Evaluation of Cognitive Impairment in Older Adults

Combining Brief Informant and Performance Measures

James E. Galvin, MD, MPH; Catherine M. Roe, PhD; John C. Morris, MD

Objective: To combine the AD8, a brief informant interview, with performance measures to develop a brief screening tool to improve detection of cognitive impairment and dementia in general practice.

Design: The AD8 was administered to informants. Clinicians conducted independent patient evaluations and administered the Clinical Dementia Rating Scale and a 30-minute neuropsychological battery. Logistic regression was used to determine the best combination of brief tests to correctly classify patients as having no dementia, uncertain dementia, or dementia. The area under the receiver operator characteristic curve (AUC) evaluated the discriminative ability of the combined tests.

Patients/Setting: Patients (n=255) were consecutive referrals to a dementia clinic. Patients had a mean±SD age of 73.3±11.3 years, with 13.7±3.0 (mean±SD) years of education. The sample was 56% women; 77% of patients were white.

Main Outcome Measure: Dementia classification.

Results: A model combining the AD8 interview (odds ratio, 1.91; 95% confidence interval, 1.6-2.3) and the Consortium to Establish a Registry for Alzheimer Disease 10-item Word List Recall (odds ratio, 1.43; 95% confidence interval, 1.2-1.7) predicted dementia with 91.5% correct classification (AUC=0.968; 95% confidence interval, 0.93-0.99). A cutoff of 2 or greater on the AD8 and less than 5 items remembered on the Word List Recall was sensitive (94%) and specific (82%). For cognitive impairments not meeting dementia criteria, combining AD8 (odds ratio, 2.31; 95% confidence interval, 1.3-4.0) and Word List Recall (odds ratio, 1.42; 95% confidence interval, 1.1-1.8) was most predictive (AUC=0.91; 95% confidence interval, 0.8-1.0). Using the same cutoffs as those used for dementia gave the best combination of sensitivity (85%) and specificity (84%).

Conclusion: Combining the AD8 interview with the Word List Recall improves the ability to detect the presence of dementia. The AD8 can be administered to an informant and, when combined with Word List Recall, is a powerful yet brief method of detecting cognitive impairment.

Arch Neurol. 2007;64:718-724
mentia Rating Scale (CDR), may be more sensitive than brief performance measures that rely on interindividual norms to detect cognitive change. We used this premise to develop a brief interview, the AD8, which distinguishes individuals with very mild dementia from those without dementia, regardless of etiology. The AD8, which is based on intradividual decline, has been demonstrated to be a valid and reliable screening tool for dementia. We explored the potential added value of neuropsychological testing combined with the AD8 in developing a brief screening battery for use in general practice to improve clinicians’ ability to detect cognitive disorders at the earliest possible stage.

### METHODS

#### STUDY PARTICIPANTS

Participants were drawn from a consecutive series of referrals to the Memory Diagnostic Center, a dementia specialty practice at Washington University School of Medicine, for evaluation of cognitive, behavioral, and mood disorders. Diagnoses ranged from no dementia through all levels of dementia severity. When calling for an appointment, the patient identified an informant to provide additional information on cognitive and functional change. A total of 235 patient-informant dyads agreed to participate. No patient-informant dyad contributed more than 1 visit to the data set. The Washington University Human Studies Committee approved all procedures.

#### ADMINISTRATION OF THE AD8

The AD8 contains 8 questions (yes or no) that ask the informant to rate change in cognition and function. After informed consent, the informant rated the patient, and the number of yes answers was totaled to obtain the AD8 score. The Memory Diagnostic Center physicians were blinded to the results of AD8 administration. (A copy of the AD8 table with scoring rules may be found at http://alzheimer.wustl.edu/AboutUs/PDFs/AD8form2005.pdf.)

#### CLINICAL ASSESSMENT

The Memory Diagnostic Center physicians conducted independent, semistructured interviews with the patient and a knowledgeable collateral source (usually the spouse or a close family member). Each patient-caregiver dyad was interviewed to generate a diagnosis and CDR score. The diagnostic criteria for AD were consistent with the definitions from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition and of “probable AD” with the National Institute of Neurological and Communication Disorders and Stroke-Alzheimer’s Disease and Related Disorders Association criteria. Published criteria were used for other dementing disorders.

The CDR was used to determine the presence or absence of dementia and to stage its severity. A CDR score of 0 indicates no dementia; CDR 0.5 represents very mild dementia or, in cases where cognitive impairment does not meet dementia criteria, uncertain dementia; and CDR 1, 2, and 3 corresponds to mild, moderate, and severe dementia, respectively. The sum of CDR boxes represents the standard for cognitive impairment in this study.

In many individuals, the CDR 0.5 rating equates with very mild dementia and is the threshold of demented status. A subset of participants with a CDR score of 0.5 had cognitive impairments that did not meet criteria for dementia. We operationalized these individuals as having uncertain dementia comparable with MCI.

#### NEUROPSYCHOLOGICAL EVALUATION

Each patient was administered a 30-minute test battery at the time of his or her office visit. Episodic memory was assessed by the logical memory subtest of the Wechsler Memory Scale, and the CERAD 10-item Word List Recall, immediate and delayed. The Animal Fluency Test assessed semantic memory, and the 15-item Boston Naming Test assessed confrontation naming. Three measures addressed psychomotor, visuospatial, and executive abilities: the digit symbol subtest of the Wechsler Adult Intelligence Scale, the Trail-Making A Test, and the Trail-Making B Test. Brief global measures included the MMSE and the Short Blessed Test (SBT).

#### STATISTICAL ANALYSIS

All analyses were performed using SPSS, version 13.0 (SPSS Inc, Chicago, Ill). Descriptive statistics were used to report the demographic characteristics of the patients and informants, neuropsychological tests, and dementia stages. Group means were compared using analysis of variance; post hoc comparisons were made using Tukey’s Honestly Significant Difference test. Logistic regression models were developed to determine the best combination of brief tests (defined here as taking <7 minutes to complete) to correctly classify patients as having no dementia (CDR score 0=0) or dementia (CDR score 0.5 or greater=1). Continuous test scores were used to initially identify the tests and combinations of tests that significantly predicted dementia. Once the predictive tests were identified, we determined the best cutoff scores for use in clinical practice. Psychometric tests included in the logistic regression analyses were the MMSE, SBT, Word List Recall (immediate and delayed), and the Animal Fluency, Boston Naming, and Trail-Making A tests. The Trail-Making B Test, and the Wechsler Memory Scale logical memory and Wechsler Adult Intelligence Scale digit symbol subtests were excluded from the analyses because of the length of time needed to administer them and the complexity of scoring and interpretation. We used 3 approaches to test the validity of the models. First, all variables (AD8 and psychometrics) were entered simultaneously to determine which variables independently predicted a CDR score greater than 0 when adjusting for scores on the remaining tests. We also used 2 stepwise approaches (forward and backward) with the AD8 forced into the model and the psychometric tests as candidates for stepwise entry. The probability was 0.05 for stepwise entry and 0.10 for removal. Because similar models were elicited using forward and backward stepwise methods, only results from the forward stepwise regressions are reported. The odds ratios (ORs) and confidence intervals (CIs) were reported for each measure; the percentage correctly classified as demented is reported. Receiver operator characteristic curves and the area under the receiver operator characteristic curve (AUC) were generated to graphically and quantitatively reflect the ability of the AD8 and each of the models derived from logistic regression to discriminate between patients without dementia (CDR score 0) and patients with dementia (CDR score ≥0.5). Analyses were repeated to determine discriminative properties of the AD8 and each of the models between patients without dementia (CDR score 0) and patients with uncertain dementia (CDR score ≥0.5).
**RESULTS**

**SAMPLE CHARACTERISTICS**

A total of 255 patient-informant dyads were evaluated between October 1, 2003, and September 30, 2004. Patients’ mean±SD age at the time of assessment was 73.3±11.3 years (range, 40-102 years), with 13.7±3 (mean±SD) years of education (range, 6-20 years). The sample was composed of 56% women; 77.1% of patients were white. Fifty-three percent of collateral sources were spouses, 37% were children, and 10% were other (relatives, friends, and paid caregivers). Twelve percent of the sample was nondemented (CDR score=0). Eleven percent had cognitive impairments that were not sufficient to interfere with everyday function or were potentially reversible (CDR score=0.5), comparable with MCI.15,33 Dementia diagnoses included AD (64%), and vascular (8%), Lewy body (8%), frontotemporal (7%), progressive aphasia (5%), and other (8%) dementias. The mean±SD MMSE score for the sample was 19.3±7.8, and the mean±SD SBT score was 12.8±8.1.

Demographic characteristics, dementia staging, and performance on neuropsychological tests for the no dementia, uncertain dementia, and dementia groups are provided (Table 1). The dementia group was older than the no dementia (P=.002) and uncertain dementia (P=.04) groups. The dementia group was also less educated than the no dementia (P<.001) and uncertain dementia (P=.02) groups. The dementia group performed worse than the other groups on all of the psychometric tests. The no dementia group differed from the uncertain dementia group in AD8 (P=.02), digit symbol subtest (P=.04), and Word List Recall (immediate [P=.04] and delayed [P<.001]) scores.

**MODELING INFORMANT AND PSYCHOMETRIC PERFORMANCE MEASURES TO DETECT DEMENTIA**

Logistic regression combining the AD8 and brief psychometric tests was performed to determine which combination of tests best discriminated individuals with dementia from those without dementia (Table 2). The AD8 was a significant predictor of group membership (step 1; Wald χ²=44.3; OR, 1.91; 95% CI, 1.6-2.3; P<.001). With each 1-point increase in AD8 score, patients were 2 times more likely to be demented; 87.9% of patients were correctly classified as having dementia using AD8 scores (dementia prevalence, 77%). The addition of the Word List Recall (step 2; Wald χ²=14.7; P<.001) improved classification (91.5% correct classification; OR, 1.43; 95% CI, 1.2-1.7) (Table 2). As the number of words recalled decreased, the more likely it became that the patient was demented. The addition of the Boston Naming Test (step 3) further increased correct classification to 95.2%. The MMSE, SBT, and the Animal Fluency, Trail-Making A, and Word List Recall (immediate) tests did not enter the stepwise models.

**DISCRIMINATIVE ABILITY OF AD8 AND BRIEF PSYCHOMETRIC PERFORMANCE MEASURES IN DETECTING DEMENTIA**

Receiver operator characteristic curves were generated to determine the ability of the AD8, alone and combined with the psychometric tests from the stepwise regression, to best discriminate between individuals with and individuals without dementia (Figure 1). The AUC for the AD8 alone was 92.6% (95% CI, 0.88-0.96). The addition of the Word List Recall increased discriminative...
tion (AUC=0.968; 95% CI, 0.93-0.99). The addition of the Boston Naming Test (step 3; AUC=0.927; 95% CI, 0.85-1.0) did not improve discrimination. To make combined use of AD8 and Word List Recall scores practical in a clinical setting, we determined the sensitivity and specificity values yielded using different cutoff scores for each of the tests. Combining a cutoff of 2 or greater on the AD8 and less than 5 items remembered on the Word List Recall gave the best combination of sensitivity (94.1%) and specificity (81.8%).

**ALTERNATIVE MODELS USING TRADITIONAL BRIEF OFFICE MEASURES**

We determined whether the inclusion of 2 commonly used brief performance measures of cognition (the MMSE and SBT) increased dementia detection when combined with the AD8 (Table 3). The AD8 alone (89.7% of patients correctly classified) performed equally as well as the MMSE alone (89.3% of patients correctly classified) and the SBT alone (89.5% of patients correctly classified). The addition of the MMSE and/or the SBT to the AD8 did not increase dementia detection.

**COMBINING BRIEF INFORMANT AND PERFORMANCE MEASURES TO DETECT MILD IMPAIRMENTS OF COGNITION**

We examined the ability of combined brief measures to detect uncertain dementia (n=26) using forward stepwise logistic regression. The AD8 (Wald $\chi^2=9.20$; OR, 2.31; 95% CI, 1.3-4.0; $P=.002$) and Word List Recall (Wald $\chi^2=8.98$; OR, 1.42; 95% CI, 1.1-1.8; $P=.003$) were significant predictors of group membership (76% correct classification). Unlike the models generated for predicting dementia, the Boston Naming Test did not enter the stepwise models. Receiver operator characteristic curves were constructed for the AD8 alone and the combined battery of the AD8 and Word List Recall (Figure 2). The AUC for the AD8 was 0.77 (95% CI, 0.6-0.9; $P=.004$) and was 0.91 for the combined battery (95% CI, 0.8-1.0; $P<.001$). Combining a cutoff of 2 or greater on the AD8 and less than 5 words on the Word List Recall gave the best combination of sensitivity (85.0%) and specificity (84.2%).

Similar to dementia diagnoses, the MMSE did not provide useful information concerning the uncertain dementia group. The MMSE by itself was not a significant predictor of cognitive impairment (OR, 0.99; 95% CI, 0.9-1.0; $P=.87$) and was not as effective at discriminating individuals without dementia from those with the mildest forms of cognitive impairment (AUC=0.71; 95% CI, 0.5-0.8). The model combining the MMSE with the AD8 did not perform better (AUC=0.78; 95% CI, 0.6-0.9) than the AD8 alone. Combining a cutoff score of 2 or greater on the AD8 and less than 28 on the MMSE gave the best combination of sensitivity (73.9%) and specificity (69.6%) for the combined tests, but was less sensitive and specific than combining the AD8 with Word List Recall.

**COMMENT**

We previously demonstrated that the AD8 is a brief informant interview that is able to discriminate cognitively normal older adults, regardless of age, sex, race, or education, from those with even the mildest stages of cognitive impairment. Here, we find that combining the AD8 with brief psychometric tests improves prediction of the presence of dementia. The addition of the 10-item Word List Recall improved discrimination to 97%
between adults without dementia and those with dementia. Although the addition of a test of confrontational naming (Boston Naming) slightly increased the detection of dementia, we believe that the increased time needed to administer and score this additional test offsets the slight improvement in detection. Instead, we suggest that when there is limited time available for clinical assessment, the AD8 combined with Word List Recall results in a brief battery that optimizes the detection of dementia. Using cutoff scores of 2 or greater on the AD8 and 5 or fewer words on the Word List Recall resulted in excellent sensitivity (94%) and specificity (82%). Common brief screening instruments, such as the MMSE and SBT, did not improve dementia detection.

We also examined whether the same brief battery would be effective in detecting cognitive changes not meeting dementia criteria. The combination of the AD8 and the Word List Recall detected more than 90% of these individuals using the same cutoffs as for the dementia groups with very good sensitivity (85%) and specificity (84%). The addition of the Boston Naming Test did not contribute to discrimination, a finding similar to other reports noting that naming impairments may not characterize persons with MCI, a group operationally identical to the uncertain dementia group described here. The MMSE did not contribute to detection or discrimination but performed worse than the AD8 alone, with unacceptable low sensitivity (74%) and specificity (70%).

The neuropsychological profile of AD and amnestic MCI is well documented; impairments in episodic memory predominate in conjunction with deficits in associative learning and semantic ability. Individuals without dementia followed up longitudinally who later develop AD have poorer initial performance on a number of tasks, including the Rey Auditory Verbal Learning Task (a delayed recall of word lists), and the Animal Fluency and Wechsler Adult Intelligence Scale information tests. The logical memory subtest of the Wechsler Memory Scale may be the single most useful test in detecting episodic memory impairment; however, the task is lengthy and requires specialized training to administer and interpret.

In developing this project, we focused on brief tests that required little specialized training or equipment, which could be carried out in most settings. In a recent study, investigators combined a brief test of episodic memory (the John Brown, 42 Market Street, Chicago subtest from the Short Blessed Test) with the 1-minute Animal Fluency Test, which discriminated individuals with dementia from those without dementia with similar degrees of sensitivity and specificity as the MMSE. Brief batteries, such as the one proposed here—the AD8 and Word List Recall (with or without confrontational nam-
A number of brief screening measures, such as the MMSE\(^1^0\) and SBT,\(^2^2\) are already available, but these performance-based measures may not be able to detect or quantify change from previous levels of function, particularly in very high-functioning individuals or those with poorer long-term abilities. Furthermore, many cognitive tests are culturally insensitive and may underestimate the abilities of African American individuals and other minority groups.\(^4^6\) There is also little available data about how these brief measures perform in non-AD dementia. Clock drawing\(^7^7\) is also commonly used; however, the clock lacks the sensitivity to detect MCI or mild dementia, regardless of the scoring method used.\(^8^8\) We have demonstrated that the AD8 combined with Word List Recall reliably detects all forms of cognitive impairment. Although a small proportion of individuals without dementia may screen for dementia using our brief screening battery, further evaluation should exclude these individuals.

There are limits to this study. The sample is drawn from patients referred to an academic specialty clinic and may not be representative of the general population. However, other than educational attainment, the demographic attributes of the sample are similar to US census reports for the St Louis metropolitan area. The mixture of patients in this sample had a diversity of sex and race; included multiple medical comorbidities; had a combination of cognitive, behavioral, and affective disorders; and had collateral sources that varied in terms of relationship and exposure to the patients. In this setting, the AD8 is a brief screening tool that reliably discriminates healthy older adults from those with MCI or very mild dementia. The AD8 can be administered to an informant, and, when combined with Word List Recall, is a powerful yet brief method for detecting cognitive impairment. This brief battery consisting of an informant interview and a performance measure may be used as a screening tool in community settings, primary care practices, or as part of epidemiological studies to detect dementia in older adults.

Accepted for Publication: December 5, 2006.

Correspondence: James E. Galvin, MD, MPH, Alzheimer Disease Research Center, Washington University School of Medicine, 4488 Forest Park Ave, Suite 130, St Louis, MO 63108 (galvinj@neuro.wusl.edu).

Author Contributions: Study concept and design: Galvin and Morris. Acquisition of data: Galvin and Roe. Analysis and interpretation of data: Galvin and Roe. Drafting of the manuscript: Galvin. Critical revision of the manuscript for important intellectual content: Galvin, Roe, and Morris. Statistical analysis: Galvin and Roe. Obtained funding: Galvin and Morris. Administrative, technical, and material support: Morris. Study supervision: Galvin. All 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.

Financial Disclosure: None reported.

Funding/Support: This study was supported by grants from the Longer Life Foundation; the National Institute on Aging (K08 AG20764, P01 AG03991, and P50 AG05681); the American Federation for Aging Research; and the Alan A. and Edith L. Wolff Charitable Trust (St Louis, Mo). Dr Galvin is a recipient of the Paul Beeson Physician Faculty–Scholar in Aging Research Award.

Acknowledgment: We thank the physicians, nurse clinicians, and staff at the Memory Diagnostic Center for assistance in completing this study.

REFERENCES

16. Román GC. Vascular dementia may be the most common form of dementia in the elderly. J Neurol Sci. 2002;203:7-10.