Screening for predementia AD

Time-dependent operating characteristics of episodic memory tests

Carol A. Derby, PhD
Leah C. Burns, MPH
Cuiling Wang, PhD
Mindy J. Katz, MPH
Molly E. Zimmerman,
PhD
Gilbert L'Italien, PhD
Zhenchao Guo, PhD
Robert M. Berman, MD
Richard B. Lipton, MD

Correspondence to
Dr. Derby:
carol.derby@einstein.yu.edu

ABSTRACT

Objective: Data from the Einstein Aging Study (EAS) were used to prospectively evaluate the free recall score from the Free and Cued Selective Reminding Test (FCSRT-FR) and Logical Memory I immediate recall (LM-IR) subtest of the Wechsler Memory Scale–Revised for prediction of incident Alzheimer disease (AD) dementia among individuals from a community-based cohort with memory complaints.

Methods: Analyses included 854 participants, age ≥ 70 years, who initially had no dementia, and had memory complaints. Clinic evaluations were completed annually and AD dementia was diagnosed using standard criteria ($n = 86$ cases; average follow-up 4.1 years). Time-dependent receiver operating characteristic analysis was used to evaluate the prognostic ability of FCSRT-FR and LM-IR for incident AD over various durations of follow-up.

Results: For identifying those with memory complaints who will develop incident AD dementia over 2–4 years, the FCSRT-FR had better operating characteristics than LM-IR. *APOE* $\epsilon 4$ status, age, and education did not affect cut points; however, positive predictive values were higher among *APOE* $\epsilon 4$ -positive individuals.

Conclusions: For follow-up intervals of 2–4 years, the FCSRT-FR is more predictive than the LM-IR for identifying individuals with memory complaints who will develop incident AD. *APOE* $\epsilon 4$ status improves positive predictive value, but does not affect the choice of optimal cuts. *Neurology*® 2013;80:1307–1314

GLOSSARY

AD = Alzheimer disease; **AUC** = area under the receiver operating characteristic curve; **CDR** = Clinical Dementia Rating; **CERAD** = Consortium to Establish A Registry for Alzheimer's Disease; **CI** = confidence interval; **DSM-IV** = *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition; **EAS** = Einstein Aging Study; **FCSRT-FR** = free recall score from the Free and Cued Selective Reminding Test; **HSA** = Health Self-Assessment Form; **LM-IR** = Logical Memory I immediate recall; **MCI** = mild cognitive impairment; **PPV** = positive predictive value; **ROC** = receiver operating characteristic.

Diagnostic guidelines for Alzheimer disease (AD) have been revised to reflect the view of the disease as a continuum.^{1,2} The long preclinical course of AD dementia suggests the potential for preventive measures aimed at the preclinical stages, and the search for biomarkers to identify asymptomatic individuals at high risk is underway. In the absence of definitive biomarkers to facilitate primary prevention, secondary prevention trials target persons without dementia with memory impairment and aim to delay or prevent onset of clinical AD dementia.^{1,2} A critical step in the design of these trials is the ability to discriminate between those with memory complaints attributable to underlying AD and those with impairment due to other causes.^{3–5} There are different approaches to conceptualizing the symptomatic predementia phase of AD. Definitions of mild cognitive impairment (MCI) rely on norming cognitive performance relative to an age- and education-matched sample. However, if the goal is to identify individuals with high probability of developing AD dementia, the ideal cognitive cut scores for cognitive screening could be determined prospectively based on ability to predict incident AD dementia. Demographic and genetic covariates might enhance the predictive ability of these longitudinally derived screening cut scores. Because intervention studies may involve various

Editorial, page 1274

Supplemental data at
www.neurology.org

From the Department of Neurology, Division of Cognitive Aging and Dementia and the Einstein Aging Study (C.A.D., C.W., M.J.K., M.E.Z., R.B.L.), and Department of Epidemiology and Population Health (C.A.D., C.W., R.B.L.), Albert Einstein College of Medicine, Bronx, NY; Global Health Economics and Outcomes Research (L.C.B.), Bristol-Myers Squibb R&D, Princeton, NJ; and Bristol-Myers Squibb R&D (G.L., Z.G., R.M.B.), Wallingford, CT.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

lengths of follow-up, it is important to determine whether the cut scores that optimize prediction depend upon length of follow-up.

The Einstein Aging Study (EAS) cohort provides the opportunity to evaluate the ability of 2 previously validated tests, the free recall component of the Free and Cued Selective Reminding Test (FCSRT-FR)^{6,7} and Logical Memory I immediate recall (LM-IR) of the Wechsler Memory Scale–Revised,⁸ to identify which community-based older individuals with cognitive complaints are likely to develop AD dementia over specified follow-up periods. We determine the sensitivity, specificity, and positive predictive values, the cut points that optimize prediction of incident AD dementia over 2–4 years of follow-up, and assess whether *APOE* ϵ 4 status, a known genetic risk factor for AD dementia,⁹ age, or education level affect test operating characteristics.

METHODS Study population. The analysis is based on longitudinal data from the EAS. The cohort includes systematically recruited adults from a multiethnic, community-dwelling population in Bronx County, New York. Enrolled participants were at least 70 years of age, Bronx residents, noninstitutionalized, and English-speaking. Exclusions included visual or auditory impairments, active psychiatric symptomatology that interfered with assessments, and nonambulatory status.¹⁰

Standard protocol approvals, registrations, and patient consents. Written informed consent was obtained according to protocols approved by the institutional review board.

Clinical assessments. In-person evaluations were completed at baseline and annually. Assessments included demographic information, medical history, medications, and health behaviors. Participants underwent a standard neurologic examination adapted from the Unified Parkinson's Disease Rating Scale.¹¹ The neurologist documented clinical impression of cognitive impairment using the Clinical Dementia Rating (CDR).¹² Subjective cognitive complaints were ascertained using the Clinical History–Cognitive Impairment/Dementia Questionnaire of the Consortium to Establish A Registry for Alzheimer's Disease (CERAD)–self¹³ and the Albert Einstein Health Self-Assessment Form (HSA), which inquires about current memory problems and changes in memory compared to 1 and 10 years ago. Global cognitive status was ascertained by the Blessed Information-Memory-Concentration test.¹⁴ Memory was measured using the picture version of the FCSRT-FR⁶ and the LM-IR test.⁸ Each test is scored such that a higher score indicates better performance (FCSRT-FR range = 0–48; LM-IR range = 0–50).

The FCSRT begins with a study phase in which participants are asked to name pictures of 16 common items. They are then presented with these pictures in a series of pages, each with 4 objects arranged in a grid. Participants orally identify each picture following a categorical prompt. After all 16 are correctly identified, immediate recall is tested by removing the visual stimulus. In the free recall condition, the participant recalls as many of the pictures as possible. If the participant fails to recall a picture, he or she is provided with a category cue to test cued recall. There are 3 free and 3 cued recall trials. A 20-second interference

(counting backward) is administered between each.^{6,7} This analysis utilizes the sum of free recall across the 3 trials (score range 0–48). Free recall has been shown to be predictive of incident dementia, and has been correlated with hippocampal measures in the EAS sample.^{6,15} Free rather than total recall (sum of free and cued items recalled) was selected as total recall demonstrates ceiling effects in our community-based sample, and because FCSRT-FR has been shown to be independent of education and race.⁶

The LM-IR assesses the ability to learn and retrieve contextually related verbal information through immediate free recall of 2 short stories, each containing 25 elements of information.⁸

Dementia diagnosis. Dementia diagnosis was based on standardized criteria from the *DSM-IV*¹⁶ and required impairment in memory plus at least one additional cognitive domain, accompanied by functional decline. Diagnoses were assigned at consensus case conferences which included comprehensive review of all cognitive tests, neurologic signs and symptoms, and functional status by 2 neurologists and a neuropsychologist. Memory impairment was defined as scores in the impaired range on any memory test in the neuropsychological battery (FCSRT-FR $\leq 24^6$ or 1.5 SDs below the age-adjusted mean on LM-IR). Functional impairment was determined by scores on the Lawton Brody Scale,¹⁷ clinical evaluation, and informant questionnaires. AD dementia was diagnosed in dementia cases meeting established criteria for probable or possible AD.¹⁸ AD refers to AD alone or in combination with other dementia disorders. A subset of individuals who participated in the EAS have come to autopsy, providing an important quality control for diagnostic accuracy. Based on an autopsy sample of 175, a clinical diagnosis of dementia had a positive predictive value (PPV) of 96% for significant pathology upon autopsy. A clinical diagnosis of possible or probable AD had a PPV of 79% for the presence of NIA-Reagan¹⁹ intermediate or high likelihood Alzheimer-type pathology, while 71% of those with Braak stage ≥ 3 had been assigned a clinical diagnosis of AD.

Because both LM and FCSRT are part of the neuropsychological battery reviewed at consensus case conference, there is potential for diagnostic circularity to bias this assessment of the tests. Therefore we repeated analyses using CDR > 1.0 to define dementia independent of these measures. Results using either outcome were not materially different, suggesting that circularity did not confound our results. Furthermore, diagnoses are based on the entire range of neuropsychological and neurologic assessments across multiple time points, while this analysis is comparing prediction of baseline memory assessments. The analyses presented are for the outcome of AD diagnosis. Results of analyses using CDR diagnosis are presented as supplemental tables on the *Neurology*[®] Web site at www.neurology.org for reference.

***APOE* ϵ 4 status.** DNA was extracted from whole blood, or was isolated from buffy coat that had been stored at -70°C using the Puregene DNA Purification System (Gentra System, Minneapolis, MN). Amplification and sequencing primers for genotyping of the target *APOE* single nucleotide polymorphisms, rs429358 (position 112) and rs7412 (position 158), were designed using PSQ version 1.0.6 software (Biotage, Uppsala, Sweden); in each case, the reverse primer was biotinylated. Genotyping was performed using a Pyrosequencing PSQ HS 96A system (<http://www.pyrosequencing.com>). Individuals were classified as *APOE* ϵ 4 carriers if they had at least one ϵ 4 allele.

Statistical analysis. Baseline characteristics were reported using descriptive statistics. The Wilcoxon rank sum test for continuous variables and χ^2 test for categorical variables were used for the comparison between those who developed AD during follow-up and those who did not. To evaluate the prognostic capacity of baseline FCSRT-FR and LM-IR for predicting incident AD over various durations of follow-up, we applied time-dependent receiver operating characteristics (ROC) analysis²⁰ using the Nearest Neighbor Estimation²¹ of the joint

distribution of the cognitive test and time to AD. The method can accommodate censored data. Areas under the ROC curves (AUC) were calculated. Bootstrap techniques were used for confidence interval (CI) estimation and hypothesis testing. We sampled the observations in the data with replacements for 1,000 bootstrap samples. For each sample, we computed AUC for predicting AD dementia at 2–4 years of follow-up for each test as well as the difference in AUC between the 2 tests. ROC curves could not be fit for a follow-up interval of 1 year, as there were insufficient cases that developed in this interval. The 95% CI for each AUC and the difference was then obtained. Youden's index, the sum of sensitivity and specificity minus one, is used to select the optimal cutoff value.²² The independent contributions of FCSRT-FR and LM-IR were assessed in Cox proportional hazards models adjusted for age, sex, education, and race. The risk score from each model was calculated as a weighted sum of the predictors in the model, with weights determined by the estimated coefficients of the Cox model. Time-dependent ROC analysis was then applied to the risk scores. Interactions of each cognitive test with age, education, and *APOE* ϵ 4 status were tested in separate Cox models to determine whether these factors alter the relation of cognitive test to incident AD dementia.

RESULTS A total of 1,187 participants had at least one annual follow-up evaluation, and were free of dementia at baseline. Of these, 896 reported memory complaints on either the CERAD or the HSA, and 854 had complete data on FCSRT-FR and LM-IR and were included in the analysis. Analyses including *APOE* ϵ 4 status are based on the subset of 564 individuals with available information on *APOE* ϵ 4 status. The groups with and without genotyping were similar with respect to age, sex, ethnicity, medical comorbidities, baseline FCSRT-FR and LM-IR, and risk of incident AD dementia.

Those who developed incident AD dementia were slightly older at baseline, with less education and worse performance on baseline FCSRT-FR and LM-IR, but were not significantly different with respect to sex, race, medical index, and follow-up time (table 1).

Results of time-dependent ROC analyses are shown in table 2. These data allow for comparisons of specificity of each test for a fixed level of sensitivity, and vice versa. FCSRT-FR consistently shows higher specificity for a given sensitivity, and also demonstrates higher sensitivity for a given specificity. Based on Youden's index,²¹ the FCSRT-FR and LM-IR cuts that optimize prediction of incident AD dementia are similar over the range of 2–4 years of follow-up. For FCSRT-FR, cuts were 2 years, 28; 3 years, 29; 4 years, 27. LM-IR cuts that optimize prediction of incident AD over 2, 3, or 4 years are 16, 17, and 16, respectively.

Figure 1 shows the time-dependent ROC curves for FCSRT-FR and LM-IR for prediction of AD over 2, 3, or 4 years of follow-up, along with optimal cuts derived by Youden's index, and the corresponding AUC. Regardless of length of follow-up, the ROC curve for FCSRT-FR dominates that for LM-IR across the entire range of false-positive rates, demonstrating that for any fixed specificity, FCSRT-FR is a more sensitive marker for AD. AUCs at 2, 3, and 4 years of follow-up are 0.87, 0.88, and 0.89, respectively, for FCSRT-FR and 0.78, 0.77, and 0.75 for LM-IR. The bootstrap 95% CIs for the difference in AUC between FCSRT-FR and LM-IR at years 2, 3, and 4 are (0.025, 0.18), (0.036, 0.16), and (0.070, 0.21), respectively, indicating a statistically

Table 1 Descriptive characteristics of analysis population, by incident Alzheimer dementia status during follow-up, Einstein Aging Study cohort

Baseline characteristics	Analysis sample (n = 854)	Developed incident AD (n = 86)	No AD (n = 768)	p Value for AD vs no AD
Age, y, mean $\pm$ SD	78.8 $\pm$ 5.4	81.4 $\pm$ 5.7	78.5 $\pm$ 5.3	<0.0001
Female, %	60.5	61.6	60.4	0.83
Race, %				
Caucasian	69.0	67.4	69.1	0.95
Black	26.6	27.9	26.7	
Other	4.3	4.7	4.3	
Education, y, %				
<6	3	3.5	3.0	0.38
7–12	43	50.0	42.7	
>12	54	46.5	54.3	
FCSRT-FR, mean $\pm$ SD	31.0 $\pm$ 5.6	23.9 $\pm$ 7.3	30.3 $\pm$ 6.2	<0.0001
LM-IR, mean $\pm$ SD	19.8 $\pm$ 7.0	15.4 $\pm$ 6.8	19.4 $\pm$ 7.1	<0.0001
Years of follow-up, mean $\pm$ SD ^a	4.2 $\pm$ 2.9	3.7 $\pm$ 2.9	4.1 $\pm$ 2.9	0.0895
<i>APOE</i> ϵ 4+, % ^b	21.9	35.1	23.2	0.0253

Abbreviations: AD = Alzheimer disease; FCSRT-FR = free recall score from the Free and Cued Selective Reminding Test; LM-IR = Logical Memory I immediate recall.

^aTime from baseline to dementia diagnosis.

^b*APOE* ϵ 4 information available for 564 subjects.

Table 2 Cut scores, sensitivity, specificity, and positive predictive values for predicting AD with the FCSRT-FR and LM-IR tests for specified follow-up times among persons with memory complaints, Einstein Aging Study Cohort (n = 854)

Cut scores	2 Years, %			3 Years, %			4 Years, %		
	Sensitivity	Specificity	PPV	Sensitivity	Specificity	PPV	Sensitivity	Specificity	PPV
FCSRT-FR									
32	94.1	47.0	7.4	96.2	48.2	11.5	97.2	49.6	16.0
31	91.3	55.7	8.4	94.3	57.1	13.3	95.9	58.8	18.6
30	89.7	60.8	9.3	92.0	62.2	14.5	93.8	64.0	20.4
29	88.3	68.6	11.1	87.6 ^a	70.0 ^a	16.9 ^a	89.6	71.8	23.8
28	85.7 ^a	73.8 ^a	12.8 ^a	82.1	75.1	18.6	84.9	77.0	26.5
27	80.8	78.4	14.3	76.4	79.6	20.6	80.9 ^a	81.7 ^a	30.0 ^a
26	73.1	82.6	15.8	67.9	83.6	22.3	73.2	85.7	33.1
25	67.6	85.5	17.2	62.8	86.5	24.4	66.7	88.3	35.6
24	63.3	88.5	19.7	58.2	89.4	27.6	59.5	90.9	38.7
23	59.1	91.3	23.3	54.6	92.3	32.8	52.5	93.4	43.4
22	53.7	94.2	28.9	49.7	95.1	40.8	44.1	95.8	49.8
LM-IR									
22	94.4	39.9	6.6	92.9	40.7	9.9	88.2	41.1	12.9
21	92.4	46.0	7.1	90.8	46.9	10.7	83.7	47.2	13.6
20	89.9	50.7	7.6	88.6	51.7	11.3	80.4	51.9	14.2
19	83.5	57.0	8.0	84.4	58.1	12.3	76.3	58.5	15.4
18	79.2	61.5	8.4	80.8	62.6	13.1	72.9	63.0	16.3
17	74.6	66.9	9.2	76.7 ^a	68.1 ^a	14.4 ^a	69.8	68.7	18.0
16	71.9 ^a	71.1 ^a	10.0 ^a	72.3	72.2	15.4	66.7 ^a	72.9 ^a	19.5 ^a
15	66.7	75.1	10.7	65.1	76.1	16.0	61.4	76.8	20.7
14	57.5	80.7	11.8	52.2	81.3	16.3	50.8	82.1	21.8
13	49.5	84.0	12.2	43.9	84.5	16.5	44.1	85.3	22.7
12	43.2	86.9	12.9	38.9	87.4	17.7	38.8	88.1	24.3
11	35.1	89.5	13.0	32.9	89.9	18.6	33.0	90.6	25.6

Abbreviations: AD = Alzheimer disease; FCSRT-FR = free recall score from the Free and Cued Selective Reminding Test; LM-IR = Logical Memory I immediate recall; PPV = positive predictive value.

^a Sensitivity, specificity, and positive predictive values for Youden's index cuts that optimize prediction of incident AD dementia over 2, 3, or 4 years of follow-up.

significant difference in classification accuracy at each point ($p \leq 0.05$).

Table 3 presents the results of Cox models predicting incident AD dementia. When added to a model with FCSRT-FR, LM-IR was only moderately associated with incident AD (models 1, 2). In the subset of participants with *APOE* $\epsilon 4$ status, having an *APOE* $\epsilon 4$ allele was significantly associated with risk of AD dementia, but this did not attenuate the association of FCSRT-FR with the outcome (models 3–5). ROC curves for the risk scores derived from the Cox models demonstrate that the addition of LM-IR and *APOE* $\epsilon 4$ status to FCSRT-FR does not appreciably improve the diagnostic accuracy of incident AD over that based on FCSRT-FR (figure 2).

To test whether test performance differs by age, education, or *APOE* $\epsilon 4$ status, we tested additional Cox models that included interaction terms for these variables

with either FCSRT-FR or LM-IR. None of the interactions tested were significant (p for interactions all > 0.10). Furthermore, stratified ROC analyses showed similar operating characteristics for LM-IR and FCSRT-FR over strata defined by age, education, or *APOE* $\epsilon 4$ status. The Youden's index optimal cuts were similar to those derived for the entire cohort (FCSRT-FR cuts 27–29 and LM-IR cuts 16 and 17) and were consistent across strata. However, the positive predictive power of either test was somewhat higher in the group with an *APOE* $\epsilon 4$ allele.

DISCUSSION The goal of this analysis was to prospectively evaluate the predictive validity of 2 commonly used memory tests, the FCSRT-FR and the LM-IR, for identifying which individuals with memory complaints will develop incident AD dementia. To date, no community-based study has prospectively compared the predictive

Figure 1 Receiver operating characteristic curves and areas under the curve for free recall score from the Free and Cued Selective Reminding Test and Logical Memory test immediate recall predicting Alzheimer dementia over 2–4 years

Optimal cuts based on Youden's index.²¹ FCSRT = Free and Cued Selective Reminding Test, free recall.

ability of these 2 tests. Furthermore, the extent to which duration of follow-up and *APOE* $\epsilon 4$ status affect the choice of screening cut points has not been evaluated.

For follow-up times of 2–4 years, the cut scores that maximized prediction were fairly consistent for FCSRT-FR or LM-IR. However, when the goal is to distinguish those who will develop AD dementia, the FCSRT-FR appears preferable. FCSRT-FR has a higher specificity for a given value of sensitivity and ROC analysis demonstrated consistently greater AUCs for FCSRT-FR compared with LM-IR. These findings are consistent with prior studies showing that FCSRT-FR has high discriminative validity for dementia.^{3,6} Recently, the FCSRT total recall score has been shown to be superior to LM for discriminating between MCI cases with and without AD biomarkers.²³ We examined free recall given that our sample was not selected for MCI and thus exhibited ceiling effects for total recall. Although both the LM-IR and FCSRT-FR assess short-term declarative verbal memory, the FCSRT-FR is unique in that it utilizes control of attention and strategy use to maximize performance and improve encoding and retrieval of learned information. In addition, when applied to a community-based sample, the test does not require age- and education-specific norming, and does not suffer from ceiling effects.⁶

The stratified analyses suggest that *APOE* $\epsilon 4$ status does not affect the choice of screening cut point for either test. Although *APOE* $\epsilon 4$ was a significant predictor of AD dementia, the addition of *APOE* $\epsilon 4$ status to a prediction model did not improve the diagnostic accuracy compared to using the FCSRT-FR alone, adjusting for demographic factors, and the relation of either FCSRT-FR or LM-IR to dementia incidence did not vary by *APOE* $\epsilon 4$ status.

These data are useful for the design of clinical trials, as they provide operating characteristics of 2 memory tests widely used for determining which patients are likely to progress to dementia in a relatively short time horizon. The ability to select eligibility criteria derived prospectively over specified time frames should increase efficiency of screening for participants with memory complaints who are likely to develop dementia over specified follow-up periods. The information provided allows for comparison of the tradeoffs between sensitivity and specificity. This balance depends upon the context for the screening. In a situation where potential study participants will undergo a 2-stage screen, the choice would depend on the extent to which the second screen is costly or invasive. When the second test is relatively noninvasive and inexpensive, the choice may be to maximize sensitivity of the cognitive screen and follow it with a more specific assay to exclude persons with memory complaints not related to underlying AD dementia. If, however, the second test is costly or invasive, such as PET scanning or lumbar

Table 3 Hazard ratios from Cox proportional hazards models predicting incident AD, Einstein Aging Study Cohort^a

	Model 1	Model 2	Model 3	Model 4	Model 5
One unit decrease in FCSRT-FR	1.25 (<0.0001)	1.23 (<0.001)	1.26 (<0.0001)	1.24 (<0.0001)	1.23 (<0.0001)
One unit decrease in LM-IR		1.04 (0.03)		1.04 (0.06)	1.04 (0.07)
APOE ε4+					1.91 (0.03)

Abbreviations: AD = Alzheimer disease; FCSRT-FR = free recall score from the Free and Cued Selective Reminding Test; LM-IR = Logical Memory I immediate recall.

^a Age: <80 vs ≥80 years. Education: ≤high school vs >high school. Among the total sample (n = 864): model 1: FCSRT-FR adjusted for age, sex, education, race; model 2: model 1 + LM-IR. Among the subgroup with APOE ε4 status available (n = 564): model 3: FCSRT-FR adjusted for age, sex, education, race; model 4: model 3 + LM-IR; model 5: model 4 + APOE ε4 status.

puncture, the goal may be to maximize specificity of the cognitive screen to minimize secondary testing of those unlikely to develop incident dementia. Such tradeoffs must also be evaluated in the context of risk vs benefit for the intervention. In the clinical setting, the information provided here has utility for estimating the time frame for follow-up screening. In this context, the goal may be to maximize the PPV.

A strength of this study is that the population was community-based and not derived from a clinic setting. Thus, results may be generalizable to trials enrolling subjects from the community. The EAS cohort is representative of the Bronx, New York population with respect to demographics and is clinically well-characterized.¹⁰ We were able to prospectively assess the operating characteristics of the screening tests with respect to specified lengths of follow-up. This

comparison is restricted to the FCSRT-FR and LM-IR tests. While these are widely used, we acknowledge that a number of other instruments exist; however, they were not in our test battery. In particular, other studies utilize the delayed rather than the immediate recall condition of the LM.²⁴ This was not available for this analysis, and results should be interpreted with this distinction in mind.

The consistency of our results using either case conference diagnosis or CDR diagnosis alleviates concerns regarding diagnostic circularity (tables e-1 and e-2 and tables 2 and 3). In addition, case conference diagnosis is based on information from a broad neuropsychological battery including other tests of episodic memory, a clinical evaluation, informant reports, as well as neurologic examinations and functional data. The influence of any single cognitive test at the baseline assessment is thus likely to be small.

Figure 2 Receiver operating characteristic curves for prediction scores derived from Cox models in table 3, for prediction of Alzheimer dementia over 3 years of follow-up

AUC = area under the curve; CI = 95% bootstrap confidence intervals around areas under the curve; cov = covariates in Cox models (age, sex, race, education); FCSRT = Free and Cued Selective Reminding Test, free recall; LM = Logical Memory test.

For distinguishing individuals with memory complaints who will develop dementia over 2–4 years, the FCSRT-FR test is a better predictor than LM-IR. These data are useful for planning clinical trials with fixed duration, or for determining the follow-up intervals for rescreening in the clinical setting. While *APOE* $\epsilon 4$ is a significant predictor of incident AD dementia, the choice of cut scores did not vary by *APOE* $\epsilon 4$ status. Thus, when screening for a clinical trial, the cognitive test alone may be sufficient. In the clinical setting, knowledge of *APOE* $\epsilon 4$ status increases the PPV and thus may be of benefit for the patient.

AUTHOR CONTRIBUTIONS

Dr. Derby participated in study concept and design, design and interpretation of the statistical analysis, and in drafting and revising the manuscript. Ms. Burns participated in study concept and design, revising the manuscript for content, and interpretation of data. Dr. Wang participated in the design of the analyses, performed the statistical analyses, and participated in the interpretation of results and the drafting and revising of the manuscript. Ms. Katz participated in study design, data collection, conceptualizing the analysis plan, interpreting the statistical results, and revising the manuscript. Dr. Zimmerman participated in the data collection and interpreting the statistical results and revising the manuscript. Dr. L'Italien participated in the analysis and interpretation of data. Dr. Guo participated in drafting and revising the manuscript for content, including medical writing for content, the analysis and interpretation of data, and statistical analysis of data. Dr. Berman participated in drafting and revising the manuscript for content, including medical writing for content, and the analysis and interpretation of data. Dr. Lipton participated in study design, data collection, conceptualizing the analysis plan, interpreting the statistical results, and revising the manuscript.

ACKNOWLEDGMENT

The authors thank Charlotte Magnotta, Diane Sparracio, and April Russo for assistance with participant recruitment; Betty Forro, Alicia Gomez, Wendy Ramratan, and Mary Joan Sebastian for assistance in clinical and neuropsychological assessments; Michael Potenza for assistance with data management; and all of the study participants who generously gave their time in support of this research.

STUDY FUNDING

Supported by a research grant from Bristol-Myers Squibb, the S and L Marx Foundation, and the National Institutes of Health (NIA-AG03949).

DISCLOSURE

C.A. Derby receives research support from NIH P01 AG039409 (project leader), P01AG027734 (coinvestigator), 5R01 AG22374 (site principal investigator), 2U01AG012535-16 (principal investigator), NCR R 5UL1RR025750-03 (investigator), Bristol-Myers Squibb, and the S and L Marx Foundation. L.C. Burns is an employee of Bristol-Myers Squibb. C. Wang receives research support from NIH P01 AG039409 (investigator), R01 AG039330 (investigator), R01 AG036921 (investigator), R01 HL094581 (investigator), and Bristol-Myers Squibb. M.J. Katz receives research support from NIH-NIA P01 AG027734 (investigator), R01 AG022374 (investigator), NIH-NIA R01 AG034119 (investigator), NIH-NIA R01 AG012101 (investigator), NIH-NIA P01 AG039409 (investigator), NIH-NIA AG038651 (investigator), and Bristol-Myers Squibb. M.E. Zimmerman receives research support from the NIH [P01 AG03949 (coinvestigator), P01AG027734 (coinvestigator)] and the Alzheimer's Association [NIRG-11-206369 (principal investigator)] and has received research support from Bristol-Myers Squibb and Merck. G. L'Italien is an employee of and holds stock options from Bristol-Myers Squibb. Z. Guo is an employee of and holds stock options from Bristol-Myers Squibb. R.M. Berman is an employee of and holds stock options from Bristol-Myers Squibb. R.B. Lipton receives research support from the NIH [P01 AG03949 (program director), P01AG027734 (project leader), RO1AG025119 (investigator), RO1AG022374-06A2 (investigator), RO1AG034119 (investigator),

RO1AG12101 (investigator), K23AG030857 (mentor), K23NS05140901A1 (mentor), and K23NS47256 (mentor)] and the S and L Marx Foundation; serves on the editorial boards of *Neurology*[®] and as senior advisor to *Headache*; has reviewed for the NIA and NINDS; holds stock options in eNeura Therapeutics; and serves as consultant, advisory board member, or has received research support from Allergan, American Headache Society, Autonomic Technologies, Boston Scientific, Bristol-Myers Squibb, Cognimed, Colucid, Eli Lilly, Endo, eNeura Therapeutics, GlaxoSmithKline, MAP, Merck, Nautilus Neuroscience, Novartis, NuPathe and Pfizer. Go to Neurology.org for full disclosures.

Received June 13, 2012. Accepted in final form November 8, 2012.

REFERENCES

1. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:280–292.
2. Jack CR, Albert MS, Knopman DS, et al. Introduction to the recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:257–262.
3. Sarazin M, Berr C, DeRotrou J, et al. Amnesic syndrome of the medial temporal type identifies prodromal AD: a longitudinal study. *Neurology* 2007;69:1859–1867.
4. Jelic V, Kivipelto M, Winblad B. Clinical trials in mild cognitive impairment: lessons for the future. *J Neurol Neurosurg Psychiatry* 2006;77:429–438.
5. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007;6:734–746.
6. Grober E, Lipton RB, Hall CB, et al. Memory impairment on free and cued selective reminding predicts dementia. *Neurology* 2000;54:827–832.
7. Buschke H. Cued recall in amnesia. *J Clin Neuropsychol* 1984;6:433–440.
8. Wechsler D. Wechsler Memory Scale: Revised. San Antonio: The Psychological Corporation; 1987.
9. Saunders AM. Apolipoprotein E and Alzheimer disease: an update on genetic and functional analyses. *J Neuropathol Exp Neurol* 2000;59:751–758.
10. Katz MJ, Lipton RB, Hall CB, et al. Age-specific and sex-specific prevalence and incidence of mild cognitive impairment, dementia, and Alzheimer dementia in blacks and whites: a report from the Einstein Aging Study. *Alzheimer Dis Assoc Disord Epub* 2011 Dec 7.
11. Fahn S, Elton R. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Goldstein M, et al, eds. *Recent Developments in Parkinson's Disease*. Vol. 2. Florham Park, NJ: Macmillan Health Care Information; 1987:153–163, 293–304.
12. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412–2414.
13. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): part I: clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989;39:1159–1165.
14. Blessed G, Tomlinson BE, Roth M. The association between qualitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry* 1968;114:797–811.
15. Zimmerman ME, Pan JW, Hetherington HP, et al. Hippocampal neurochemistry, neuromorphometry, and verbal memory in nondemented older adults. *Neurology* 2008;70:1594–1600.

16. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, DSM-IV. Washington, DC: American Psychiatric Association; 1994.
17. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179–186.
18. McKhann G, Drachan D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services task force on Alzheimer's disease. *Neurology* 1984;34:939–944.
19. The NIA-Reagan Working Group. Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. *Neurobiol Aging* 1997;18(suppl 4):S1–S2.
20. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;56:337–344.
21. Akritas MG. Nearest neighbor estimation of a bivariate distribution under random censoring. *Ann Stat* 1994;22:1299–1327.
22. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–35.
23. Wagner M, Wolf S, Reischies FM, et al. Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology* 2012;78:379–386.
24. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's disease Neuroimaging Initiative (ADNI) clinical characterization. *Neurology* 2010;74:201–209.

Share Your Artistic Expressions in *Neurology* 'Visions'

AAN members are urged to submit medically or scientifically related artistic images, such as photographs, photomicrographs, and paintings, to the "Visions" section of *Neurology*®. These images are creative in nature, rather than the medically instructive images published in the *NeuroImages* section. The image or series of up to six images may be black and white or color and must fit into one published journal page. Accompanying description should be 100 words or less; the title should be a maximum of 96 characters including spaces and punctuation.

Learn more at www.aan.com/view/Visions, or upload a Visions submission at submit.neurology.org.

Spring AAN Webinars: Help for Your Practice, CME for Your Career

The American Academy of Neurology offers cost-effective Practice Management Webinars that can be attended live or through convenient recordings posted online five days after the event. AAN members can save 25% on all regular webinars! Plus, physicians can earn 1.5 valuable CME credits for each webinar, and administrators receive a certificate of completion. Mark your calendar for upcoming programs and register today for these and other 2013 webinars at www.aan.com/view/pmw13:

- April 9: Correct Coding for Chemodenervation
- April 24: Working with NPs and PAs to Maximize Office Productivity
- May 14: Remaining Relevant in the Changing Health Care Payment and Care Delivery Systems
- June 5: Coding Accurately for Epilepsy
- June 25: E/M: Minimize Mistakes, Maximize Reimbursement