

A standardized diagnostic assessment was performed on 141 *completely screened* patients with  $\geq 1$  abnormal screening result to identify underlying causes that would validate the screening result of cognitive impairment (i.e., a true positive result). If no underlying cause could be identified, the impaired screening test result was classified as a *false positive*. Clinical diagnosis of an identifiable cause of cognitive impairment was therefore the reference standard for determining screening test validity. To date, no patients have been autopsied for pathologic confirmation.

Clinical practice restrictions typical for community settings precluded diagnostically assessing the 74 patients who screened normally on all three tests (*unconfirmed normal*). Also among 141 *completely screened* patients with  $\geq 1$  abnormal screening results, 20 of them failed to comply in completing their diagnostic assessment (*unconfirmed impaired*). The *unconfirmed impaired* patients were excluded from further analysis because they had both normal and abnormal screening test outcomes, thus requiring diagnosis to distinguish true from false positive cases. Of the 121 completely screened and diagnostically assessed patients (*confirmed diagnosis*), clinical diagnosis was *confirmed impaired* in 51 and *confirmed normal* in 70. Standardized diagnostic assessment consisted of MRI, labora-

tory studies (B12, folate and thyroid deficiencies, homocysteine, lipids, C-reactive protein, metabolic and hematologic disorders), family history of a dementing disorder and risk factor assessment history. NINDS-ADRDA criteria for AD [29,38], ADDTC criteria for vascular dementia [12], and DSM-IV criteria for other dementing disorders [1], were applied to these data to determine etiologic diagnosis. The requirement of dementia was relaxed because  $> 90\%$  of the diagnostically assessed patients were not demented. For the *confirmed impaired* patient group, clinical diagnosis was approximately distributed as follows: possible AD = 43%, cerebrovascular disease = 36%, cerebrovascular disease + AD = 1%, alcohol abuse=9%, traumatic brain injury = 4%, depressive pseudodementia=3%, metabolic disorders=2%, Parkinson's disease = 1%, and other causes=1%. For screening test validation, clinical diagnosis was classified as either *confirmed impaired* ( $N = 51$ ) or *confirmed normal* ( $N = 70$ ). The validation results for each screening test are shown in Table 2.

#### STATISTICAL ANALYSES

To test for demographic differences among clinical diagnosis groups (incompletely screened, confirmed normal, confirmed impaired, unconfirmed normal, and unconfirmed impaired), we performed a multinomial logistic regression using clinical diagnosis group as the dependent variable, and using age, gender and education as the independent variables (Stata 9.0, ologit). ANOVA and chi-square tests were performed to identify age, gender, and educational differences across the clinical diagnosis groups. Significant group differences in age or education were evaluated by pair-wise one-way ANOVA comparisons with Bonferroni adjustment for multiple comparisons (Stata 9.0, oneway). Significant gender differences between groups were evaluated by subtracting their chi-square values and assessing significance with 1 degree of freedom.

Validity was determined by measuring agreement (kappa coefficient) between each screening test and clinical diagnosis group with and without adjusting for selection bias (see *Adjustment for Potential Selection Bias* section).

Screening performance was evaluated by computing sensitivity, specificity, positive and negative predictive values between each screening test and clinical diagnosis group. In addition, non-parametric ROC curves were generated to measure the 95% confidence inter-

vals of overall accuracy of each screening test relative to clinical diagnosis (Stata 9.0, roctab). The area under an ROC curve is an unbiased estimate of a screening test's overall accuracy.

#### *Adjustment for potential selection bias*

We used two strategies to adjust for possible selection bias among the 74 *unconfirmed normal* FAST stage 1 patients (not diagnostically assessed due to practice restrictions), then computed the upper and lower bounds of the screening test performance statistics on the full sample of 195 patients (74 *unconfirmed normal* + 121 *confirmed diagnosis* patients). For Strategy 1, we applied the highest reported MCI prevalence over 65 years old (0.20) [20,?,28,32] to the FAST stage 1 patient group ( $N = 158$ ), which gave a maximum of 32 false negatives. As there were 22 FAST stage 1 *confirmed impaired* patients, 10 *unconfirmed normal* patients with the lowest screening test scores were assigned a clinical diagnosis of *impaired*, and the remaining 64 were assigned a clinical diagnosis of *normal*. For Strategy 2, all 74 *unconfirmed normals* were assigned a clinical diagnosis of normal because of the very low prevalence of MCI among FAST stage 1 individuals. The results from Strategies 1 and 2 were then combined with those of the 121 *confirmed diagnosis* patients, and the upper and lower bounds of the screening test performance statistics were computed. To compute positive and negative predictive values (PPV, NPV), a mid-range prevalence of 16% for MCI—mild dementia  $\geq 65$  years old was used. To estimate the lower and upper bounds of MCI prevalence in this cohort (*impaired* patients in FAST stages 1–3), we used the total number of impaired cases estimated by Strategies 2 and 1 respectively, and divided by the total number of FAST stages 1–3 patients in the sample.

#### *Instrument reliability*

MMSE test-retest reliabilities range from 0.54 for community dwelling elderly subjects with a high school education to 0.65 among those without one [30]. CDT test-retest reliability when administered in a hospital-based geriatric outpatient clinic was 0.82 [71].

The MCIS test-retest reliability was assessed on an independent sample of 30 normal aging-to-moderately demented subjects, who were retested within six months and had not undergone a change in treatment. A neuropsychologist administered the pencil-and-paper CERAD Wordlist (from which part of the MCIS was

derived). On a separate occasion, an office assistant administered the online MCIS to the same patients. Test administrator order was randomized. The within-subject differences were computed for the CERAD Wordlist sub-scores (immediate and delayed free recall trials). Test-retest reliability was evaluated using the hypothesis of no within-subject differences between test sub-scores on the two occasions (paired t-tests with Bonferroni adjustment for multiple comparisons). Within-subject test-retest correlation for each sub-score was also computed. Mean within-subject differences on immediate free recall trials 1–3 and delayed free recall were 0.23, 0.80, 0.30, and 0.33 words. Paired t-tests showed that these differences did not significantly differ from zero (all  $p > 0.027$  based on the Bonferroni-adjusted significance level = 0.013). Test-retest correlations for immediate and delayed free recall trials were 0.79, 0.59, 0.75 and 0.90, and for all trials combined was 0.83. The MCIS test-retest reliability was therefore better than both the MMSE and CDT.

#### *Comparison with more contemporary screening tests*

While the present study compares the MCIS to the most widely used tests in clinical practice, other recently developed screening tests may also be useful once they have been properly validated. An unbiased benchmark for screening test performance is the area under the ROC curve, which measures overall accuracy. For reference, Appendix B lists the overall ROC accuracies in discriminating normal aging from MCI for all peer-reviewed studies listed in PubMed on or before January 15, 2007, that provided such data. This review found that the MCIS had the highest overall ROC accuracy plus the largest validation sample.

## RESULTS

For the 39 *incompletely screened* patients, ANOVA and chi-square tests showed that their age, gender and education did not significantly differ from any of the *completely screened* groups ( $p > 0.31$ ), except that the *incompletely screened* were more educated than the *confirmed impaired* ( $p = 0.014$ , mean (SD) = 13.8 (3.2) vs. 11.7 (3.2) years education). While the overall chi square comparing the proportions of *incompletely vs. completely screened* patients across the FAST stages was marginally significant ( $p = 0.043$ ), no group differences were found in any single FAST stage ( $p > 0.05$ , Bonferroni adjusted). The *incompletely and com-*

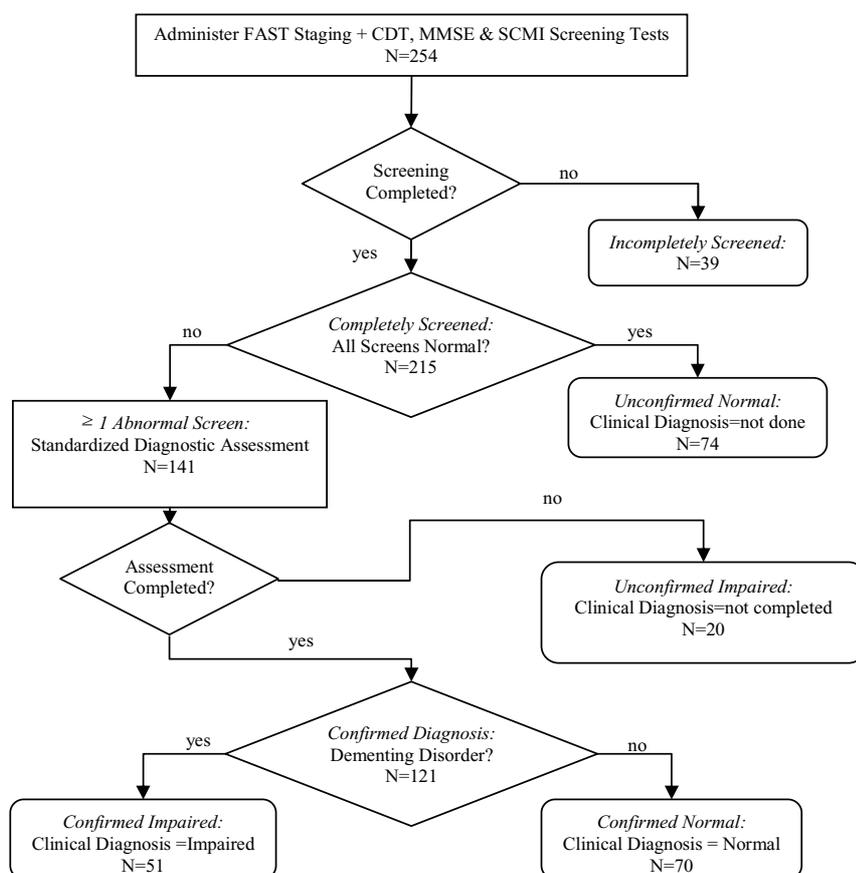


Fig. 1. Process for Screening and Clinical Diagnosis.

pletely screened groups therefore did not substantively differ demographically or by level of cognitively related functional impairment.

Figure 1 shows that the 215 completely screened patients were sub-classified as *unconfirmed normal* ( $N = 74$ ), *unconfirmed impaired* ( $N = 20$ ), *confirmed impaired* ( $N = 51$ ) and *confirmed normal* ( $N = 70$ ). After excluding the 20 *unconfirmed impaired* patients (see Methods), screening performance statistics were computed on the *confirmed diagnosis* (CDx) sample ( $N = 121$ ), and on the *confirmed diagnosis* plus *unconfirmed normal* sample ( $N = 195$ ) using bias adjustment Strategies 1 and 2.

Table 1 shows the sample characteristics of the 215 completely screened patients. Multinomial logistic regression found that age and gender did not differ across clinical diagnosis groups, and that, while education significantly differed across them ( $p < 0.001$ ), the magnitude of the effect was small (only 3% of total variance explained). One-way ANOVA plus chi-square tests for demographic differences between clinical diagnosis

groups found: a) *unconfirmed normals* were younger and more educated than all other groups ( $p < 0.004$ ); and b) *confirmed normals* were more educated than the *confirmed impaired* ( $p=0.012$ ). One-way ANOVA of CDT scores across clinical diagnosis groups found: a) *confirmed normals* did not differ from either *confirmed* or *unconfirmed impaired* groups ( $p > 0.33$ ); and b) *unconfirmed normals* scored higher than all other groups ( $p < 0.02$ ). One-way ANOVA of MMSE scores across clinical diagnosis groups found: a) *confirmed normals* scored higher than *confirmed impaired* patients ( $p=0.008$ ), did not differ from *unconfirmed impaired* patients ( $p = 0.32$ ), and scored lower than *unconfirmed normals* ( $p < 0.001$ ); and b) *unconfirmed normals* scored higher than all other groups ( $p < 0.001$ ). One-way ANOVA of MCIS scores across clinical diagnosis groups found: a) no differences between *confirmed* and *unconfirmed normals* ( $p=0.49$ ); and b) both *confirmed* and *unconfirmed normals* scored higher than the *confirmed* and *unconfirmed impaired* groups ( $p < 0.001$ ).

Table 1  
Patient Demographic, Screening Test, and FAST Staging Characteristics and Bias-Adjusted Numbers of Normal and Impaired Patients from Non-Diagnosed Patients Stratified by Clinical Diagnosis Group

| Measure            | Clinical diagnosis Group (n = 215)* |                           |                               |                             | p§       |
|--------------------|-------------------------------------|---------------------------|-------------------------------|-----------------------------|----------|
|                    | Unconfirmed Normal (n = 74)         | Confirmed Normal (n = 70) | Unconfirmed Impaired (n = 20) | Confirmed Impaired (n = 51) |          |
| Age (years)        | 68.5 ± 8.4 <sup>a</sup>             | 73.4 ± 11.4 <sup>b</sup>  | 77.6 ± 5.7 <sup>b</sup>       | 75.0 ± 9.9 <sup>b</sup>     | < 0.0001 |
| Gender: Female     | 40 (54.1%) <sup>a</sup>             | 29 (41.4%) <sup>a</sup>   | 6 (30.0%) <sup>a</sup>        | 24 (47.1%) <sup>a</sup>     | 0.20     |
| Education (years)  | 14.9 ± 2.9 <sup>a</sup>             | 13.2 ± 3.3 <sup>b</sup>   | 12.7 ± 2.4 <sup>b,c</sup>     | 11.7 ± 3.2 <sup>c</sup>     | < 0.0001 |
| CDT (#/15)         | 14.3 ± 0.8 <sup>a</sup>             | 11.4 ± 3.0 <sup>b</sup>   | 12.5 ± 2.1 <sup>b</sup>       | 10.8 ± 3.4 <sup>b</sup>     | < 0.0001 |
| MMSE (#/30)        | 29.5 ± 0.5 <sup>a</sup>             | 27.7 ± 1.8 <sup>b</sup>   | 26.7 ± 3.4 <sup>b,c</sup>     | 26.5 ± 2.9 <sup>c</sup>     | < 0.0001 |
| MCIS (-1 to 4)     | 2.4 ± 0.7 <sup>a</sup>              | 2.2 ± 0.7 <sup>a</sup>    | 0.4 ± 0.6 <sup>b</sup>        | 0.4 ± 0.8 <sup>b</sup>      | < 0.0001 |
| FAST Stage 1       | 74 (100.0%) <sup>a</sup>            | 62 (88.6%) <sup>a</sup>   | 14 (70.0%) <sup>a</sup>       | 22 (43.1%) <sup>b</sup>     | < 0.004  |
| FAST Stage 2       | 0 (0.0%) <sup>a</sup>               | 8 (11.4%) <sup>b</sup>    | 6 (30.0%) <sup>b</sup>        | 11 (21.6%) <sup>b</sup>     | < 0.0003 |
| FAST Stage 3       | 0 (0.0%) <sup>a</sup>               | 0 (0.0%) <sup>a</sup>     | 0 (0.0%) <sup>a</sup>         | 10 (19.6%) <sup>b</sup>     | < 0.0001 |
| FAST Stage 4       | 0 (0.0%) <sup>a</sup>               | 0 (0.0%) <sup>a</sup>     | 0 (0.0%) <sup>a</sup>         | 8 (15.7%) <sup>b</sup>      | < 0.0001 |
| CDx: NL/IMP        | -                                   | 70/0                      | -                             | 0/51                        | N/A      |
| Strategy 1: NL/IMP | 55/19                               | 70/0                      | -                             | 0/51                        | N/A      |
| Strategy 2: NL/IMP | 74/0                                | 70/0                      | -                             | 0/51                        | N/A      |

Note: Continuous variables are summarized as mean ± standard deviation; discrete variables are summarized as frequency and column percentage.

§p-values were obtained from tests for overall clinical diagnosis group differences (chi-square for discrete and one-way ANOVA for continuous variables). Means and percentages with different letters (a,b,c, ...) are significantly different, according to pairwise comparisons with Bonferroni corrections for multiple comparisons.

\*Unconfirmed normal category consists of those who screened entirely normally and did not have a diagnostic assessment due to clinical practice restrictions. Unconfirmed impaired category consists of patients who screened abnormally on ≥ 1 test and did not complete diagnostic assessment. Confirmed normal and confirmed impaired categories are those who screened abnormally on ≥ 1 test and completed their diagnostic assessment.

CDx = Confirmed diagnosis sample (N = 121 = 51 confirmed impaired + 70 confirmed normal).

NL/IMP = numbers of patients assigned a clinical diagnosis of normal and impaired, respectively.

'-' = categories not used for analysis.

N/A = not applicable.

Table 2 summarizes, overall, for FAST stage 1 and for FAST stages 1–3, the validity (kappa coefficient) of the CDT, MMSE and MCIS tests in predicting clinical diagnosis group with and without bias adjustment strategies. The lower and upper bounds of the kappa values of the CDT, MMSE and MCIS for FAST stage 1 were respectively -0.06–0.16 ( $p < 0.74$ ), -0.04–0.19 ( $p < 0.67$ ), and 0.68–0.89 ( $p < 0.0001$ ); for FAST stages 1–3, the corresponding kappa values were -0.07–0.20 ( $p < 0.80$ ), 0.007–0.27 ( $p < 0.46$ ), and 0.79–0.92 ( $p < 0.0001$ ); and for FAST stages 1–4 the corresponding kappa values were -0.02–0.25 ( $p < 0.61$ ), 0.06–0.33 ( $p < 0.23$ ), and 0.82–0.93 ( $p < 0.0001$ ). The CDT and MMSE therefore had zero to low validity for early detection within any FAST stage, while the MCIS had high validity for early detection within all FAST stages.

Table 3 shows the performance of each screening test with and without bias adjustment. The lower and upper bounds of sensitivity (Sn) for the CDT, MMSE and MCIS were respectively 49–59%, 59–71%, and

79–94%, while those for specificity (Sp) were 39–70%, 36–69%, and 97–99%. At a mid-range prevalence for MCI–mild dementia of 0.16, the corresponding values for PPV were respectively 16–27%, 17–30%, and 86–95%, while those for NPV were respectively 83–90%, 87–93%, and 96–99%. For the unbiased estimate, overall accuracy using ROC methods, the widest 95% confidence intervals for the CDT, MMSE and MCIS were respectively 43–78%, 51–83%, and 91–100%. The upper and lower bounds of the mean overall accuracies for the CDT, MMSE and MCIS were respectively 53.7–69.4%, 61.5–75.8%, and 95.5–96.2%. The MCIS had a significantly higher overall accuracy than both the CDT and MMSE ( $p < 0.00001$ ), while the accuracies of the CDT and MMSE did not significantly differ ( $p = 0.22$ ) (Stata 9.0, roccomp). A point estimate of screening performance was obtained by holding specificity constant at 95.0%, and comparing the lower and upper bound sensitivities of the CDT, MMSE and MCIS tests, which were respectively 8.8–13.6%, 21.9–26.6%, and 86.0–94.1%. After bias adjustment (see Methods), the

Table 2  
Validity of the CDT, MMSE and MCIS Stratified by FAST Staging and Adjusted For Possible Selection Bias from Non-Diagnosed Patients

| Kappa Coefficient ± Standard Error |          |     |                                     |                                     |                                     |
|------------------------------------|----------|-----|-------------------------------------|-------------------------------------|-------------------------------------|
| FAST Staging                       | Strategy | N   | CDT§                                | MMSE**                              | MCIS§§                              |
| FAST Stages 1–4                    | CDx      | 121 | −0.02 ± 0.09<br>( <i>p</i> = 0.61)  | 0.06 ± 0.08<br>( <i>p</i> = 0.23)   | 0.92 ± 0.09<br>( <i>p</i> < 0.0001) |
| FAST Stages 1–4                    | 1        | 195 | 0.16 ± 0.07<br>( <i>p</i> = 0.01)   | 0.23 ± 0.07<br>( <i>p</i> = 0.0004) | 0.82 ± 0.07<br>( <i>p</i> < 0.0001) |
| FAST Stages 1–4                    | 2        | 195 | 0.25 ± 0.07<br>( <i>p</i> = 0.0001) | 0.33 ± 0.07<br>( <i>p</i> < 0.0001) | 0.93 ± 0.07<br>( <i>p</i> < 0.0001) |
| FAST Stages 1–3                    | CDx      | 113 | −0.07 ± 0.09<br>( <i>p</i> = 0.80)  | 0.007 ± 0.08<br>( <i>p</i> = 0.46)  | 0.91 ± 0.09<br>( <i>p</i> < 0.0001) |
| FAST Stages 1–3                    | 1        | 187 | 0.11 ± 0.07<br>( <i>p</i> = 0.07)   | 0.17 ± 0.07<br>( <i>p</i> = 0.008)  | 0.79 ± 0.07<br>( <i>p</i> < 0.0001) |
| FAST Stages 1–3                    | 2        | 187 | 0.20 ± 0.07<br>( <i>p</i> = 0.002)  | 0.27 ± 0.07<br>( <i>p</i> < 0.0001) | 0.92 ± 0.07<br>( <i>p</i> < 0.0001) |
| FAST Stage 1                       | CDx      | 84  | −0.06 ± 0.09<br>( <i>p</i> = 0.74)  | −0.04 ± 0.08<br>( <i>p</i> = 0.67)  | 0.87 ± 0.11<br>( <i>p</i> < 0.0001) |
| FAST Stage 1                       | 1        | 158 | 0.04 ± 0.08<br>( <i>p</i> = 0.26)   | 0.07 ± 0.08<br>( <i>p</i> = 0.17)   | 0.68 ± 0.08<br>( <i>p</i> < 0.0001) |
| FAST Stage 1                       | 2        | 158 | 0.16 ± 0.07<br>( <i>p</i> = 0.01)   | 0.19 ± 0.07<br>( <i>p</i> = 0.003)  | 0.89 ± 0.08<br>( <i>p</i> < 0.0001) |

Note: Strategy = bias adjustment strategy to account for possible diagnostic bias due to the 74 unconfirmed normal patients.

§CDT: Clock Drawing Test.

\*\*MMSE: Mini-Mental Status Exam.

§§MCIS: MCI Screen.

CDx = Confirmed diagnosis sample (*N* = 121: 51 confirmed impaired + 70 confirmed normal).

Strategy 1: bias adjustment used a maximum published MCI prevalence of 0.20 to estimate maximum false negatives among FAST stage 1 patients. This adjustment assigned 10 unconfirmed normals a diagnosis of impaired.

Strategy 2: bias adjustment assuming no false negatives among the unconfirmed normals.

lower and upper bounds of MCI prevalence for FAST stages 1 and 1–3 were respectively 13.9–20.3% and 23.0–28.3%.

## DISCUSSION

### Major findings

The MCIS results of this study independently validate its previously reported overall accuracies of 97% and 99%, respectively, in discriminating MCI and mild dementia from normal aging [61]. Both with and without bias adjustment, the MCIS had high validity, sensitivity, specificity, overall accuracy, PPV and NPV for MCI and mild dementia at all levels of functional severity (FAST staging). In contrast, the corresponding measures for the CDT and MMSE were uniformly poor except for that of NPV, and indicate that the MMSE and CDT are not valid for early detection of AD/DRD in primary care. These findings are noteworthy given that the proportions of the cohort without symptoms

(FAST stage 1) or without objective functional impairment (FAST stages 1–2) were 81% and 91% respectively, and that 43–50% of the impaired patients were asymptomatic.

The absence of symptoms in about half of the impaired patients means that the criterion of subjective decline – required for some definitions of MCI [47–49] – would have excluded 50% of the MCI patients in this primary care sample. In this regard, our findings agree with those of the Iowa EPESE study, a population-based, longitudinal, 10 year cohort of 3,673 persons ≥ 65 years old [50]. They found that 69% of subjects objectively classified as having MCI had no subjective complaints. These data support the conclusion that subjective cognitive complaints, when present, may help identify MCI, but that they are not necessarily present in MCI patients.

Age, gender and education did not appear to significantly influence the MCIS screening results. In previous analysis of the MCIS, the total variance in normal vs. MCI patients explained by these demographics was 17%. However, once the optimal score of the MCIS

Table 3  
Screening Test Performance Statistics Adjusted for Possible Selection Bias from Non-Diagnosed Patients

| Strategy | Sample<br>N | CDT   |        | MMSE   |        | MCIS   |         |
|----------|-------------|---|--------|--------|--------|--------|---------|
|          |             | Sn  | Sp     | Sn     | Sp     | Sn     | Sp      |
| CDx      | 121         | 59%   | 39%    | 71%    | 36%    | 94%    | 97%     |
| 1        | 195         | 49%   | 68%    | 59%    | 66%    | 79%    | 99%     |
| 2        | 195         | 59%   | 70%    | 71%    | 69%    | 94%    | 99%     |
| All      | –           | 49–59%  | 39–70% | 59–71% | 36–69% | 79–94% | 97–99%  |
| Strategy | N           | PPV   | NPV    | PPV    | NPV    | PPV    | NPV     |
| CDx      | 121         | 16%   | 83%    | 17%    | 87%    | 86%    | 99%     |
| 1        | 195         | 23%   | 88%    | 25%    | 89%    | 94%    | 96%     |
| 2        | 195         | 27%   | 90%    | 30%    | 93%    | 95%    | 99%     |
| All      | –           | 16–27%  | 83–90% | 17–30% | 87–93% | 86–95% | 96–99%  |
| Strategy | Sample<br>N | 95% Confidence Interval of Overall ROC Curve Accuracy |        |        |        |        |         |
|          |             | CDT   |        | MMSE   |        | MCIS   |         |
| CDx      | 121         | 43%   | 64%    | 51%    | 72%    | 91%    | 100%    |
| 1        | 195         | 56%   | 73%    | 61%    | 77%    | 92%    | 99%     |
| 2        | 195         | 61%   | 78%    | 68%    | 83%    | 92%    | 100%    |
| All      | –           | 43–61%  | 64–78% | 51–68% | 72–83% | 91–92% | 99–100% |

CDT = Clock Drawing Test; MMSE = Mini-Mental Status Exam; MCIS = MCI Screen; Sn = sensitivity; Sp = specificity; PPV = positive predictive value; NPV = negative predictive value. CDx = Confirmed diagnosis sample (N = 121: 51 confirmed impaired + 70 confirmed normal). Strategy 1: bias adjustment used a maximum published MCI prevalence of 0.20 to estimate maximum false negatives among FAST stage 1 patients. This adjustment assigned 10 unconfirmed normals a diagnosis of impaired.

Strategy 2: bias adjustment assuming no false negatives among the unconfirmed normals.

(derived from correspondence analysis of each patient's item response pattern on the CERAD Wordlist) was added to the logistic regression equation, gender and education no longer accounted for any of the variance, and age only accounted for 1% [61]. The optimal score of the MCIS therefore accounts for almost all of the variance explained by age, gender and education.

In this primary care cohort, and in our previous study [61], the MCIS detected a variety of disorders causing MCI, including AD, cerebrovascular disease, alcohol abuse, traumatic brain injury, metabolic disorders, depressive pseudo-dementia and Parkinson's disease (see Methods for proportional breakdown). The high PPV and NPV of the MCIS therefore allow it to be used to screen a broad spectrum of disorders that can cause MCI or dementia.

The mean duration of FAST stages 2 and 3 for AD are 15 and 7 years respectively [52,58], which implies that the MCI patients in our study were detected anywhere from 1 to 22 or more years before their dementia will develop. These large differences in time to dementia for MCI patients in different FAST stages may explain the wide range of MCI prevalence values reported, as well as explain why "time to dementia" may be an unreliable measure for detecting differences between placebo and treated groups in clinical trials.

The relatively high prevalence of MCI found here reflects: a) the use of an unselected primary care co-

hort; and b) detection of very early stages of MCI (i.e., FAST stage 1) [61]. Even our relatively high prevalence estimates may be low if recent data from the Religious Orders autopsy study approximate AD prevalence among the asymptomatic population (i.e. approximately 50%) [7].

Given the use of routine screening in primary care for disorders such as breast and colorectal cancer, which have lower prevalence rates than those for MCI, we think that screening high risk groups to detect AD early will be more commonly done as better therapies become available.

Historically, it has been difficult to detect all forms of MCI with a single test [46]. The present study, however, shows that almost all MCI cases were detected with the MCIS. Furthermore, an independent study of the MCIS in a sample of 214 subjects found it to be 97% accurate in discriminating normal aging from both amnesic and mixed cognitive domain types of MCI [61]. These MCIS results show that a single test, appropriately designed and scored, can detect most patients with MCI.

#### Study strengths

Current findings are strengthened by the manner in which the data were obtained – namely, in an unse-

Table 4  
Summary of the MCI Screen Component Tasks

| Task Order | Function To Be Tested | Tasks  | Instructions on the Screen  | Comments   |
|------------|-----------------------|--|---|--|
| 1          | Iconic memory         | The 10-word list – Registration*                           | “Repeat each word so I know you heard it correctly.”  | The examiner reads each word to the subject until it is correctly repeated in order to eliminate recall errors due to comprehension or hearing problems.   |
| 2          | Working memory        | The 10-word list – Immediate Free Recall*                  | “Recall as many of the 10 words as you can.”  | After each of the 10 words are correctly repeated, the subject recalls as many of the 10 words as they can, and the words are recorded in the order in which they are recalled.  |
| 3          | Meta-memory           | Judgment about one’s ability to perform a free recall task | “I would like you to guess or estimate how many of the 10 words you will recall on your own in about five minutes.”                                   | The subject’s estimate of their own episodic memory ability is recorded and compared to their actual performance.  |
| 4          | Judgment              | Judged Triadic Comparison of Animals                       | “Which of these three animals is most different from the other two?”  | The examiner reads three animals to the subject who chooses the animal they consider to be most different from the other two for whatever reason. The subject’s response is recorded for 12 different sets of three animals, which are arranged in a balanced lambda-1 design. These judged comparisons are largely independent of culture and language.       |
| 5          | Episodic Memory       | Rehearsed Free Recall                                      | Delayed<br>“Do you remember that earlier, I asked you to remember a list of 10 words? Recall as many of them as you can.”                             | If the subject remembers the task, they then recall as many of the 10 words as they can, and the words are recorded in the order in which they are recalled.   |
| 6          | Episodic Memory       | Delayed Cued Recall  | “I will read you a list of words. After each word, tell me if it was one of the words from the list that you were asked to remember.”                 | 20 words are presented including those from the 10-word list randomly interspersed with 10 distractor words. The examiner records the subject’s response to each word.   |
| 7          | Episodic Memory       | Unrehearsed Free Recall                                    | Delayed<br>“Do you remember the task where I asked you to compare three animals? Now I would like you to recall as many of those animals as you can.” | The examiner records the subject’s responses in the order in which they are recalled. Since the subject was not asked to learn the animals for recall later, test anxiety due to intentional learning does not occur. Test anxiety may have impaired performance when the delayed free recall score for animals is much higher than that for the 10-word list. |

\*The registration and immediate free recall tasks are consecutively repeated three times to learn the 10-word list.

lected primary care sample of all previously undetected patients over 65 years old. The resulting screening test performance statistics are therefore more likely to resemble those found in primary care settings than would have been obtained otherwise. An additional strength of this cohort is minimal study dropouts due to low migration rates for Hancock County, Maine.

#### Study limitations

Excluding the 20 *unconfirmed impaired* patients could have biased results. However, the data from Table 1 show that the *unconfirmed impaired* and *confirmed impaired* patients were quantitatively similar in all respects. Another potential bias arises from lack of diagnostic assessment for the 74 *unconfirmed normal* (FAST stage 1). Bias adjustment strategy 1 provided worst case lower bounds for all measures report-

ed by applying the highest published MCI prevalence to the least impaired group (Table 1 shows that they were quantitatively more normal than the *confirmed normal* patients in terms of younger age, higher education, higher CDT, MMSE, and MCIS scores, and lower FAST stage). These potential biases are therefore unlikely to have altered the screening performance results. Further validation will come from ongoing longitudinal follow-up of this cohort.

## CONCLUSIONS

The present study’s findings independently replicate the previously reported screening performance of the MCIS, and support its validity and accuracy in a primary care setting for early ADRD detection due to a variety of disorders among patients over 65 years old. To

Table 5  
The MCI Screen Sensitivity and Specificity for 471 community dwelling patients (adopted from Shankle et al., PNAS, 2005 [39])

|                   | n   | ROC overall accuracy | 95% CI    | Sensitivity | Specificity |
|-------------------|-----|----------------------|-----------|-------------|-------------|
| MCI vs. normal    | 214 | 97                   | (94, 99)  | 95          | 88          |
| MCI/MD vs. normal | 471 | 98                   | (97, 99)  | 97          | 88          |
| MD vs. normal     | 376 | 99                   | (98, 100) | 96          | 99          |

MD: mild dementia; CI: confidence interval.

date, the MCIS has the highest, published overall ROC accuracy for discriminating normal aging from MCI. In contrast, the most commonly used tests in clinical practice, the MMSE and CDT, were not valid for early ADRD detection. Provided that patients are adequately informed of the risks and benefits of early ADRD detection, the MCIS permits accurate and practical screening of at-risk patients.

## ACKNOWLEDGMENTS

We wish to thank Drs. Barry Reisberg, Susan Weller and Junko Hara for their valuable critique and feedback, Erin Hoschouer for technical assistance, Jane Harris for medical library research assistance, Cynthia Wixon for cognitive testing and data entry.

## APPENDIX A

The MCIS is an online, examiner-administered test of iconic, working and episodic memory, metamemory (judgment about one's own memory ability), judged comparisons, comprehension and speech. It takes about 10 minutes. For each task, the examiner reads the instructions shown on the computer monitor to the subject, and then enters the subject's response into the computer. The instructions for each task are given in quotes, along with each task's description (Table 4).

At the end of the test, the MCIS's scoring algorithm automatically analyzes the subject's item responses and generates an immediate result. The primary result is "normal or impaired". The scores for each task are provided in a table along with age, gender, race and education adjusted z-scores. The test results are summarized in terms of cognitive localization, differential diagnostic considerations, and suggestions for further evaluation when appropriate.

The algorithm used in the MCIS was developed from prior analysis of an independent sample of 471 well-characterized normal aging ( $N = 119$ ) vs. MCI ( $N = 95$ ) (FAST 3) and mild dementia ( $N = 257$ ) patients

who were given the National Institute of Aging ADRD CERAD neuropsychological test battery [72,73]. The wordlist portion of the CERAD battery, when combined with correspondence analysis, logistic regression and receiver operating characteristics (ROC) curve methods, gave overall classification accuracy of 97.3 % and 99 % in differentiating normal from MCI and from mild dementia, respectively (Table 5) [61]. To classify a patient's MCIS results, the algorithm derived from this analysis uses a point on the ROC curve corresponding to 96% sensitivity for MCI-mild dementia and 91% specificity for normal aging.

## APPENDIX B

Summary of all published studies reporting the unbiased measure of test performance, overall accuracy, for discriminating normal aging from MCI using ROC methods.

| Test   | AUC (NL vs. MCI) | NL  | MCI | Total Cases |
|--|------------------|-----|-----|-------------|
| MCI Screen (called "CWL with CA" in article) [61]  | 97.3%            | 119 | 95  | 214         |
| Alzheimer's disease Cooperative Study scale for ADL in MCI [44]  | 97.0%            | 30  | 45  | 75          |
| MCI Screen (MCIS) [Present Study]  | 95.6%            | 130 | 53  | 183         |
| Alzheimer's disease Assessment scale, cognitive subscale (ADAS-Cog) [44]                                 | 93.0%            | 30  | 45  | 75          |
| Wechsler Logical Memory Delayed Recall (LM-II) [35]  | 91.7%            | 88  | 37  | 125         |
| The Memory Screen [60]   | 89.0%            | 198 | 244 | 442         |
| 4-item Instrumental Activities of Daily Living Scale+Single-item informant report of memory problem [35] | 87.2%            | 88  | 37  | 125         |
| 60 second verbal fluency task [14]   | 87.0%            | 46  | 45  | 91          |
| MRI: Diffusion Tensor Left Hippocampal High Mean Diffusivity [41]  | 87.0%            | 18  | 18  | 36          |
| Wechsler Logical Memory Delayed Recall:LM-II [35]  | 86.1%            | 88  | 37  | 125         |
| AD8 Questionnaire [22]   | 84.7%            | 28  | 112 | 140         |
| NeuroTrax Mindstreams [15]   | 82.3%            | 39  | 30  | 69          |
| Wechsler Logical Memory Immed. Recall (LM-I) [35]  | 81.2%            | 88  | 37  | 125         |

| Test  | AUC (NL vs. MCI) | NL  | MCI | Total Cases |
|---|------------------|-----|-----|-------------|
| Single-item informant report of memory problem [35] | 79.5%            | 88  | 37  | 125         |
| Mini-Mental Status Exam (MMSE) [8]                  | 78.5%            | 25  | 26  | 51          |
| NeuroTrax Mindstreams (Best Reported Result) [16]   | 78.3%            | 39  | 30  | 69          |
| 18FDG PET Hippocampal Glucose Metabolism [40]       | 78.0%            | 11  | 13  | 24          |
| 4-item Instrumental ADL Scale [35]                  | 76.9%            | 88  | 37  | 125         |
| Wechsler Logical Memory Immed. Recall (LM-I) [14]   | 76.0%            | 46  | 45  | 91          |
| CNS Vital Signs (Best Result) [26]                  | 74.1%            | 89  | 36  | 125         |
| MRI: Left Hippocampal Low Normalized Volume [41]    | 72.0%            | 18  | 18  | 36          |
| Mini-Mental Status Exam (MMSE) [35]                 | 67.6%            | 88  | 37  | 125         |
| MRI: Hippocampal Atrophy [43]                       | 65.9%            | 59  | 65  | 124         |
| AB Cognitive Screen (Verbal Fluency) [64]           | 64.6%            | 174 | 166 | 340         |
| Enhanced Delayed Cued Recall [54]                   | 62.5%            | 33  | 40  | 73          |
| AB Cognitive Screen (Standardized MMSE) [64]        | 57.2%            | 174 | 166 | 340         |
| Behavioral Dyscontrol Scale [6]                     | 55.0%            | 40  | 40  | 80          |
| California Verbal Learning Test [25]                | NR               | 65  | 65  | 130         |
| Hopkins Verbal Learning Test [57]                   | NR               | 54  | 19  | 73          |
| Buschke Selective Reminding Test [45]               | NR               | 76  | 234 | 310         |
| Montreal Cognitive Assessment (MoCA) [42]           | NR               | 39  | 49  | 88          |
| Hopkins Verbal Learning Test+ Placing Test [33]     | NR               | 51  | 29  | 80          |

NL = Normal; AUC = Area Under the Curve; NR = Not Reported.

## References

- American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, Washington, DC, 2000.
- J.W. Ashford, S. Borson, R. O'Hara, P. Dash, L. Frank, P. Robert, W.R. Shankle, M.C. Tierney, H. Brodaty, F.A. Schmitt, H.C. Kraemer and H. Buschke, Should older adults be screened for dementia? *Alzheimer's & Dementia* **2** (2006), 76–85.
- S.R. Auer, S.G. Sclan, R.A. Yaffee and B. Reisberg, The neglected half of Alzheimer disease: cognitive and functional concomitants of severe dementia, *J Am Geriatr Soc* **42** (1994), 1266–1272.
- L. Backman, B.J. Small and L. Fratiglioni, Stability of the preclinical episodic memory deficit in Alzheimer's disease, *Brain* **124** (2001), 96–102.
- U. Beinhoff, V. Hilbert, D. Bittner, G. Gron and M.W. Riepe, Screening for cognitive impairment: a triage for outpatient care, *Dement Geriatr Cogn Disord* **20** (2005), 278–285.
- H.G. Belanger, K. Wilder-Willis, P. Malloy, S. Salloway, R.F. Hamman and J. Grigsby, Assessing motor and cognitive regulation in AD, MCI, and controls using the Behavioral Dyscontrol Scale, *Arch Clin Neuropsychol* **20** (2005), 183–189.
- D.A. Bennett, J.A. Schneider, Z. Arvanitakis, J.F. Kelly, N.T. Aggarwal, R.C. Shah and R.S. Wilson, Neuropathology of older persons without cognitive impairment from two community-based studies, *Neurology* **66** (2006), 1837–1844.
- A.D. Benson, M.J. Slavin, T.T. Tran, J.R. Petrella and P.M. Doraiswamy, Screening for Early Alzheimer's Disease: Is There Still a Role for the Mini-Mental State Examination? *Prim Care Companion J Clin Psychiatry* **7** (2005), 62–69.
- C.M. Callahan, H.C. Hendrie and W.M. Tierney, Documentation and evaluation of cognitive impairment in elderly primary care patients, *Ann Intern Med* **122** (1995), 422–429.
- P. Chen, G. Ratcliff, S.H. Belle, J.A. Cauley, S.T. DeKosky and M. Ganguli, Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study, *Arch Gen Psychiatry* **58** (2001), 853–858.
- M.M. Cherrier, A.M. Matsumoto, J.K. Amory, S. Asthana, W. Bremner, E.R. Peskind, M.A. Raskind and S. Craft, Testosterone improves spatial memory in men with Alzheimer disease and mild cognitive impairment, *Neurology* **64** (2005), 2063–2068.
- H.C. Chui, J.I. Victoroff, D. Margolin, W. Jagust, R. Shankle and R. Katzman, Criteria for the Diagnosis of Ischemic Vascular Dementia Proposed by the State of California Alzheimer's Disease Diagnostic and Treatment Centers, *Neurology* **42** (1992), 473–480.
- A. Collie and P. Maruff, The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment, *Neurosci Biobehav Rev* **24** (2000), 365–374.
- A. Cunje, D.W. Molloy, T.I. Standish and D.L. Lewis, Alternate forms of logical memory and verbal fluency tasks for repeated testing in early cognitive changes, *Int Psychogeriatr* **19** (2006), 65–75.
- G.M. Doniger, D.M. Zucker, A. Schweiger, T. Dwolatzky, H. Chertkow, H. Crystal and E.S. Simon, Towards practical cognitive assessment for detection of early dementia: a 30-minute computerized battery discriminates as well as longer testing, *Curr Alzheimer Res* **2** (2005), 117–124.
- T. Dwolatzky, V. Whitehead, G.M. Doniger, E.S. Simon, A. Schweiger, D. Jaffe and H. Chertkow, Validity of the Mindstreams computerized cognitive battery for mild cognitive impairment, *J Mol Neurosci* **24** (2004), 33–44.
- M.F. Elias, A. Beiser, P.A. Wolf, R. Au, R.F. White and R.B. D'Agostino, The preclinical phase of Alzheimer disease: A 22-year prospective study of the Framingham Cohort, *Arch Neurol* **57** (2000), 808–813.
- C. Esteban-Santillan, R. Praditsuwan, H. Ueda and D.S. Geldmacher, Clock drawing test in very mild Alzheimer's disease, *J Am Geriatr Soc* **46** (1998), 1266–1269.
- M.R. Farlow and M.L. Lilly, Rivastigmine: an open-label, observational study of safety and effectiveness in treating patients with Alzheimer's disease for up to 5 years, *BMC Geriatr* **5** (2005), 3.
- H. Feldman, A.R. Levy, G.Y. Hsiung, K.R. Peters, A. Donald, S.E. Black, R.W. Bouchard, S.G. Gauthier, D.A. Guzman, D.B. Hogan, A. Kertesz and K. Rockwood, ACCORD Study Group, A Canadian Cohort Study of Cognitive Impairment and Related Dementias (ACCORD): Study Methods and Baseline Results, *Neuroepidemiology* **22** (2003), 265–274.
- M.F. Folstein, S.E. Folstein and P.R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician, *J Psychiatr Res* **12** (1975), 189–198.
- J.E. Galvin, C.M. Roe, C. Xiong and J.C. Morris, Validity and reliability of the AD8 informant interview in dementia, *Neurology* **67** (2006), 1942–1948.
- M. Ganguli, E. Rodriguez, B. Mulsant, S. Richards, R. Pandav, J.V. Bilt, H.H. Dodge, G.P. Stoehr, J. Saxton, R.K. Morycz,

- R.T. Rubin, B. Farkas and S.T. DeKosky, Detection and management of cognitive impairment in primary care: The Steel Valley Seniors Survey, *J Am Geriatr Soc* **52** (2004), 1668–1675.
- [24] R. Goekoop, P. Scheltens, F. Barkhof and S.A. Rombouts, Cholinergic challenge in Alzheimer patients and mild cognitive impairment differentially affects hippocampal activation—a pharmacological fMRI study, *Brain* **129** (2006), 141–157.
- [25] M.C. Greenaway, L.H. Lacritz, D. Binegar, M.F. Weiner, A. Lipton and C. Munro Cullum, Patterns of verbal memory performance in mild cognitive impairment, Alzheimer disease, and normal aging, *Cogn Behav Neurol* **19** (2006), 79–84.
- [26] C.T. Gualtieri and L.G. Johnson, Neurocognitive testing supports a broader concept of mild cognitive impairment, *Am J Alzheimers Dis Other Demen* **20** (2005), 359–366.
- [27] T. Hanninen, M. Hallikainen, S. Tuomainen, M. Vanhanen and H. Soininen, Prevalence of mild cognitive impairment: a population-based study in elderly subjects, *Acta Neurol Scand* **106** (2002), 148–154.
- [28] L.E. Hebert, P.A. Scherr, J.L. Bienias, D.A. Bennett and D.A. Evans, Alzheimer disease in the US population: prevalence estimates using the 2000 census, *Arch Neurol* **60** (2003), 1119–1122.
- [29] C. Holmes, N. Cairns, P. Lantos and A. Mann, Validity of current clinical criteria for Alzheimer's disease, vascular dementia and dementia with Lewy bodies, *Br J Psychiatry* **174** (1999), 45–50.
- [30] G.A. Hopp, R.A. Dixon, M. Grut and L. Backman, Longitudinal and psychometric profiles of two cognitive status tests in very old adults, *J Clin Psychol* **53** (1997), 673–686.
- [31] D.B. Howieson, A. Dame, R. Camicioli, G. Sexton, H. Payami and J.A. Kaye, Cognitive markers preceding Alzheimer's dementia in the healthy oldest old, *J Am Geriatr Soc* **45** (1997), 584–589.
- [32] G.Y. Hsiung, A.D. Sadovnick and H. Feldman, Apolipoprotein E4 genotype as a risk factor for cognitive decline and dementia: data from the Canadian Study of Health and Aging, *CMAJ* **171** (2004), 863–867.
- [33] C.A. De Jager, E. Hogervorst, M. Combrinck and M.M. Budge, Sensitivity and specificity of neuropsychological tests for mild cognitive impairment, vascular cognitive impairment and Alzheimer's disease, *Psychol Med* **33** (2003), 1039–1050.
- [34] J. Koontz and A. Baskys, Effects of galantamine on working memory and global functioning in patients with mild cognitive impairment: a double-blind placebo-controlled study, *Am J Alzheimers Dis Other Demen* **20** (2005), 295–302.
- [35] M. Li, T.P. Ng, E.H. Kua and S.M. Ko, Brief informant screening test for mild cognitive impairment and early Alzheimer's disease, *Dement Geriatr Cogn Disord* **21** (2006), 392–402.
- [36] W.R. Markesbery, F.A. Schmitt, R.J. Kryscio, D.G. Davis, C.D. Smith and D.R. Wekstein, Neuropathologic substrate of mild cognitive impairment, *Arch Neurol* **63** (2006), 38–46.
- [37] A.C. McKee, R. Au, H.J. Cabral, N.W. Kowall, S. Seshadri, C.A. Kubilus, J. Drake and P.A. Wolf, Visual association pathology in preclinical Alzheimer disease, *J Neuropathol Exp Neurol* **65** (2006), 621–630.
- [38] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price and E.M. Stadlan, Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease, *Neurology* **34** (1984), 939–944.
- [39] M.S. Mittelman, S.H. Ferris, E. Shulman, G. Steinberg and B. Levin, A family intervention to delay nursing home placement of patients with Alzheimer disease. A randomized controlled trial, *JAMA* **276** (1996), 1725–1731.
- [40] L. Mosconi, W.H. Tsui, S. De Santi, J. Li, H. Rusinek, A. Convit, Y. Li, M. Boppana and M.J. de Leon, Reduced hippocampal metabolism in MCI and AD: automated FDG-PET image analysis, *Neurology* **64** (2005), 1860–1867.
- [41] M.J. Muller, D. Greverus, C. Weibrich, P.R. Dellani, A. Scheurich, P. Stoeter and A. Fellgiebel, Diagnostic utility of hippocampal size and mean diffusivity in amnesic MCI, *Neurobiol Aging* (2006) [online publication ahead of press].
- [42] Z.S. Nasreddine, N.A. Phillips, V. Bedirian, S. Charbonneau, V. Whitehead, I. Collin, J.L. Cummings and H. Chertkow, The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment, *J Am Geriatr Soc* **53** (2005), 695–699.
- [43] C. Pennanen, M. Kivipelto, S. Tuomainen, P. Hartikainen, T. Hanninen, M.P. Laakso, M. Hallikainen, M. Vanhanen, A. Nissinen, E.L. Helkala, P. Vainio, R. Vanninen, K. Partanen and H. Soininen, Hippocampus and entorhinal cortex in mild cognitive impairment and early AD, *Neurobiol Aging* **25** (2004), 303–310.
- [44] R. Pernecky, C. Pohl, C. Sorg, J. Hartmann, K. Komossa, P. Alexopoulos, S. Wagenpfeil and A. Kurz, Complex activities of daily living in mild cognitive impairment: conceptual and diagnostic issues, *Age Ageing* **35** (2006), 240–245.
- [45] R.C. Petersen, G.E. Smith, S.C. Waring, R.J. Ivnik, E.G. Tangalos and E. Kokmen, Mild cognitive impairment: clinical characterization and outcome, *Arch Neurol* **56** (1999), 303–308.
- [46] R.C. Petersen, R. Doody, A. Kurz, R.C. Mohs, J.C. Morris, P.V. Rabins, K. Ritchie, M. Rossor, L. Thal and B. Winblad, Current concepts in mild cognitive impairment, *Arch Neurol* **58** (2001), 1985–1992.
- [47] R.C. Petersen, J.C. Stevens, M. Ganguli, E.G. Tangalos, J.L. Cummings and S.T. DeKosky, Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review) Report of the Quality Standards Subcommittee of the American Academy of Neurology, *Neurology* **56** (2001), 1133–1142.
- [48] R.C. Petersen and J. O'Brien, Mild cognitive impairment should be considered for DSM-V, *J Geriatr Psychiatry Neurol* **19** (2006), 147–154.
- [49] R.C. Petersen, J.E. Parisi, D.W. Dickson, K.A. Johnson, D.S. Knopman, B.F. Boeve, G.A. Jicha, R.J. Ivnik, G.E. Smith, E.G. Tangalos, H. Braak and E. Kokmen, Neuropathologic features of amnesic mild cognitive impairment, *Arch Neurol* **63** (2006), 665–672.
- [50] J.L. Purser, G.G. Fillenbaum and R.B. Wallace, Memory complaint is not necessary for diagnosis of mild cognitive impairment and does not predict 10-year trajectories of functional disability, word recall, or short portable mental status questionnaire limitations, *J Am Geriatr Soc* **54** (2006), 335–338.
- [51] B. Reisberg, S.H. Ferris, E.H. Franssen, E. Shulman, I. Monteiro, S.G. Sclan, G. Steinberg, A. Kluger, C. Torossian, M.J. de Leon and E. Laska, Mortality and temporal course of probable Alzheimer's disease: a 5-year prospective study, *Int Psychogeriatr* **8** (1996), 291–311.
- [52] B. Reisberg, S. Ferris, M.J. de Leon, C. Torossian, S. Kadiyala and W. Zhu, Subjective Cognitive Impairment: The pre-Mild Cognitive Impairment stage of brain degeneration: Longitudinal outcome after a mean of 7 years follow-up, *Neuropsychopharmacology* **30** (2005), S81.
- [53] E.H. Rubin, M. Storandt, J.P. Miller, D.A. Kinscherf, E.A. Grant, J.C. Morris and L. Berg, A prospective study of cog-

- nitive function and onset of dementia in cognitively healthy elders, *Arch Neurol* **55** (1998), 395–401.
- [54] E. Saka, E. Mihci, M.A. Topcuoglu and S. Balkan, Enhanced cued recall has a high utility as a screening test in the diagnosis of Alzheimer's disease and mild cognitive impairment in Turkish people, *Arch Clin Neuropsychol* **21** (2006), 745–751.
- [55] S. Salloway, S. Ferris, A. Kluger, R. Goldman, T. Griesing, D. Kumar and S. Richardson, Donepezil 401 Study Group, Efficacy of donepezil in mild cognitive impairment: a randomized placebo-controlled trial, *Neurology* **63** (2004), 651–657.
- [56] M. Sano, C. Ernesto, R.G. Thomas, M.R. Klauber, K. Schafer, M. Grundman, P. Woodbury, J. Growdon, C.W. Cotman, E. Pfeiffer, L.S. Schneider and L.J. Thal, A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. The Alzheimer's Disease Cooperative Study, *New Eng J Med* **336** (1997), 1216–1222.
- [57] A.M. Schrijnemaekers, C.A. de Jager, E. Hogervorst and M.M. Budge, Cases with mild cognitive impairment and Alzheimer's disease fail to benefit from repeated exposure to episodic memory tests as compared with controls, *J Clin Exp Neuropsychol* **28** (2006), 438–455.
- [58] S.G. Sclan and B. Reisberg, Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality, *Int Psychogeriatr* **4** (1992), 55–69.
- [59] B. Seltzer, P. Zolnouni, M. Nunez, R. Goldman, D. Kumar, J. Ieni and S. Richardson, Donepezil 402 Study Group, Efficacy of donepezil in early-stage Alzheimer disease: a randomized placebo-controlled trial, *Arch Neurol* **61** (2004), 1852–1856.
- [60] W.R. Shankle, S. Mani, M.J. Pazzani and P. Smyth, Dementia Screening with Machine Learning methods, in: *Intelligent Data Analysis in Medicine and Pharmacology*, E. Keravnou, N. Lavrac and B. Zupan, eds, Kluwer Academic Publishers, 1997, pp. 149–165.
- [61] W.R. Shankle, A.K. Romney, J. Hara, D. Fortier, M.B. Dick, J.M. Chen, T. Chan and X. Sun, Methods to improve the detection of mild cognitive impairment, *Proc Natl Acad Sci USA* **102** (2005), 4919–4924.
- [62] B.J. Small, L. Fratiglioni, M. Viitanen, B. Winblad and L. Backman, The course of cognitive impairment in preclinical Alzheimer disease: three- and 6-year follow-up of a population-based sample, *Arch Neurol* **57** (2000), 839–844.
- [63] M.R. Somerfield, C.S. Weisman, W. Ury, G.A. Chase and M.F. Folstein, Physician practices in the diagnosis of dementing disorders, *J Am Geriatr Soc* **39** (1991), 172–175.
- [64] T.I. Standish, D.W. Molloy, A. Cunje and D.L. Lewis, Do the ABCS 135 short cognitive screen and its subtests discriminate between normal cognition, mild cognitive impairment and dementia? *Int J Geriatr Psychiatry* (2006) [online publication ahead of press].
- [65] J.E. Storey, J.T. Rowland, D. Basic and D.A. Conforti, A comparison of five clock scoring methods using ROC (receiver operating characteristic) curve analysis, *Int J Geriatr Psychiatry* **16** (2001), 394–399.
- [66] J.E. Storey, J.T. Rowland, D. Basic and D.A. Conforti, Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients, *Int Psychogeriatr* **14** (2002), 259–271.
- [67] M.C. Tierney, J.P. Szalai, W.G. Snow, R.H. Fisher, A. Nores, G. Nadon, E. Dunn and P.H. St George-Hyslop, Prediction of probable Alzheimer's disease in memory-impaired patients: A prospective longitudinal study, *Neurology* **46** (1996), 661–665.
- [68] J. Touchon and K. Ritchie, Prodromal cognitive disorder in Alzheimer's disease, *Int J Geriatr Psychiatry* **14** (1999), 556–563.
- [69] V.G. Valcour, K.H. Masaki, J.D. Curb and P.L. Blanchette, The detection of dementia in the primary care setting, *Arch Intern Med* **160** (2000), 2964–2968.
- [70] G.S. Watson, B.A. Cholerton, M.A. Reger, L.D. Baker, S.R. Plymate, S. Asthana, M.A. Fishel, J.J. Kulstad, P.S. Green, D.G. Cook, S.E. Kahn, M.L. Keeling and S. Craft, Preserved cognition in patients with early Alzheimer disease and amnesic mild cognitive impairment during treatment with rosiglitazone: a preliminary study, *Am J Geriatr Psychiatry* **13** (2005), 950–958.
- [71] Y.I. Watson, C.L. Arfken and S.J. Birge, Clock completion: an objective screening test for dementia, *J Am Geriatr Soc* **41** (1993), 1235–1240.
- [72] K. Welsh, N. Butters, J. Hughes, R. Mohs and A. Heyman, Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures, *Arch Neurol* **48** (1991), 278–281.
- [73] K.A. Welsh, N. Butters, R.C. Mohs, D. Beekly, S. Edland, G. Fillenbaum and A. Heyman, The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery, *Neurology* **44** (1994), 609–614.
- [74] K. Yaffe, K. Krueger, S.R. Cummings, T. Blackwell, V.W. Henderson, S. Sarkar, K. Ensrud and D. Grady, Effect of raloxifene on prevention of dementia and cognitive impairment in older women: the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial, *Am J Psychiatry* **162** (2005), 683–690.