James E. Galvin, M.D.

"The Key Features Inventory: Early Diagnosis of Dementia to Improve Quality of Life"

Abstract

Dementia is a significant public health problem that is under-recognized and under-diagnosed. Reliance on objective performance in office-based cognitive tests such as the Mini-Mental Status Exam is hampered by insensitivity to the early stages of dementia and contains biases based on gender, race and educational attainment. Informant-based assessments on the other hand, are unbiased and more sensitive to early cognitive change but are time-consuming.

To address the absence of brief but sensitive measures of cognitive change, we undertook a pilot study of individuals enrolled in a longitudinal study of memory and aging to identify clinical variables that appeared to reliably differentiate cognitively normal individuals from those with very mild and mild dementia.

Following funding we developed a 55-item battery of informant queries regarding an individual's cognitive status was derived from a semistructured interview and a consensus panel of dementia experts. The battery was evaluated with informants for 189 consecutive participants of a longitudinal study of memory and aging and compared with an independently-obtained Clinical Dementia Rating (CDR) rating for the participant. Multiple regression and ROC curves assessed subsets of the items to discriminate between CDR 0 (no dementia) and CDR 0.5 (very mild dementia).

The final version (AD8) querying memory, orientation, judgment, and function was administered to an additional sample of 112 CDR 0 and 68 CDR 0.5 participants. Using a cut-off of 2 items endorsed, the area under the curve was 0.834, suggesting good to excellent discrimination; sensitivity = 74% and specificity = 86% (prevalence of 0.38 for very mild dementia). Inclusion of 56 additional individuals with mild to severe dementia (increasing dementia prevalence to 0.53) increased sensitivity to 85%.

Lay Summary (1 – 2 paragraphs) Short summary of research questions, methods, results and conclusions in language that can be understood by lay readers.

The AD8 was developed to serve as brief, informant-based assessment to increase the ability of the practicing clinician to detect dementia. Since dementia work-ups are time-consuming and insensitive, it is often difficult for health care providers to pick-up cognitive change in older adults and diagnose dementia. We used an expert panel to develop questions that tapped into memory, orientation, activities of daily living, judgment and problem-solving abilities. These questions were then prospectively testing in a sample of older adults participating in a longitudinal research project.

The AD8 was very sensitive in discriminating individuals with normal cognition from those with even the mildest of cognitive decline. It was very brief (less than 2 minutes) and was not influenced by education, gender, race or relationship of the informant. We have also completed a study looking at the validity and reliability of the AD8 in a clinic sample that will better represent the community practice of medicine, and are testing the ability of the AD8 to serve as a self-rating tool.

Introduction/ Brief Literature Review

Alzheimer disease (AD) and other dementias are both underrecognized and underdiagnosed in the community. This may be due, in part, to the lack of brief measures that can discriminate normal aging from very mild dementia. Tests such as the Mini-mental state exam (MMSE) have a ceiling effect that makes them insensitive to early signs of dementia, especially in highly educated individuals, and they may not be culturally sensitive. Other scales such as the Cognitive Abilities Screening Instrument (CASI) and the General Practitioner

Assessment of Cognition (GPCOG) have less cultural bias but require extensive training to administer and generally are too lengthy for use in general practice. Brief measures such as the Short Blessed Test (SBT) and the Memory Impairment Screen are heavily weighted towards memory deficits and may not detect non-amnestic dementias. The Clock-Drawing task or cube-copying are similarly limited to a single cognitive domain and may not be useful in detecting mild cases of dementia.

Published criteria for AD diagnosis such as those developed by the Work Group of the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) require standard assessment of patient. Comparison of individual performance on cognitive test measures to normative values, however, may not detect declines that occur in very mild dementia, particularly in high-functioning individuals. Further, brief objective testing may falsely identify as demented people with life-long poor cognitive functioning. Informant-based assessments, on the other hand, may reveal early cognitive change because of a longitudinal perspective, have face validity, and are established in studies characterizing AD and in AD clinical trials. Global rating systems such as the Clinical Dementia Rating (CDR) and the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) incorporate information from a collateral source to assess change in the patient's cognitive ability to conduct their accustomed activities. These systems do not require a baseline assessment for comparison and minimize issues such as practice effects and educational and sociocultural influences that confound interpretation of performance-based assessments. These informant-based assessments, however, are lengthy and require interpretation by an experienced clinician; therefore, they are difficult to use in community practice. Given the brief period available to primary care physicians in a standard office visit, there will likely be some acceptable trade-off to the clinician, sacrificing specificity and sensitivity to detect dementia to keep the clinical tool brief. Given that requirement, our goal was to develop a dementia screening instrument that could be completed in a just in a few minutes with maximally balanced sensitivity and specificity.

We developed a brief dementia assessment that includes both performance-based assessment of the patient and a brief informant interview. We identified clinical variables that we believed would distinguish individuals with very mild dementia from those without dementia based on a review of the literature and our experience with the semi-structured interview of the informant in the research protocol used to derive the CDR (we also use information from the participant to derive the CDR). We identified 55 questions asked of the informant about the cognitive function of the participant that were thought to assist in dementia detection, and when combined with a brief objective assessment of the participant, might serve as a diagnostic screening tool. The results were compared with an independently derived CDR rating of the participant. Data from these participants were then used to develop an eight-item informant interview (the AD8) to discriminate between CDR 0 (nondemented) and CDR 0.5 (very mildly impaired conditions). The psychometric and discriminative properties of the AD8 were then tested prospectively in a second sample. Evaluation of the AD8 in combination with the brief participant assessment is ongoing; here we describe the development and validation of the AD8 and its correlation with other dementia screening measures.

Methods

All participants were volunteers who enrolled in a longitudinal study (initiated in 1979 and studying to date 2049 individuals) of healthy aging and dementia. The Washington University Human Studies Committee approved all procedures. Experienced clinicians (neurologists, psychiatrists, geriatricians, and masters-prepared geriatric nurse specialists) conducted independent semistructured interviews with the participant and a knowledgeable collateral source (usually the spouse or close family member) to capture features suggestive of a dementing disorder.

The CDR was used as the "Gold Standard" to determine the presence or absence of dementia and to stage its severity. Using all information from the clinical assessment protocol but without reference to the individual's psychometric performance, the CDR rates cognitive function in each of six categories (memory, orientation, judgment and problem solving, and performance

in community affairs, home and hobbies and personal care). The global CDR is derived from the individual ratings in each of the six categories where CDR 0 indicates no dementia. CDR 0.5 may represent very mild dementia or in some cases with minimal impairment, uncertain or questionable dementia. CDR 1, 2, or 3 corresponds to mild, moderate, or severe dementia. The CDR is useful to detect the change in cognitive abilities from a prior level of function and also to assess interference with accustomed activities.

The AD8 was developed using a combination of expert clinical judgment and statistical modeling. A list of 55 questionnaire items was developed based on an extensive review of the literature, our experience with semistructured informant interviews, the consensus opinion of 4 dementia experts, and the results of a previously collected telephone survey of community-dwelling older adults (unpublished data). Three response options were offered with instructions: (1) Yes, a change; (2) No, no change; or (3) N/A, don't know. Between June 3, 2002, and February 21, 2003, the 55-item questionnaire was administered to 290 consecutive collateral sources (CS) of study participants at their annual assessment. None of the 55 items were contained within the structured part of the interview. An open-ended part of the interview allows clinicians to ask non-standard questions that probe changes in memory, orientation, and other cognitive domains specific to that participant. To maximize its ability to discriminate between normal cognitive aging and very mild dementia, the AD8 was developed using data from the CS responses of participants who were aged 55 years or above, and who received a CDR score of 0 (N=86) or 0.5 (N=103) at their assessment. There were no significant differences between the CDR groups with regards to age, gender, race, education, or the relationship between the CS and the participant.

Because the AD8 was conceptualized as a brief instrument, the RSQUARE selection method (SAS version 8 for Sun OS, Cary, NC) was used to reduce the number of items. This method finds subsets of independent variables that best predict a dependent variable using multiple linear regression. After specifying the subset size, the procedure computes all possible regressions and their R^2 value so the models' predictive ability can be compared. The procedure was carried out within the CDR 0.5 sample (N=103) using the CDR-SB as the dependent variable. All 55 items were used as predictor variables, and subset size was set to 10. Thus, regressions were calculated for all possible 10-item combinations of the 55 predictor variables. The model with the highest correlation coefficient was chosen ($R^2 = 0.744$) and revised by the research team to improve face validity. Ultimately an 8-item scale was developed. The predictive ability of the 8-item scale (area under curve [AUC] = 0.870) was superior to the 10-item scale (AUC = 0.853) and comparable to the 55-item scale from which it was originally derived. (AUC = 0.895).

The AD8 was then administered for a 6-month period to a new sample from our longitudinal study at their annual assessment. The AD8 was administered to the CS prior to the clinician's interview with the CS and assessment of the participant. The clinician was not told the answers given by the CS. At the end of the assessment, the clinician generated an independent CDR rating for the participant according to our standard procedures.

Descriptive statistics were used to summarize sample characteristics as well as AD8 scores at each CDR level. In order to demonstrate how the AD8 scores compare with the independently-generated CDR rating and brief office-based objective measures, correlations between AD8, CDR, MMSE, and SBT scores were determined using Pearson product-moment correlation coefficients. We hypothesized that higher AD8 scores would represent more severe stages of dementia and correlate with higher CDR stages and worsening performance on the MMSE and SBT. ROC curves and AUC were generated to reflect graphically and quantitatively the ability of the AD8 to discriminate between participants with CDR 0 and CDR 0.5 and between those with CDR 0 and CDR $\geq$ 0.5. Classification tables were constructed, and the sensitivity, specificity, positive and negative predictive values of the AD8 were calculated. All statistical analyses were conducted using SAS.

Results

The AD8 was administered to 236 participants between October 6, 2003, and April 2, 2004. The participant's mean age at time of assessment was 78.1 $\pm$ 9.2 y (range 55-102), and

53% were women. The collateral sources were spouses (52%), children (29%), friends (10%), and other sources such as social workers, case managers, and health aides (9%). The participant's cognitive status ranged from nondemented (47% CDR 0) through all stages of dementia (very mild, CDR 0.5 = 29%; mild, CDR 1 = 18%; moderate, CDR 2 =5%; and severe, CDR 3 = 1%). Clinical diagnoses from the standard assessment included nondemented elders (47% of sample), uncertain or questionable dementia (12%), and demented individuals (41%). Dementia diagnoses were largely made up of DAT; however other dementia diagnoses included mixed dementia, VaD, and DLB. The AD8 was administered to the CS by the center's secretarial staff, taking on average less than 3 minutes to complete.

AD8 Scores compared with CDR stages

Total AD8 scores were compared by CDR stages determined independently by experienced clinicians. Those individuals who were rated as CDR 0 had a mean AD8 score of 0.6 (range 0-5) compared with individuals who were rated as having at least very mild dementia (CDR 0.5 or greater) who had a mean score of 2.91 or greater (range 0-8).

Discriminative Ability of the AD8

ROC curves were generated to measure the effectiveness of the AD8 in classifying CDR 0 (nondemented) versus CDR 0.5 (very mild dementia) participants. The AUC is 83%, suggesting good to excellent ability to discriminate between CDR 0 and CDR 0.5 groups. Using a cut-off score of 2 or greater on the AD8 to predict dementia yielded the most desirable combination of sensitivity (74%) and specificity (86%). With a prevalence of 38%, the positive predictive value (the probability that someone with an AD8 score $\geq$ 2 has dementia) was 76%, whereas the negative predictive value (the probability that someone with an AD8 score < 2 is nondemented) was 84%.

When predicting membership in the demented group (CDR $\geq$ 0.5) versus status as a nondemented participant (CDR 0), the ROC area under the curve was 90% suggesting excellent discriminative ability. Sensitivity increased to 85%, with specificity remaining at 86%. The positive predictive value increased (87%) and negative predictive value (84%) remained essentially the same with the increased prevalence rate (53%) and inclusion of more severely demented individuals.

Correlation of AD8 to other measures of cognitive status

Strong correlations were demonstrated between the AD8 and the CDR staging ($r = 0.74$, $p<0.0001$) and between AD8 and SBT scores ($r = 0.58$, $p<0.0001$). An inverse correlation was found between the AD8 and MMSE scores ($r = -0.64$, $p<0.0001$).

Discussion, including implications, intended next steps, and potential long-term extensions

The AD8 is a brief informant-based measure that reliably differentiates nondemented from demented individuals and is sensitive to the earliest signs of cognitive change as reported by an informant. The AD8 is highly correlated with our gold standard, the CDR, as well as performance on brief objective measures such as the MMSE and SBT. The AD8 took the CS less than three minutes to complete and was administered by clerical staff. The goals of any screening test are to separate people with a high probability of having the disease from those with a low probability and to presumptively identify unrecognized disease. Diagnostic confirmation is generally required. An effective screening test should be inexpensive, and its characteristics should include reliability, sensitivity, specificity, social acceptability, safety, and brevity. We have presented data indicating that the AD8 meets all of these requirements and discriminates cognitively healthy older adults from individuals at the very earliest symptomatic stage of dementing illnesses. The AD8 is currently being used as part of a brief dementia detection instrument (in conjunction with a patient assessment) to evaluate patients in community settings. If the instrument were used to

select research participants, a different cut-off score might be chosen depending on the costs of identifying individuals (sensitivity) versus the costs of screening and excluding them (specificity).

A number of brief screening measures are already in use, but most are based on patient performance and are therefore unable to detect or quantify change from previous levels of function. Some performance-based measures are also insensitive to subtle changes in high functioning individuals who may score well within the normal range through the early stages of dementia. These same measures also may prevent detection of dementia in individuals with poorer lifelong abilities. Further, many cognitive tests are culturally insensitive and may underestimate the abilities of African-American and other minority groups.

Informant-based assessments provide an opportunity for the clinician to assess change from the patients' prior level of function and determine interference with the patient's accustomed functioning in daily tasks. The use of an informant permits the use of patients as their own control while eliminating the need for baseline assessments. The time required to complete the informant interview to derive the CDR, however, is greater than an average office visit permits. Other informant-based interviews such as the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) are also too lengthy for general use. The Observation List for early signs of Dementia (OLD) is a brief 12-item measure; however it is performed by the physician, rather than caregivers or family members. The physician does not have as much opportunity to observe the patient as does a family member.

A potential drawback of the AD8 is that knowledgeable informants may not be readily available. The validity of using informant-based assessments has been addressed by several authors. In general, informants who live with the patient are able to give more accurate reports than informants who do not. Spouses are generally more accurate than other informant relationship types. The patient's educational level, social status, or neuropsychiatric symptoms do not appear to affect informant accuracy. In the study we report data from several different informant relationships, ranging from spouse to close friends, social workers or paid caregivers.

The sample for our study was not population-based; as with any volunteer sample, selection biases limit generalization of the results. Our sample was largely Caucasian, so it is unknown if these results generalize to other ethnicities. Because the AD8 only requires comment on observable change in the patient's cognitive abilities, however, it is less apt to be biased by gender, education, or ethnicity. Our convenience sample includes community volunteers rather than referrals from memory disorder clinics, and the participants' gender and education attributes reflect those of the similarly aged population in the St Louis metropolitan area. The sample is well characterized, enabling comparison of the AD8 to the longer semistructured interview used to derive the CDR as a gold standard.

We have completed testing the psychometric properties of the AD8 in a clinic population alone and in combination with brief objective measures of the participant's cognitive function and this manuscript is in preparation. We are near completion of a secondary aim, to test the ability of the AD8 as a self-rating tool. The use of the AD8 in conjunction with a brief cognitive assessment could improve diagnostic accuracy in general practice and may be applicable for dementia screening in clinical trials, community surveys, and epidemiological studies.