Do Long Tests Yield a More Accurate Diagnosis of Dementia Than Short Tests?

A Comparison of 5 Neuropsychological Tests

Donald T. Stuss, PhD; Nachshon Meiran, PhD; D. Antonio Guzman, MD; Ginette Lafleche, PhD; Jonathan Willmer, MD

Objective: To provide comparative evidence for a valid and practical measure of mental-status functioning that could be used in dementia clinics.

Design: Five mental-status neuropsychological tools for dementia screening were administered to patients in a memory disorder clinic. These included the Mini-Mental State Examination, the Dementia Rating Scale, the 6-item derivative of the Orientation-Memory-Concentration Test, a short Mental Status Questionnaire, and a composite tool we labeled the Ottawa Mental Status Examination, which assessed orientation, memory, attention, language, and visual-constructive functioning. The tools were compared using various criteria, including the statistical factors of sensitivity and reliability; effects of gender, native language, and language of testing; the utility of these tests for the differential diagnosis of Alzheimer-type and vascular dementia; and sensitivity to cognitive decline in the entire sample and among patients with severe dementia.

Results: All of the tests were highly intercorrelated, suggesting that they are interchangeable.

Conclusion: The comparisons along the various criteria indicate that if the objective is to have a general index of dementia of the Alzheimer type, short tests are at least as good and sometimes better than the longer tests.

Arch Neurol. 1996;53:1033-1039

From the Rotman Research Institute of Baycrest Centre for Geriatric Care (Drs Stuss and Meiran) and the Departments of Medicine (Neurology) and Psychology (Dr Stuss), University of Toronto, Ontario; the Department of Medicine (Neurology), University of Ottawa, Ontario (Drs Stuss, Guzman, and Willmer); the Memory Disorder Clinic, Elizabeth Bruyere Health Centre, Ottawa (Drs Guzman and Willmer); and the Department of Neuropsychology, New England Sinai Hospital and Rehabilitation Center, Stoughton, Mass (Dr Lafleche).

DIAGNOSIS OF dementia of the Alzheimer type (DAT) requires examination of the cognitive status of the person. The question for busy clinicians is which test of mental status provides the best results in sensitivity, accuracy, and time. We investigated this question in a prospective study in which we compared 5 mental-status tests.

Neuropsychological tools are used to detect deterioration in overall cognitive functioning and are at the heart of the early diagnosis of dementia.[1,2] In general, the various tests have reasonable sensitivity and specificity.[3] The choice among the many different mental-status tests[4] becomes arbitrary unless the tests are formally compared. Relatively few studies compared 2 or more mental-status tests.[5-17] What is still lacking is a large-scale study in which different mental-status tests are compared along many criteria. We report a study conducted on the database collected in the Memory Disorder Clinic of the Ottawa (Ontario) General Hospital. The study was planned when the clinic was established. As part of their clinical assessment, the patients were administered 5 mental-status tests. Three of the tests are among the most widely used: Mini-Mental State Examination (MMSE),[18] Dementia Rating Scale (DRS)[19] and the abbreviated 6-item Orientation-Memory-Concentration (OMC) Test adapted from Blessed et al[20] by Katzman et al.[21] The 2 additional tests were a 10-item Mental Status Questionnaire (MSQ),[22] and the Ottawa Mental Status Examination (OMSE), a composite measure of the 3 shorter examinations plus additional measures, designed to see if more items would enhance reliability. Three of the tests were short—MMSE, OMC, and MSQ—and 2 were minibatteries—DRS and OMSE.

Our research objective was to compare the 5 tests using many criteria, including sensitivity, specificity, reliability, sensitivity to cognitive decline, and utility for differential diagnosis. Our approach had several advantages. First, all

See Patients and Methods on next page

PATIENTS AND METHODS

PATIENTS

All data were collected from patients who were referred to the Memory Disorder Clinic of the Ottawa General Hospital. Patients returned to the clinic for follow-up approximately every 12 to 16 months until diagnosis was confirmed or until clinical help could no longer be offered, or both. The initial diagnosis was made independently by 2 neurologists (D.A.G., J.W.) using the results of their examinations, including a neurobehavioral investigation, because the prospective purpose was to assess the sensitivity and specificity of the neuropsychological tests. The neuropsychological test results were reviewed clinically to see if discrepancies occurred, but the quantitative measures were not used by the neurologist. All patients received a full workup with repeated neurological, neurobehavioral, and laboratory tests. The final diagnosis was determined at a consensus meeting by the neurologists (D.A.G., J.W.), neuropsychologists (D.T.S., G.L.), and radiologists who had tested the patients, using the results of the repeated testing. Four types of dementia were considered in the analysis: probable DAT, possible DAT, vascular dementia (VaD), and a group of mixed types. Diagnosis of probable and possible DAT was based on National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria.[2] The 2 most common types were DAT and VaD. A fifth group included about a fourth of the patients referred for memory problems and suspected DAT who were eventually found not to have dementia.

DIAGNOSTIC CRITERIA

Probable and Possible DAT

Patients were diagnosed as having probable or possible DAT based on the NINCDS-ADRDA criteria.[2]

Vascular Dementia

Patients were diagnosed as having VaD if they met the following criteria: Hachinski[23] score more than 7, focal neurological findings, evidence of stroke on computed tomography scan, acute onset and stepwise progression of disease, and a history of hypertension or of cardiac or vascular disease.

A group of mixed dementia types included patients who were diagnosed as having DAT and another dementia, or DAT and depression. This group also included patients who were diagnosed as having other dementia, including Pick disease, other frontal dementias, Parkinson disease, progressive supranuclear palsy, spinocerebellar degeneration, multisystem atrophy, hydrocephalus, metabolic and toxic disorders, trauma, neoplasm, and demyelinating disorders.

Among the patients in the final list, 283 patients (of whom 238 were tested in their native language of English or French) were tested at least twice; of these, 77 were tested 3 times (63 in their native language). Of those patients, 17 were tested 4 times (14 in their native language) and of the 17, 3 patients were tested 5 times (all in their native language). The average intertest delay was 14.6 months. The remaining patients were tested only once.

The aforementioned mental-status tests were used. They were given in the standard format, with the exception that Canadian content replaced American content (eg, "What is the name of the prime minister of Canada?").

tests were collected from the same set of patients, providing the basis for direct comparison. Second, patients in the clinic were seen successively at clinic follow-up, allowing confirmation of diagnoses. Third, our cohort included patients who had been referred

See also page 1043

for possible dementia, but who on successive evaluations turned out not to have dementia. In our view, this group constitutes a better control group than would persons selected randomly from the community, because the function of a memory disorder clinic is to discriminate among those referred to it for possible dementia from those who do not have dementia. The "cost" of this feature is that generalization of sensitivity and specificity to community-based studies is unwarranted.

RESULTS

Unless stated otherwise, the analyses were performed on patients whose native language was French or English and who were tested in their native language.

ARE THE TESTS INTERCHANGEABLE?

To assess if the tests measure similar abilities, we conducted a principal component analysis on the 5 tests. This analysis was not conducted to assess underlying abilities (in which case tests that include overlapping questions should not be analyzed together). All of the tests loaded heavily on 1 component, which explained 87.7% of the variance. The proportions of variance accounted for by the single component were 0.80 for DRS, 0.96 for OMSE, 0.87 for MSQ, 0.83 for OMC, and 0.93 for MMSE. The different tests measure similar abilities and are, therefore, interchangeable. Consequently, any consideration of preferring one test over others should be related to other factors such as administration time and psychometric properties.

RETEST RELIABILITIES

Retest reliabilities were computed by comparing the scores that patients obtained in their first and second tests. Reliabilities were computed twice. The first was completed on the data of patients who were diagnosed as not having dementia, using a semipartial correlation "partialling out" the effect of intertest delay from the second

test. The advantage of this analysis is that patients who do not have dementia should not show any change in cognitive functioning between tests. The second computation of reliability was based on the entire sample, this time partialling out the effects of delay and type of dementia (DAT pooled across probable and possible, VaD, other dementia, and no dementia) from the second test.

Reliability results for both analyses are given in **Table 1**. For both computations, DRS, which is the longest test, had the lowest overall reliability. In general, MMSE had the most satisfactory reliability. Subscale scores should be interpreted with caution, if at all.

Table 1. Retest Reliabilities for the 5 Screening Tests*

Test	Proportions of Variance: Nondemented Group	Proportions of Variance: All Groups
MMSE	.85 (47)	.98 (236)
OMC	.77 (47)	.80 (235)
MSQ	.87 (47)	.90 (236)
DRS	.76 (35)	.79 (130)
Attention	.17	.68
Initiation	.71	.70
Construction	.97	.64
Conceptualization	.41	.49
Memory	.82	.72
OMSE	.88 (47)	.80 (236)
Orientation	.71	.83
Memory	.62	.71
Attention	.84	.82
Language	.45	.69
Copy	.47	.64

**The estimates for the nondemented group were made after "partialling out" the linear effect of retest delay. The estimates for the entire sample were made after partialling out retest delay and cause (expressed as 3 dummy variables) (β weights predict second test). Number of patients are given parenthetically for 5 screening tests. MMSE indicates Mini-Mental State Examination; OMC, Orientation-Memory-Concentration Test; MSQ, Mental Status Questionnaire; DRS, Dementia Rating Scale; and OMSE, Ottawa Mental Status Examination.*

INFLUENCE OF ETIOLOGY, EDUCATION, AGE, GENDER, AND TESTING LANGUAGE

An ideal test for the diagnosis of dementia, especially for the differential diagnosis of DAT, should be maximally sensitive to etiology (type of dementia) and insensitive to gender, language, and education. We estimated the effects on test scores of gender, testing language (English speakers tested in English, French speakers tested in French; and French or other non-English speakers tested in English), education (up to 8 years, 9-12 years, undergraduate level, graduate level), and etiology (DAT, VaD, other dementias, and no dementia). Age also was included as a category variable after initial examination indicated that, although the effect of age on test score is monotonic, it is nonlinear. Given the uneven distribution of patients across these categories, we could not include interactions in the models and restricted them to include main effects only. In every case, we report the effect of a variable only after the effects of the remaining variables were controlled for, as though it entered last into the model.

Follow-up pairwise comparisons were conducted on least-square means, which are corrected for the effects of the remaining independent factors. The results indicated a large main effect of dementia type for all tests: MMSE, $F(3,598)=67.02$; OMC, $F(3,594)=77.80$; MSQ, $F(3,595)=64.96$; DRS, $F(3,513)=52.24$; and OMSE, $F(3,596)=70.25$ (all $P<.001$). These tests can differentiate among the defined types of dementia, at least at a group level. Follow-up pairwise tests indicated that patients with DAT performed worst, patients with VaD and other dementia performed at an intermediate level, and patients without dementia performed best. The effect of education also was significant in all analyses, but it was considerably smaller than the effect of etiology: MMSE, $F(3,598)=4.97$, $P<.005$; OMC, $F(3,594)=6.57$, $P<.001$; MSQ, $F(3,595)=3.56$, $P<.05$; DRS, $F(3,513)=21.45$, $P<.001$; OMSE, $F(3,596)=5.56$, $P<.001$. In every case, the pairwise comparisons indicated that the source of variation was a lower performance among patients with 8 years of education or less compared with the other groups. For the other levels of education (high school and above), no systematic difference was seen. Age was a significant main effect for all tests except the OMSE: MMSE, $F(4,598)=2.54$, $P<.05$; OMC, $F(4,594)=3.39$, $P<.01$; MSQ, $F(4,595)=3.87$, $P<.005$, and DRS, $F(4,513)=4.95$, $P<.001$. Despite being significant, the effect of age was even smaller than the effect of education. The age pattern obtained, with older subjects performing worse, was similar to that reported by Crum et al[24] and Schmidt et al.[25] Gender was a significant factor only for 2 tests: MSQ, $F(1,595)=10.32$, $P<.005$; and OMSE, $F(1,596)=4.81$, $P<.05$; indicating that the remaining tests are preferable if gender effects are an issue. In both cases, men obtained better scores than did women. Results of language of testing were not statistically significant. In summary, the type of the dementia (at least for the general classifications considered) affected cognitive functioning as measured by all of the tests used. Age and education had similar results, albeit small. Gender effects were test-specific.

We wanted to determine if correcting for the effects of language, education, age, and sex would noticeably increase the diagnostic value of the tests. We compared the percent variance in test scores accounted for by etiology (Eta^2) before and after the effects of the remaining variables were statistically controlled. Eta^2 scores were 0.127, MMSE; 0.153, OMC; 0.125, MSQ; 0.126, DRS; and 0.131, OMSE before controlling for the effects of age, sex, and other variables. These values increased to 0.132, MMSE; 0.161, OMC; 0.133, MSQ; 0.139, DRS; and 0.138, OMSE when these variables were controlled. The increases in explained variance due to the inclusion of sex, age, education, and testing language were negligible, less than 1% in 4 tests and 1.3% for DRS. These results also can be used to assess the signal-to-noise ratio, where signal refers to etiology-related variance, and noise refers to variance that is not etiology-related. Signal-to-noise ratios are similar across tests, with OMC (a short test) showing a slight advantage over the other tests.

Another way of assessing the usefulness of correction was to examine the maximum diagnostic value of

Table 2. Sensitivity, Specificity, and Diagnostic Values at Recommended Cutoff Points*

Test	Range†	Cutoff‡	Sensitivity	Specificity	Diagnosticity
		Probable Dementia of the Alzheimer Type			
MMSE	23-24	24/25	.85	.90	.89
OMC	16	16/15	.74	.93	.87
MSQ	7	7/8	.84	.89	.87
DRS	116-117	117/118	.76	.91	.87
Attention	33	33/34	.28	.94	.72
Initiation	25-26	26/27	.53	.94	.81
Construction	5	5/6	.51	.90	.77
Conceptualization	32-33	33/34	.87	.80	.83
Memory	16-17	17/18	.89	.85	.86
OMSE	25-30	30/31	.77	.92	.87
Orientation	8	8/9	.70	.99	.88
Memory	4-5	5/6	.71	.92	.84
Attention	4-5	5/6	.68	.85	.78
Language	6	6/7	.48	.90	.74
Copy	0	0/1	.49	.94	.77
		Dementia in General			
MMSE	28	28/29	.93	.58	.84
OMC	7-8	7/6	.91	.63	.83
MSQ	8	8/9	.81	.78	.80
DRS	131-135	135/136	.95	.55	.83
OMSE	36	36/37	.92	.62	.84

*Test names are given in Table 1.
†Indicates range of scores where diagnostic value was maximal.
‡Given as the best score for demented group/worst score for nondemented group.

the test before and after correction.[26] For the diagnosis of dementia in general (based on a comparison of all patients who had dementia with patients who did not have dementia), the correction led to a slight decrease (rather than improvement) in the maximal diagnostic value in all tests but OMC, in which it remained unchanged. For the detection of probable DAT, the correction did not affect the maximal diagnostic value of MMSE, OMC, and MSQ, and improved it slightly for DRS (0.87-0.88) and OMSE (0.87-0.89). Based on these results, the correction formulas we used are not recommended, at least with patient populations similar to this one.[27] Sufficient diagnostic accuracy is provided by the test results alone.

SENSITIVITY, SPECIFICITY, AND DIAGNOSTIC VALUE

Cutoff points, while not recommended,[28] may be useful in some circumstances. In determining recommended cutoff scores, we based our decision on 3 variables: sensitivity, specificity, and the diagnostic value.[26] Sensitivity is the proportion of independently identified patients with dementia who are diagnosed as having dementia by the test. Specificity is the proportion of patients without dementia who are diagnosed as not having dementia. The diagnostic value of the test is the proportion of correctly diagnosed patients, with and without dementia, among all of the patients. The cutoffs were positioned in the range of scores where the diagnostic value of the test was maximal and, within that range, the cutoff score was where sensitivity was maximal. Sensitivity is more important than specificity (as long as the overall diagnostic value remains high), because patients who are suspected of having dementia are likely to be tested more thoroughly, so false positives will have an additional chance of being detected. The cutoff points recommended by our present results are given in **Table 2**. We also compared the results using previously recommended cutoffs applied to the present data. These are given in **Table 3**.

All tests reach a diagnostic value of probable DAT of 0.87 to 0.89. When specificity is about 0.90, the tests that show the highest sensitivity to DAT are MMSE (0.85) and MSQ (0.84), while the remaining tests have sensitivities of about 0.75. Not surprisingly, the diagnostic value of the tests is lower when patients without dementia are compared with the total group of patients with dementia, rather than with patients with probable DAT. When it is maximal, sensitivity is high (0.90 or above) for all tests except for MSQ, but specificity is low (0.55-0.63) for all tests except for MSQ, for which it is higher. These results suggest that MSQ is the best test as far as this comparison of cutoff scores informs us, because it reaches maximal diagnostic value at a point where sensitivity and specificity are still reasonable. Two DRS subscales (conceptualization and memory) and 2 subscales of OMSE (orientation and memory) reach acceptable levels of sensitivity, specificity, and diagnostic value of DAT.

DIFFERENTIAL DIAGNOSIS

For the diagnosis of dementia in general, a single score from a short screening test seems to be more reliable, sensitive, and specific than from long tests and may be sufficient for differential diagnosis of dementia. A long test that is composed of subscales, however, may be superior for differential diagnosis if the 2 patient groups showed different profiles of deficiencies. To assess for this possibility, we conducted our analysis in 2 steps. In the

Table 3. Sensitivity, Specificity, and Diagnostic Values for Previously Recommended Cutoff Points*

	Dementia of the Alzheimer Type				Dementia in General		
Test	Cutoff†	Sensitivity	Specificity	Diagnosticity	Sensitivity	Specificity	Diagnosticity
MMSE	24/25	.85	.90	.89	.70	.90	.75
OMC	10/9	.95	.77	.83	.83	.77	.82
DRS	122/123	.87	.84	.85	.72	.84	.76
Attention	31/32	.17	.97	.71	...	...	...
Initiation	28/29	.55	.87	.76	...	...	...
Construction	3/4	.15	.96	.69	...	...	...
Conceptualization	31/32	.77	.85	.82	...	...	...
Memory	18/19	.91	.81	.85	...	...	...

*Abbreviations are given in Table 1.
†Defined in Table 2.

Table 4. Differences Between Probable DAT and Vascular Dementia on the Subscales of DRS and OMSE*

	Probable DAT		Vascular Dementia				
Subscale (Test)	Mean	SD	Mean	SD	*t*†	*df*	*P*
Conceptualization (DRS)	25.2	8.7	29.6	7.2	2.73	97	<.01
Memory (DRS)	11.5	5.3	15.1	5.3	3.38	97	<.01
Orientation (OMSE)	6.8	4.0	9.4	3.7	3.76	120	<.001
Memory (OMSE)	3.9	2.2	5.6	1.8	4.42	119	<.001
Attention (OMSE)	3.1	2.8	4.9	2.2	3.92	119	<.001
Language (OMSE)	5.8	2.2	6.7	1.4	2.49	120	<.05

*Test names are given in Table 1. DAT indicates dementia of the Alzheimer type.
†All tests are 2-sided.

first step, we compared the patient groups on every subscale individually. At the next phase, we tried to identify a weighted combination of the subscale scores that best differentiated between the 2 groups. We compared the patients diagnosed as having probable DAT with those diagnosed as having VaD, using the subscales of DRS and OMSE and the total test scores of all 5 tests. Forty-seven patients with probable DAT and 52 patients with VaD had information on DRS subscales, and 63 patients with probable DAT and 59 patients with VaD completed the OMSE subscales. The significant subscales differences are given in **Table 4**. Two DRS subscales, conceptualization and memory, differentiated between the defined dementia types. Four subscales of OMSE differentiated between the types of dementia: orientation, memory, attention, and language. Similar comparisons between patients with VaD and without dementia showed significant differences on all subscales. Stepwise discriminant analysis was conducted to identify a weighted combination of the subscales that best discriminated between the groups. Only memory entered as a predictor when DRS scales were used, whereas both memory and attention predicted dementia type when the subscales of OMSE were analyzed (criterion for entry, $\alpha = .15$). The results for DRS show that having subscales confers no advantage for differential diagnosis, because the optimal weighted combination comprised only 1 subscale, which also differentiates between patients with and without dementia. The same conclusion is correct for OMSE, because the 2 subscales that constitute the weighted combination also differentiate patients with dementia from patients without dementia. Furthermore, the rate of classification errors

Table 5. Results From Discriminant Analysis Comparing Patients With Probable DAT and Patients With Vascular Dementia*

	% of Patients	
Tests	DAT Classified as VaD	VaD Classified as DAT
MMSE	40	33
OMC	39	30
MSQ	42	39
DRS	64	18
DRS subscales	38	29
OMSE	39	32
OMSE subscales	37	32

*Test names are given in Table 1. DAT indicates dementia of the Alzheimer type; VaD, vascular dementia.

(**Table 5**) is roughly the same for the total score of OMSE as it is for the weighted combination.

We then compared the groups on the tests that normally yield only 1 score, such as MMSE (Table 5). These tests can differentiate between the 2 groups just as efficiently. We also conducted a series of discriminant analyses in which the total score on a given test was entered as a predictor of patient group (probable DAT vs VaD) or the subscales that were identified in the stepwise discriminant analysis. We drew 3 conclusions from these analyses. First, the number of misclassified patients makes the usefulness of screening tests for differential diagnosis questionable. Second, the number of misclassified pa-

Table 6. Estimated Annual Rate of Cognitive Decline Based on the Second and Third Sessions*

Dementia	First to Second Sessions				First to Third Sessions			
	No. of Patients	Mean Score	SD	P	No. of Patients	Mean Score	SD	P
				MMSE				
DAT	84	3.4	5.6	<.001	23	2.8	2.0	<.001
VaD	36	1.7	3.4	.007	9	1.8	2.3	.03
Nondemented	47	−0.4	1.5	.07	16	0.4	0.8	.04
				OMC				
DAT	84	3.4	4.8	<.001	23	3.4	2.1	<.001
VaD	36	2.6	6.1	.01	8	1.9	1.4	.007
Nondemented	47	3.5	21.6	.27	16	0.4	1.1	.02
				MSQ				
DAT	84	1.3	1.8	<.001	23	1.5	0.8	<.001
VaD	36	0.8	1.6	.004	9	0.7	0.7	.02
Nondemented	47	−0.9	5.4	.25	16	0.3	0.5	.02
				DRS				
DAT	35	8.4	9.2	<.001	12	6.7	3.7	<.001
VaD	25	3.5	12.6	.18	6	4.2	5.6	.12
Nondemented	35	−4.2	25.3	.33	14	1.4	4.0	.21
				OMSE				
DAT	84	4.8	6.0	<.001	23	4.3	2.3	<.001
VaD	36	2.3	4.6	.005	9	2.3	2.7	.04
Nondemented	47	−1.9	10.8	.23	16	0.8	1.1	.02

*Test names are given in Table 1. DAT indicates dementia of the Alzheimer type; VaD, vascular dementia.

tients is roughly the same regardless of which test is used. Finally, using subscales increases the discrimination only for DRS. The DRS total score provides poor discrimination, but the subscales yield as good a discrimination as achieved by the total scores of the other tests. In summary, the analyses did not indicate that longer tests are more useful than short tests in making a differential diagnosis between the types of dementia compared. In fact, none of the tests should be used for that purpose. Stated positively, the tests compared in this study are useful for the general diagnosis of dementia, and not for the differential diagnosis among types of dementia, at least between DAT and VaD.

SENSITIVITY TO COGNITIVE DECLINE

Another possible advantage of long screening tests over short ones is that they may include easy items, allowing for the detection of cognitive decline among patients with severe dementia.[15] We evaluated whether the tests detected cognitive decline, comparing the difference between the first and second tests and between the first and third tests. These differences were expressed as annual rates of decline to equate patients for intertest delay (**Table 6**).

All of the tests detected cognitive decline among patients with DAT (pooled across probable and possible DAT to increase sample size), and all but one test (DRS) detected notable decline among patients with VaD. The discrepancy for DRS may reflect the small number of subjects rather than a genuine sensitivity difference. None of the tests detected a significant decline from the first to second tests among patients without dementia. The results for the difference between first and third tests are noisier because of the smaller sample size. Nevertheless, the decline in cognitive functioning was largest for patients with DAT and smallest for patients without dementia. The patients without dementia in our sample were elderly, and the small decline in cognitive functioning observed between the first and third testing sessions may be attributed to deterioration that is associated with normal aging.

COMMENT

Our study can be used as a guideline for the selection of a specific dementia screening test for use in a general dementia clinic. The 5 tests were compared using various criteria. First, the tests are interchangeable, because all of them measure similar attributes or processes, as indicated in the principal component analysis. With respect to reliability, MMSE was found to be the most reliable test and DRS was the least reliable test. About half of the subscales of DRS and OMSE did not reach an acceptable level of reliability and are therefore not recommended for use as separate indices of dementia. None of the tests was sensitive to language (French or English) or language of testing (whether the patients were tested in their native language). The OMSE and MSQ and some subsets of DRS and OMSE, were sensitive to gender differences. As for signal-to-noise ratio (to what degree test scores are related to type of dementia as opposed to premorbid differences), the tests were similar to one another. When specificity is about 0.90, MMSE and MSQ have a sensitivity to diagnose DAT at about 0.85, while the remaining tests were about 0.75. None of the tests proved useful in the differential diagnosis of probable DAT and VaD. All of the tests were similarly sensitive to cog-

nitive decline, including detecting decline among patients with severe dementia.

Our data should be interpreted in the context of our population. The effect of education was significant. Our clinical experience suggests that the sensitivity of the screening tests early in a dementing process in highly educated and intelligent people remains to be validated. To evaluate this, longitudinal testing of "normal" persons, perhaps starting as early as age 55 or 60 years, may be required.

To summarize, we could not identify a single advantage of long tests over short tests. Given the limited time allowed for diagnosis, if the goal is to detect dementia as a general diagnosis, short tests should be preferred over long tests. This is because the added length does not result in higher reliability or more detailed information. Since, of the short tests, MSQ was shown to be gender-sensitive, either of the 2 other short tests is recommended—MMSE or OMC.[8] Longer tests may be helpful in differentiating specific subtypes of dementia, especially if they are designed to assess specific characteristics. Methods for the interpretation of test scores and formulas for translating a score in one test into its equivalent on another test are discussed in another article.[28]

Accepted for publication May 10, 1996.

This project was possible by funding from a Rotman Research Fellowship, Toronto (Dr Meiran); research grants from the Alzheimer Society of Canada, the Ontario Mental Health Foundation, and the Medical Research Council of Canada (Dr Stuss); the Alzheimer Society of Ottawa—Carleton (infrastructure funds for the Memory Disorder Clinic at the Ottawa General Hospital); and the Ottawa General Hospital.

We are grateful to the patients of the clinic and their families, to the many staff members of the clinic who were necessary to the clinic's success in attracting patients, and to the neuropsychology interns and technicians who helped to test the patients.

Reprints: Donald T. Stuss, PhD, Rotman Research Institute, Baycrest Centre for Geriatric Care, 3560 Bathurst St, Toronto, Ontario, Canada M6A 2E1.

REFERENCES

1. Corey-Bloom J, Thal LJ, Galasko D, et al. Diagnosis and evaluation of dementia. *Neurology.* 1995;45:211-218.
2. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan M. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology.* 1984;34:939-944.
3. Siu AL. Screening for dementia and investigating its causes. *Ann Intern Med.* 1991;115:122-132.
4. Ritchie K. The screening of cognitive impairment in the elderly: a critical review of current methods. *J Clin Epidemiol.* 1988;41:635-643.
5. Bobholz JH, Brandt J. Assessment of cognitive impairment: relationship of the Dementia Rating Scale to the Mini-Mental State Examination. *J Geriatr Psychiatry Neurol.* 1993;6:210-213.
6. Clarke M, Jagger C, Anderson J, Battock T, Kelly F, Campbell Stern M. The prevalence of dementia in a total population: a comparison of two screening instruments. *Age Ageing.* 1991;20:396-403.
7. Fillenbaum GG. Comparison of two brief tests of organic brain impairment, the MSQ and the Short Portable MSQ. *J Am Geriatr Soc.* 1980;8:381-384.
8. Fillenbaum GG, Heyman A, Wilkinson WE, Haynes CS. Comparison of two screening tests in Alzheimer's disease: the correlation and reliability of the Mini Mental State Examination and the Modified Blessed Test. *Arch Neurol.* 1987;44:924-927.
9. Fuhrer R, Ritchie K. Screening for dementia: a comparison of two tests using receiver operating characteristic (ROC) analysis. *Int J Geriatr Psychiatry.* 1993;8:867-871.
10. Harper RG, Chacko RC, Kotic-Harper D, Kirby HB. Comparison of two cognitive screening measures for efficacy in differentiating dementia from depression in a geriatric inpatient population. *J Neuropsychiatry Clin Neurosci.* 1992;4:179-184.
11. Hooijer C, Dinkgreve M, Jonker C, Lindboom J. Short screening tests for dementia in the elderly population, I: a comparison between AMTS, MMSE, MSQ and SPMSQ. *Int J Geriatr Psychiatry.* 1992;7:559-571.
12. Jagger C, Clarke M, Anderson J. Screening for dementia: a comparison of two tests using receiver operating characteristic (ROC) analysis. *Int J Geriatr Psychiatry.* 1992;7:659-665.
13. Lomeo CVC. Alzheimer's disease: correlational analysis of three screening tests and three behavioral scales. *Acta Neurol Scand.* 1992;86:603-608.
14. Ritchie K, Fuhrer R. A comparative study of the performance of screening tests for senile dementia using receiver operating characteristics analysis. *J Clin Epidemiol.* 1992;45:627-637.
15. Salmon DP, Thal LJ, Butters N, Heindel WC. Longitudinal evaluation of dementia of the Alzheimer type: a comparison of three standardized mental-status examinations. *Neurology.* 1990;40:1225-1230.
16. Thal LJ, Grundman M, Golden R. Alzheimer's disease: a correlational analysis of the Blessed Information-Memory-Concentration Test and the Mini-Mental State Examination. *Neurology.* 1986;36:262-264.
17. Uhlman RF, Larson EB, Bucher DM. Correlations of Mini-Mental-State Examination and modified Dementia Rating Scale to measures of transitional health status in dementia. *J Gerontol.* 1987;42:33-36.
18. Folstein MF, Folstein SE, McHugh PR. 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189-198.
19. Mattis S. *Dementia Rating Scale.* Odessa, Fla: Psychological Assessment Resources; 1988.
20. Blessed G, Tomlinson B, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry.* 1968;114:797-781.
21. Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *Am J Psychiatry.* 1983;140:734-739.
22. Robertson D, Rockwood K, Stolee P. A short mental-status questionnaire. *Can J Aging.* 1982;1:16-20.
23. Hachinksi VC, Iliff LD, Zilhka E, et al. Cerebral blood flow in dementia. *Arch Neurol.* 1975;32:632-637.
24. Crum RM, Anthony JC, Bessett SS, Folstein MF. Population-based norms for the Mini Mental State Examination by age and education level. *JAMA.* 1993;269:2386-2391.
25. Schmidt R, Freidl W, Fazekas F, et al. The Mattis Dementia Rating Scale: normative data from 1001 healthy volunteers. *Neurology.* 1994;44:964-966.
26. Baldessarini RJ, Finkelstein S, Arana GW. The predictive power of diagnostic tests and the effect of prevalence of illness. *Arch Gen Psychiatry.* 1983;40:569-573.
27. Kittner SJ, White LR, Farmer ME, et al. Methodological issues in screening for dementia: the problem of education adjustment. *J Chronic Dis.* 1986;39:163-170.
28. Meiran N, Stuss DT, Guzman DA, Lafleche G, Willmer J. Diagnosis of dementia: methods for interpretation of scores of 5 neuropsychological tests. *Arch Neurol.* 1996;53:1043-1054.