Recognition Memory and Verbal Fluency Differentiate Probable Alzheimer Disease From Subcortical Ischemic Vascular Dementia

Mary C. Tierney, PhD; Sandra E. Black, MD; John P. Szalai, PhD; W. Gary Snow, PhD; Rory H. Fisher, MB; Grant Nadon, MD; Helena C. Chui, MD

Background: Alzheimer disease (AD) and vascular dementia are among the most frequently occurring causes of dementia in the world, and their accurate differentiation is important because different pharmaceutical strategies may modify the course of each disease.

Objective: To determine which of 10 neuropsychological test scores can accurately differentiate patients with probable AD from those with subcortical ischemic vascular dementia (SIVD) for use in evidence-based clinical practice.

Design: Patients with suspected dementia were referred to the study by family physicians, geriatricians, and neurologists. All participants received a thorough assessment according to standard diagnostic guidelines. Diagnoses of probable AD (n=31) and probable SIVD (n=31) were made according to consensus criteria. The diagnosticians were blind to the results of the 10 neuropsychological test scores.

Results: There were no significant differences between the groups in age or Mini-Mental State Examination scores. Logistic regression analyses identified 2 neuropsychological tests that best distinguished the groups (sensitivity=81%; specificity=84%; positive likelihood ratio=5.1). These were the recognition memory subtest of the Rey Auditory Verbal Learning Test and the Controlled Oral Word Association Test. The AD group performed better on the oral association test, whereas the SIVD group did better on the recognition memory test.

Conclusion: Patients with probable AD and probable SIVD can be distinguished with a high degree of accuracy using these 2 neuropsychological tests.

Arch Neurol. 2001;58:1654-1659

A

LZHEIMER disease (AD) and cerebrovascular disease (CVD) are the most common causes of dementia in the world. For several reasons, however, differential diagnosis of AD vs vascular dementia (VaD) has proved difficult. First, there is still no specific biological marker for AD. Second, although structural neuroimaging studies (eg, computed tomography [CT] or magnetic resonance imaging) can detect relatively small infarcts or hemorrhages—the most specific hallmarks of CVD—they may be considered too expensive for epidemiological surveys or unnecessary for diagnosis in managed care settings. Once dementia is identified, a short neuropsychological examination to determine the likelihood of AD vs VaD or AD vs non-AD dementia would be useful. In this study, we use evidence-based methods to assess the utility of neuropsychological testing to differentiate dementia due to AD vs 1 subtype of VaD.

The impairment of memory and learning is one of the earliest signs of AD. This is attributed to the selective vulnerability of medial temporal lobe structures to neurofibrillary degeneration. The pattern of cognitive impairment associated with VaD, on the other hand, is highly variable and depends on the size and location of vascular brain injury. For example, occlusion of the middle cerebral artery may be associated with aphasia if the left hemisphere is involved or with inattention and neglect when the opposite hemisphere is affected. Whereas reasonable success has been achieved in identifying characteristic patterns of neuropsychological impairment in AD, this can be realized in VaD only by confining study to more homogeneous subgroups. Small-artery disease leads to lacunar infarcts in the subcortical gray and white matter. Although the side of hemispheric involvement remains important, more consistent neurobehavioral syndromes (eg, impaired executive function) emerge regardless of location, possibly owing to a shared disruption of frontal subcortical circuits. For these reasons, this study focuses on the neuropsychological differentiation of de-
Participants and Methods

Participants with suspected dementia were referred to the study by family physicians, geriatricians, and neurologists. Figure 1 shows a flow diagram describing the number of participants referred to the study and those meeting inclusion and exclusion criteria. Physicians were asked to refer patients with suspected AD or VaD who did not have any other medical conditions that would account for the cognitive impairment. Informed consent was obtained after the nature of the procedures had been fully explained. An experienced board-certified geriatrician (R.H.F. or G.N.) saw the patients and their caregivers. A standardized interview was administered inquiring about the history and course of cognitive problems as well as daily functioning. A thorough physical examination, a head CT scan, and laboratory tests including hematological, renal, hepatic, and metabolic function tests were conducted. To be included in this study, participants were required to be 60 years or older. The final study sample was selected according to the fulfillment of (1) medical exclusion criteria; (2) dementia inclusion criteria; and (3) dementia subtype inclusion criteria.

Medical Exclusion Criteria

We excluded patients with severe aphasia because of their inability to complete neuropsychological testing. We also excluded subjects if any of the following causes of memory impairment other than AD or VaD were present: chronic alcohol abuse, hypoxia, metabolic disorders, intracranial mass lesions, psychoses, brain trauma, or other neurological disorders including Parkinson disease. Six patients were excluded because they met these medical exclusion criteria.

Criteria for Dementia

Patients who met the medical criteria for the study were administered a neuropsychological test battery consisting of the Wechsler Memory Scale–Revised7: Information, Orientation, and Visual Reproduction, immediate and delayed recall (we also had the patients copy the designs after the delayed recall trials); California Verbal Learning Test8; Boston Naming Test9 (odd-even version); Controlled Oral Word Association Test (COWAT)10 (letters R, P, and W); Category Fluency (animal names); Wechsler Adult Intelligence Scale–Revised11: Digit Span, Similarities, and Digit Symbol subtests; Read Perceptual Closure Test12; Finger Tapping Test13; and Tokens Test14. Diagnoses of dementia were made by an experienced board-certified psychologist (W.G.S.) and were based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition (DSM-III-R).15 Twelve subjects were excluded because they did not meet the criteria for dementia.

Diagnosis of Dementia Subtype

Patients who met the criteria for dementia were further evaluated to determine whether they met the criteria for AD or SIVD. Diagnosticians were blind to the results of the neuropsychological evaluation. All CT scans (of patients with either AD or VaD) were interpreted and reported by clinical neuroradiologists in our institution who were blind to clinical status but aware that patients were being evaluated for memory problems. The Hachinski Ischemic Score16 was obtained for each subject but was not used for diagnosis of subtype.

Probable AD Group

Patients were diagnosed as having AD if they met the criteria of the National Institute of Neurological and Communicative Disorders and Stroke

Continued on next page

Results

Study Sample

The demographic characteristics of patients in the AD and SIVD groups, as well as their scores on the Mini-Mental State Examination (MMSE), the original full-scale Hachinski Ischemic Scale, and the neuropsychological tests are presented in Table 1. There were no differences between the groups in age (t_{60}=0.42; P = .68), level of education (t_{60}=1.86; P = .07), or MMSE score (t_{60}=0.66; P = .51). However, because the educational differences between the groups were marginally significant and because these differences are known to influence cognitive performance, education was used as a covariate in all subsequent analyses. As expected, the mean difference between the groups on the Hachinski Ischemic Scale was significant (t_{38}=13.99; P = .001).

All patients in the SIVD group had subcortical involvement including the basal ganglia (caudate or lenticular nucleus), thalamus, internal capsule, and periventricular white matter region. Nineteen subjects had purely subcortical infarcts (pure SIVD). Twelve had cortical lesions with involvement of the adjacent white matter in addition to subcortical lesions (mixed SIVD). Of the 31 patients in the SIVD group, 9 had unilateral lesions and 22 had bilateral lesions.

Neuropsychological Differences Between Participants with VaD and AD

Logistic regression analysis (forward stepwise procedure) was used to determine which of the tests in the full
neuropsychology battery (10 neuropsychological test scores and the covariate of education) significantly \((P<.05)\) predicted diagnostic classification. Only 3 variables met the criteria for entry into the model: COWAT \(\chi^2=12.22; P<.001\), RAVLT recognition \(\chi^2=9.77; P=.002\), and educational level \(\chi^2=4.31; P=.04\). The direction of the coefficients suggested that poor recognition memory, higher oral fluency scores, and higher education were associated with an increased probability of having AD.

We used the receiver operating characteristic (ROC) curve to determine the optimal combination of sensitivity and specificity for this 3-variable model. The optimal combination simultaneously maximizes both sensitivity and specificity. This analysis yielded a sensitivity of 81% and a specificity of 84% when the cutoff value of the predicted probability of AD was 0.54. The area under the ROC curve for the reduced model was 0.89 (95% confidence limits, 0.85-0.93). The ROC curve for the 3-variable model with associated predicted probabilities of having AD is provided in Figure 2.

The likelihood ratio of a positive test (LR+) for AD can be calculated from these results and is the quotient of 2 measures: (1) the likelihood of obtaining a given test result among people with the target disorder (specificity); and (2) the likelihood of obtaining the same test result among people without the target disorder (1 – sensitivity). The LR determines the extent to which a test result will raise or lower the pretest probability that a target disorder is present. The pretest probability refers to the diagnostician’s judgment about the presence of the disease in a patient before seeing the results of the test. The posttest probability is the diagnostician’s judgment about the presence of the disease after knowing the test results. If the test does not appreciably raise or lower the pretest probability, it is not useful diagnostically. The reader is urged to consult Sackett et al8 for a more detailed description of how to evaluate the usefulness of diagnostic tests. The LR+ for AD on these 2 tests (including the education covariate) is 5.1. Following the guide provided by Jaeschke et al,21 this LR would generate a moderate change in the pretest-to-posttest probability of AD.

We repeated these analyses with the subsample of 19 participants who had pure SIVD. The same pattern emerged: COWAT, RAVLT recognition, and level of education were included as significant variables in the forward-stepping regression model. The classification accuracy of this model was also similar to the model with patients who had mixed SIVD, with a sensitivity of 81% and a specificity of 84%. The optimal combination simultaneously maximizes both sensitivity and specificity. This analysis yielded a sensitivity of 81% and a specificity of 84% when the cutoff value of the predicted probability of AD was 0.54. The area under the ROC curve for the reduced model was 0.89 (95% confidence limits, 0.85-0.93). The ROC curve for the 3-variable model with associated predicted probabilities of having AD is provided in Figure 2.

The likelihood ratio of a positive test (LR+) for AD can be calculated from these results and is the quotient of 2 measures: (1) the likelihood of obtaining a given test result among people with the target disorder (specificity); and (2) the likelihood of obtaining the same test result among people without the target disorder (1 – sensitivity). The LR determines the extent to which a test result will raise or lower the pretest probability that a target disorder is present. The pretest probability refers to the diagnostician’s judgment about the presence of the disease in a patient before seeing the results of the test. The posttest probability is the diagnostician’s judgment about the presence of the disease after knowing the test results. If the test does not appreciably raise or lower the pretest probability, it is not useful diagnostically. The reader is urged to consult Sackett et al8 for a more detailed description of how to evaluate the usefulness of diagnostic tests. The LR+ for AD on these 2 tests (including the education covariate) is 5.1. Following the guide provided by Jaeschke et al,21 this LR would generate a moderate change in the pretest-to-posttest probability of AD.

We repeated these analyses with the subsample of 19 patients who had pure SIVD. The same pattern emerged: COWAT, RAVLT recognition, and level of education were included as significant variables in the forward-stepping regression model. The classification accuracy of this model was also similar to the model with patients who had mixed SIVD, with a sensitivity of 81%
and a specificity of 79% (LR+ = 3.9) when the cutoff value for the predicted probability of having AD was 0.66.

**COMMENT**

In this study we found that 2 tests, one a measure of verbal recognition memory and the other a measure of verbal fluency, distinguished accurately between patients with AD and those with SIVD. A double dissociation was found. The patients with AD performed worse on a recognition memory test than the patients with SIVD, whereas the SIVD group performed worse on the COWAT. Receiver operating characteristic curves were used to identify the optimal cutoff points for each test. The combination of poor recognition and preserved verbal fluency was associated with an LR+ of 5.1 when the cutoff value of the predicted probability of having AD was 0.54. This would generate a moderate change in the pretest-to-posttest likelihood of AD.

The participants in the SIVD subgroup all had subcortical lacunes. However, because 12 of these 31 participants also had cortical infarcts, we replicated the analysis in the smaller group (n = 19) with only subcortical lacunes. In addition to the same 3 variables emerging from the forward logistic analyses, the β coefficients were of a similar magnitude and in the same direction, allowing us to derive the same conclusions from both analyses. However, to further strengthen our findings, cross-validation of the discriminatory variables should be performed in a completely independent sample. In such a study, the β coefficients would be applied to the confirmatory data set, and the predicted probabilities would be correlated with the actual outcomes.

The literature comparing AD with SIVD is limited. In many studies of VaD, patients are not identified by subtype, and group differences in dementia severity, age, and level of education are not adequately controlled. In our study we focused on SIVD, a more homogeneous subgroup of VaD. The 2 groups of patients were similar in age and dementia severity as measured by the MMSE. Education was used as a covariate in all analyses to statistically remove its effects on test scores. Finally, the study was designed to separate the assignment of the target reference groups from the neuropsychological tests, in keeping with evidence-based guidelines for diagnostic tests.²¹,²²
Our results are consistent with previous patterns noted in the literature. Patients with subcortical ischemic vascular disease tend to show greater impairment in executive function and relatively better preservation of recognition memory.23,24 The COWAT is a widely used test of executive function25 and is considered to measure impairment in the dorsolateral prefrontal system, which is interconnected with the basal ganglia and thalamus in a frontal subcortical loop. The specific nuclei include the dorsolateral caudate nucleus, lateral dorsomedial globus pallidus internus, and anterior and dorsomedial nuclei of the thalamus.5,26 Our finding that letter fluency was one of the most critical tests in the battery may indicate that this circuit is frequently disrupted in SIVD at 1 or more of these subcortical loci or in the white matter tracts that interconnect them with the dorsolateral prefrontal lobe. Thus, the results of our study are consistent with earlier findings27,28 that the deficits found in patients with SIVD are primarily due to changes in subcortical structures and in the white matter of the frontal lobe.

It is not surprising that the patients with AD in this study were characterized by deficits in recognition memory. This pattern is well established in the AD literature. Poor recognition memory is considered to reflect deficits in storage caused by deficient consolidation of new memory traces.29 This, in turn, is most often linked to the degeneration of the mesial temporal areas, namely the hippocampus and amygdala. These areas are highly vulnerable to neurofibrillary degeneration, one of the classic neuropathological hallmarks of AD.30,31 Thus, poor recognition memory among patients with AD is likely due to neurofibrillary degeneration in the mesial temporal lobes, including the hippocampus.

This study focuses on a highly selected group of subjects. All patients met either NINCDS-ADRDA criteria for probable AD or NINDS-AIREN criteria for probable VaD modified for SIVD. The latter requires a history of a temporal relationship between a clinical stroke and cognitive decline. Among various available clinical criteria for VaD (Hachinski Ischemic Score, California Alzheimer Disease Diagnostic and Treatment Center criteria, DSM-III-R, and International Classification of Diseases, Ninth Revision), the NINDS-AIREN criteria are the most conservative.32-34 In addition, attempts were made to exclude subjects with mixed AD and VaD. This was done by restricting anyone with vascular lesions on a CT scan from the AD group and excluding stroke subjects without a clearly defined temporal relationship to the onset of dementia from the SIVD group. This selectivity limits the generalizability of our results to typical clinical practice settings where mixed, possible, and other subtypes of VaD are likely to be seen.

How can the results of this study be used to assess the relative likelihood of AD vs SIVD in a clinical setting? After administering the 2 neuropsychological tests, the regression coefficients provided in Table 2 can be used to calculate the probability of having AD. For example, patient A had a grade 12 education, a RAVLT recognition memory score of 60%, and a mean COWAT score of 13 for the 3 trials. The linear combination [X] of the products of the regression coefficients using this patient’s values for educational level, RAVLT recognition, and COWAT would be as follows:

\[
X = (1.05) + (0.23)(12) + (-0.11)(60) \\
+ (0.55)(13) = 4.36.
\]

Using the equation

\[
e^X/(1 + e^X),
\]

the predicted probability of AD for patient A would be 0.99.

We would like to caution physicians and neuropsychologists in the appropriate application of this procedure to their patients. First, it should be applied only to those patients who have undergone a thorough clinical workup (apart from a CT scan) to exclude other causes of cognitive impairment and for whom there are only 2 reasonable alternative diagnoses: AD or SIVD. A second caution in the application of these results is that this procedure should be applied only to patients who are similar to the study sample (ie, similar age and educational level, mild to moderate impairment, fluency in English, and the ability to hear normal speech with or without auditory aids). Finally, these tests should be conducted in a controlled, quiet environment by someone trained in their administration and who is skillful with the patient population. Adherence to these 3 recommendations is critical for the proper application of the procedures described in this article.

In conclusion, we illustrate how this type of study may be useful in deciding whether to order a neuroimaging study. In epidemiological studies35 or in managed care settings, neuroimaging may be considered an unnecessary expense.36 It would be useful if neuropsychological test results could identify patients for whom the addition of a neuroimaging study would significantly change the diagnosis or treatment.37 The utility of a diagnostic test (eg, a CT scan) depends not only on its sensitivity and specificity but also on the prior probability of disease. A test is less useful if the prior probability is either very low or very high. The results of neuropsychological testing may help to inform ascertainment of this prior probability.

For example, if a patient obtains a low predicted probability of AD after the regression coefficients are applied to his or her test scores and educational level, the patient is likely to have SIVD, and a CT scan should be ordered. If a patient obtains a high predicted probability

Table 2. Probability of Alzheimer Disease: Results of Logistic Linear Regression Analyses of the Relationship Between Diagnostic Classification and Variables in the Reduced Model*  

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regression Coefficient</th>
<th>P</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.05</td>
<td>.67</td>
<td>.35</td>
<td>.50-2.24</td>
</tr>
<tr>
<td>Level of education</td>
<td>0.23</td>
<td>.04</td>
<td>1.26</td>
<td>1.01-1.57</td>
</tr>
<tr>
<td>RAVLT, recognition</td>
<td>-0.11</td>
<td>.002</td>
<td>0.89</td>
<td>0.83-0.96</td>
</tr>
<tr>
<td>COWAT</td>
<td>0.55</td>
<td>.001</td>
<td>1.74</td>
<td>1.28-2.37</td>
</tr>
</tbody>
</table>

*RAVLT indicates Rey Auditory Verbal Learning Test; COWAT, Controlled Oral Word Association Test; and ellipses, not applicable.
of AD, the decision to order a CT scan will depend on the magnitude of the probability as well as the prevalence of the disease in the clinic (pretest probability). If the pretest probability of AD is high, one might decide to forgo the CT scan. If it is low, the test results will have less of an effect on the posttest probability, and the physician may still decide to order a CT scan. Future study is required to evaluate (1) the utility of using neuropsychological test performance to inform the physician about the prior probability of a disease; and (2) if more accurate assessment of prior probability will provide a cost-effective basis for deciding whether to order a neuroimaging study.

Accepted for publication July 18, 2001.

This research was supported by the Ontario Ministry of Health, Toronto, and by grant P01-AG 12435 from the National Institute on Aging, Bethesda, Md.

Corresponding author and reprints: Mary C. Tierney, Geriatric Research, A145, Sunnybrook and Women's Health Sciences Center, 2075 Bayview Ave, Toronto, Ontario, Canada M4N 3M5 (e-mail: mary.tierney@swchsc.on.ca).

REFERENCES


22. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature, III: how to use an article about a diagnostic test, B: what are the results and will they help me in caring for my patients? JAMA. 1994;271:703-707.


