Comparison of the Short Test of Mental Status and the Mini-Mental State Examination in Mild Cognitive Impairment

David F. Tang-Wai, MDCM; David S. Knopman, MD; Yonas E. Geda, MD; Steven D. Edland, PhD; Glenn E. Smith, PhD; Robert J. Ivnik, PhD; Eric G. Tangalos, MD; Bradley F. Boeve, MD; Ronald C. Petersen, PhD, MD

Background: The Mini-Mental State Examination (MMSE) is the most widely used brief screening measure of cognition, but it is not sensitive in detecting mild memory or other cognitive impairments. The Short Test of Mental Status (STMS) was specifically developed for use in dementia assessment and was intended to be more sensitive to problems of learning and mental agility that may be seen in mild cognitive impairment (MCI).

Objective: To compare the STMS and MMSE for detecting or predicting MCI.

Design: Comparison of STMS and MMSE scores at baseline among 4 groups of patients: 788 patients with stable normal cognition, 75 patients with normal cognition at baseline but who developed incident MCI or Alzheimer disease during follow-up, 129 patients with prevalent MCI at baseline, and 235 patients with prevalent mild Alzheimer disease. All patients and control subjects for this study were evaluated through the Mayo Alzheimer's Disease Patient Registry or the Mayo Clinic Alzheimer's Disease Research Center, Rochester, Minn, using a standardized diagnostic approach.

Results: The STMS was slightly more sensitive than the MMSE in discriminating between patients with stable normal cognition and patients with prevalent MCI. The STMS was superior to the MMSE in detecting deficits in cognition in individuals who had normal cognition at baseline but later developed incident MCI or Alzheimer disease.

Conclusions: Compared with the MMSE, the STMS was better able to document MCI and was more sensitive in detecting deficits in cognition in individuals who had normal cognition at baseline but later developed incident MCI or Alzheimer disease.

Arch Neurol. 2003;60:1777-1781

MILD COGNITIVE IMPAIRMENT (MCI) has taken on increasing clinical importance because it is a precursor of dementia. The Mini-Mental State Examination (MMSE) is currently the mainstay of bedside mental status examinations, but it has limitations, especially for detecting MCI.

The Short Test of Mental Status (STMS)6,7 (Table 1) was developed and validated as a screening bedside mental status test specifically for use in mild dementia. It covers a broad range of cognitive functions and uses a 4-word learning list with a delayed recall of approximately 3 minutes. The construction of the recall task in the STMS was intended to make it more sensitive to the problems of learning and recall in MCI and early dementia. In addition, the STMS includes test items that better assess abstract reasoning and mental agility than the MMSE.

Despite theoretical improvements, we questioned whether the STMS offered any advantages vs the MMSE. We retrospectively compared the STMS and the MMSE to determine if there was a clinically relevant difference between the 2 tests in the assessment of patients with MCI and dementia.

METHODS

SUBJECTS

Patients and control subjects for this study were recruited prospectively through the Mayo Alzheimer’s Disease Patient Registry and the Mayo Clinic Alzheimer’s Disease Research Center (Rochester, Minn) using a standardized protocol. Both projects were approved by the Mayo institutional review board. The patients were derived from 2 sources: community patients in Rochester and regional patients referred to the Mayo Clinic Alzheimer’s Disease Research Center. The community patients were recruited through the Mayo Clinic Division of Community Internal Medicine. Volunteers with and without cognitive complaints or disorders were recruited.
Each patient was evaluated by either a behavioral neurologist or a behavioral neurology fellow, who obtained a medical history from the patient and corroborating sources and performed a complete neurological examination including the STMS.8-12 The physician also completed the Clinical Dementia Rating Scale (CDR)13 based on both the mental status examination and an interview of the subject or informant. If cognitively impaired, the patient had a laboratory assessment and structural neuroimaging of the brain using either computed tomography or magnetic resonance imaging.

A consensus meeting was held weekly to review each patient’s examination results. The team consisted of nurses, a geriatrician (E.G.T.), 4 behavioral neurologists (Emre Kokmen, MD [deceased], B.F.B., D.S.K., and R.C.P.), and 2 neuropsychologists (G.E.S. and R.J.I.). Subjects were given a consensus evaluation based on all of the current information listed earlier and were evaluated as having normal cognition, MCI, or dementia according to the Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition.15

The diagnosis of MCI was made if the patient met the following criteria: memory complaint, normal activities of daily living, normal general cognitive function, abnormal memory for age, and no dementia.15 The diagnosis of probable Alzheimer disease (AD) was based on National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer’s Disease and Related Disorders Association criteria.16

**ANALYSIS**

We examined the differences between the MMSE and STMS obtained at the baseline visits. Because the comparison of the MMSE and STMS was not prospectively conceived, the original study design did not maintain strict independence between test scores and clinical diagnoses. The neurologists’ diagnoses used the STMS and the patient’s medical history and functional assessment, and the neuropsychologists used the MMSE scores and the full psychometric battery. To avoid circularity biases, we sought to isolate the diagnostic groupings from the baseline STMS and MMSE scores. We grouped subjects according to their status at follow-up. Subjects were considered to have stable normal cognition if they remained cognitively healthy during the course of follow-up and for a minimum of 2 subsequent annual evaluations. Subjects with incident MCI had normal cognition at the baseline evaluation but developed MCI or AD at a subsequent follow-up examination. By this strategy, the diagnostic groupings were based on future status; although we used the same instruments and procedures as at baseline, subjects were grouped according to clinical course. As an additional approach to avoid circularity, we performed analyses in which we grouped subjects according to their CDR score. Because the CDR included information obtained from the patient’s medical history, the global CDR score assigned to each patient contained additional information not derived from the mental status assessments.

We used t tests for bivariate comparisons. Areas under the receiver operating characteristic curves were compared using a modified Mann-Whitney U test statistic.17 All statistical analyses were performed using either SAS software (SAS Institute Inc, Cary, NC) or S-plus software (Mathsoft Engineering and Education Inc, Seattle, Wash).

**RESULTS**

The demographic features of the subjects grouped by categories are presented in Table 2. The study group consisted of 788 subjects with stable normal cognition, 75 subjects with normal cognition at baseline but who developed MCI (n=54) or AD (n=21) during the fol-

---

### Table 1. The Short Test of Mental Status

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Testing</th>
<th>Maximum Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientation</td>
<td>Name; address; current location (building); city; state; date (day); month; year</td>
<td>8</td>
</tr>
<tr>
<td>Attention</td>
<td>Digit span (present at 1 per second; record longest correct span) 2-9-6-8-3 5-7-1-9-4 6-1-5-9-3-6-2</td>
<td>7</td>
</tr>
<tr>
<td>Learning and immediate recall</td>
<td>Learn our unrelated words: apple, Mr Johnson, charity, tunnel. Record the number of trials for acquisition (maximum of 4 trials)</td>
<td>4</td>
</tr>
<tr>
<td>Calculation</td>
<td>$5 \times 13 = 65 - 7 = 58/2 = 29 \times 11 =$</td>
<td>4</td>
</tr>
<tr>
<td>Abstraction/ similarities</td>
<td>Similarities: orange/banana, dog/horse, table/bookcase.</td>
<td>3</td>
</tr>
<tr>
<td>Information</td>
<td>President; first president; number of weeks per year; define an island</td>
<td>4</td>
</tr>
<tr>
<td>Construction</td>
<td>Copy the Necker cube. Draw a clock face showing 11:10.</td>
<td>4</td>
</tr>
<tr>
<td>Recall</td>
<td>The 4 words apple, Mr Johnson, charity, tunnel.</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total Score</strong></td>
<td></td>
<td>38</td>
</tr>
</tbody>
</table>

*Total score = sum of the subtest scores – (number of trials for acquisition – 1). For example, if a patient learned all 4 words on the first trial, nothing is subtracted from the sum of the subtest scores. If a patient required 4 trials to learn the 4 words, then 3 was subtracted from the sum of the subtest score.

### Table 2. Demographics of the Sample

<table>
<thead>
<tr>
<th></th>
<th>Stable Normal Cognition</th>
<th>Incident MCI or AD</th>
<th>Prevalent MCI</th>
<th>AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>788</td>
<td>75</td>
<td>129</td>
<td>235</td>
</tr>
<tr>
<td>Women/men</td>
<td>498/290</td>
<td>53/22</td>
<td>80/49</td>
<td>170/65</td>
</tr>
<tr>
<td>Age at baseline, y</td>
<td>78.2 ± 6.9</td>
<td>81.3 ± 11.5</td>
<td>79.5 ± 7.2</td>
<td>80.9 ± 7.7</td>
</tr>
<tr>
<td>Education, y</td>
<td>13.2 ± 2.9</td>
<td>13.5 ± 3.4</td>
<td>13.3 ± 3.2</td>
<td>12.0 ± 5.1</td>
</tr>
<tr>
<td>Baseline MMSE total score</td>
<td>28.2 ± 1.6</td>
<td>28.0 ± 1.6</td>
<td>26.3 ± 2.2</td>
<td>21.3 ± 4.4</td>
</tr>
<tr>
<td>Baseline STMS total score</td>
<td>34.2 ± 2.4</td>
<td>32.5 ± 3.2</td>
<td>30.6 ± 3.1</td>
<td>23.2 ± 5.7</td>
</tr>
</tbody>
</table>

Abbreviations: AD, Alzheimer disease; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; STMS, Short Test of Mental Status.

*AD defined as a Clinical Dementia Rating Scale score of 0.5 at baseline. Values are expressed as mean ± SD unless otherwise indicated.

†The MMSE maximum score = 30; STMS maximum score = 38.
of trade-offs between sensitivity and specificity across a larger range of scores. Sensitivity and specificity did not exceed 80% for either test or any test value.

Finally, we compared the baseline STMS and MMSE scores in subjects with stable normal cognition with those of individuals with normal cognition who developed incident MCI or AD. The mean ± SD follow-up period for the 75 individuals with incident MCI or AD was 5.6 ± 3.1 years. The groups did not differ in MMSE score (P = .30) but differed significantly in STMS score (P < .001) (Table 4 and Table 5). The result was not influenced by age, sex, or education (P = .76 for MMSE and P = .001 for STMS, testing group differences by logistic regression analysis and controlling for age, sex, and education). One half of subjects with incident MCI or AD scored 2 or lower on the STMS recall substests at baseline compared with 25% of the clinically stable normal cognition group (P < .001).

As expected, the mean total scores on both the STMS and the MMSE were higher in the stable normal cognition group compared with the prevalent MCI group and in the MCI group compared with the prevalent mild AD group (Table 1). In paired comparisons of subjects with stable normal cognition vs those with prevalent MCI, stable normal cognition vs AD, and MCI vs AD, the difference in means between groups was highly statistically significant for both the MMSE and the STMS (P < .001 for all comparisons).

The Figure depicts the receiver operating characteristic curves of the STMS and the MMSE for discriminating the subjects with prevalent MCI from those with stable normal cognition. The area under the receiver operating characteristic curve was modestly but significantly better for the STMS compared with the MMSE for discriminating between diagnostic groups (stable normal cognition vs prevalent MCI). Table 3 presents the quantitative data for that contrast and for the contrast between groups defined by CDR score (CDR score of 0 vs CDR score of 0.5 to 1.5). For discriminating between prevalent MCI and AD, there was no difference in the performance of the 2 tests. To put the magnitude of the differences between the STMS and the MMSE in perspective, we also conducted an analysis in which we dropped the delayed-recall item from the MMSE. The increase in area under the curve from the full MMSE to the STMS was comparable with the difference in area under the curve from the MMSE with the delayed-recall element deleted compared with the full MMSE (lower rows of Table 3).

In our sample, a cutoff score of 24 on the MMSE had very poor sensitivity for detecting MCI, with a high specificity. A cutoff score of 29 had 82% sensitivity but a specificity of only 48%. The STMS exhibited a similar pattern of trade-offs between sensitivity and specificity across a large range of scores. Sensitivity and specificity did not exceed 80% for either test or any test value.

Table 3. Receiver Operating Characteristic Curve Values*  

<table>
<thead>
<tr>
<th>Groups compared†</th>
<th>AUC (95% CI)</th>
<th>AUC (95% CI)</th>
<th>χ² Statistic</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDRSoB = 0 vs CDRSoB = 0.5-1.5</td>
<td>0.66 (0.60-0.71)</td>
<td>0.71 (0.67-0.76)</td>
<td>3.98</td>
<td>.05</td>
</tr>
<tr>
<td>CDRSoB = 0.5-1.5 vs CDRSoB = 2.0-3.5</td>
<td>0.77 (0.69-0.85)</td>
<td>0.78 (0.70-0.85)</td>
<td>0.02</td>
<td>.88</td>
</tr>
<tr>
<td>Normal cognition vs prevalent MCI</td>
<td>0.75 (0.70-0.80)</td>
<td>0.82 (0.79-0.86)</td>
<td>9.75</td>
<td>.002</td>
</tr>
<tr>
<td>Normal cognition vs prevalent AD</td>
<td>0.96 (0.94-0.97)</td>
<td>0.97 (0.96-0.98)</td>
<td>3.05</td>
<td>.06</td>
</tr>
<tr>
<td>Prevalent MCI vs prevalent AD</td>
<td>0.86 (0.83-0.90)</td>
<td>0.88 (0.84-0.91)</td>
<td>0.65</td>
<td>.42</td>
</tr>
<tr>
<td>Normal cognition vs prevalent MCI</td>
<td>0.74 (0.70-0.79)</td>
<td>0.70 (0.64-0.75)</td>
<td>12.70</td>
<td>.004</td>
</tr>
<tr>
<td>Normal cognition vs prevalent AD</td>
<td>0.96 (0.94-0.97)</td>
<td>0.94 (0.92-0.96)</td>
<td>8.94</td>
<td>.003</td>
</tr>
<tr>
<td>Prevalent MCI vs prevalent AD</td>
<td>0.86 (0.83-0.90)</td>
<td>0.85 (0.82-0.89)</td>
<td>1.30</td>
<td>.25</td>
</tr>
</tbody>
</table>

Abbreviations: AD, Alzheimer disease; AUC, area under the curve; CDRSoB, Clinical Dementia Rating Scale score sum of boxes; CI, confidence interval; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; STMS, Short Test of Mental Status.
*Determined using the χ² statistic and P value comparing the area under the MMSE and STMS receiver operating characteristic curves or under the full MMSE and MMSE without the delayed-recall item receiver operating characteristic curves.
†The MMSE vs STMS.
‡Full MMSE vs MMSE without delayed-recall item.
The present analysis showed that the STMS was slightly more effective than the MMSE in differentiating between cognitively healthy individuals and individuals with MCI. In addition, the STMS was superior to the MMSE in detecting deficits in cognition in individuals who had normal cognition but later developed incident MCI or AD. For individuals with dementia, the STMS and the MMSE were indistinguishable.

The differences between the 2 tests were modest, and the most conservative comparison of the STMS and the MMSE would be to say that they were very similar overall in their diagnostic accuracy. However, when the distinction between normal cognition and MCI was at stake, the STMS was better than the MMSE. The additional cognitive test items offered by the STMS revealed impairments in subjects with MCI compared with those who had normal cognition and also showed lower performance in subjects with normal cognition who subsequently developed MCI or AD.

A potential limitation of the analyses was the bias introduced by the availability of the STMS to the neurologists and the MMSE to the neuropsychologists at the time that the baseline diagnoses were formulated. We attempted to minimize these potential biases by using 2 different analytic strategies, both of which showed that the STMS was modestly superior to the MMSE. Another limitation in this data set was that our subjects were relatively well educated. Our results may not generalize to elderly individuals with low educational attainment.

The introduction of calculations, verbal similarities, and fund of information in the STMS was intentional, because it was developed for use with a population with a high school education. In individuals with fewer than 9 years of formal education, there might be fewer differences between the STMS and the MMSE.

The MMSE has been the mainstay of bedside cognitive testing. We propose that the STMS is equally effective and may have some features that make it more informative than the MMSE in persons with MCI. Bedside mental status assessment is only 1 aspect in the evaluation of cognitive impairment. Neither the STMS nor the MMSE can be used alone to diagnose MCI or dementia. Clinical judgment and neuropsychological testing are integral in diagnosing MCI.

Accepted for publication August 19, 2003.

**Author contributions:** Study concept and design (Drs Tang-Wai, Knopman, Geda, Smith, Tangalos, Boeve, and Petersen); acquisition of data (Drs Tang-Wai, Knopman, Smith, Ivnik, Tangalos, Boeve, and Petersen); analysis and interpretation of data (Drs Tang-Wai, Knopman, Edland, Boeve, and Petersen); drafting of the manuscript (Drs Tang-Wai, Knopman, Geda, and Petersen); critical revision of the manuscript for important intellectual content (Drs Tang-Wai, Knopman, Edland, Smith, Ivnik, Tangalos, Boeve, and Petersen); statistical expertise (Drs Knopman, Edland, and Smith); obtained funding (Drs Smith and Petersen); administrative, technical, and material support (Drs Ivnik, Tangalos, Boeve, and Petersen); study supervision (Drs Knopman, Boeve, and Petersen).

This study was supported by grants AG 16574 and AG 07216 from the National Institute on Aging, Bethesda, Md.

Portions of this study were presented at the American Academy of Neurology Annual Meeting, April 2, 2003, Honolulu, Hawaii.

We acknowledge the contributions of Kris Johnson, RN, and the nurses, neuropsychometrists, and allied health staff at the Mayo Alzheimer’s Disease Research Center and Healthy Aging Project, Rochester, for input and efforts obtaining information used in this article. We would also like to acknowledge Tiffani Slusser, BS; Matt Plevak, BS; and Tera B. Carpenter, BS, for assistance with this project.
This article is dedicated to the memory of Emre Kokmen, MD, the developer of the Short Test of Mental Status.

Corresponding author and reprints: David S. Knopman, MD, Department of Neurology, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (e-mail: knopman@mayo.edu).

REFERENCES


