A score based on screening tests to differentiate mild cognitive impairment from subjective memory complaints

Fábio Henrique de Gobbi Porto,1 Lívia Spindola,1 Maira Okada de Oliveira,1 Patricia Helena Figuerêdo do Vale,1 Marco Orsini,2 Ricardo Nitrini,1 Sonia Maria Dozzi Brucki1
1Behavioral and Cognitive Neurology Unit, Department of Neurology, Cognitive Disorders Reference Center, das Clínicas Hospital, University of São Paulo; 2Neurology Department, Antonio Pedro University Hospital, Federal Fluminense University, Rio de Janeiro, Brazil

Abstract

It is not easy to differentiate patients with mild cognitive impairment (MCI) from subjective memory complainers (SMC). Assessments with screening cognitive tools are essential, particularly in primary care where most patients are seen. The objective of this study was to evaluate the diagnostic accuracy of screening cognitive tests and to propose a score derived from screening tests. Elderly subjects with memory complaints were evaluated using the Mini Mental State Examination (MMSE) and the Brief Cognitive Battery (BCB). We added two delayed recalls in the MMSE (a delayed recall and a late-delayed recall, LDR), and also a phonemic fluency test of letter P fluency (LPF). A score was created based on these tests. The diagnoses were made on the basis of clinical consensus and neuropsychological testing. Receiver operating characteristic curve analyses were used to determine area under the curve (AUC), the sensitivity and specificity for each test separately and for the final proposed score. MMSE, LDR, LPF and delayed recall of BCB scores reach statistically significant differences between groups (P=0.000, 0.03, 0.001 and 0.01, respectively). Sensitivity, specificity and AUC were MMSE: 64%, 79% and 0.75 (cut off <29); LDR: 56%, 62% and 0.62 (cut off <3); LPF: 71%, 71% and 0.71 (cut off <14); delayed recall of BCB: 56%, 82% and 0.68 (cut off <9). The proposed score reached a sensitivity of 88% and 76% and specificity of 62% and 75% for cut off over 1 and over 2, respectively. AUC were 0.81. In conclusion, a score created from screening tests is capable of discriminating MCI from SMC with moderate to good accuracy.

Introduction

Mild cognitive impairment (MCI) is a heterogeneous condition characterized by subjective complaints of cognitive decline, supported by objective decline in neuropsychological evaluation and relative preservation of functionality, which does not meet criteria for the diagnosis of dementia.14 It may be due to degenerative diseases,2 vascular lesions,4 be associated with psychiatric disorders,2 or have other causes and may represent a transitional state between the cognition of normal aging and mild dementia.9 Patients with subjective complaints of cognitive decline but without abnormalities in objective testing have been defined as pre-mild cognitive impairment (pre-MCI).2-4 As there are not enough data to indicate that all of these patients will become MCI, we prefer to refer to them as subjective memory complainers (SMC). Despite the fact that some patients after neurophysiological testing (NT) have been diagnosed as non- amnestic MCI, almost all patients complained of memory problems. Because of this, we called them subjective memory complainers. They represent an understudied population with unknown etiological diagnosis and long-term outcomes.14 Further studies are needed to better characterize this population until a definitive labeling of pre-MCI is possible.

Although there are no approved drugs for MCI,7 some non-pharmacological interventions such as physical exercise and cognitive training have been shown to be effective.12-15 Physical exercise and cognitive rehabilitation are examples of interventions that may delay the progression of MCI to dementia. New drugs are being studied in MCI, such as intranasal insulin and a nicotine patch.16,17 Furthermore, with the possibility of disease-modifying treatments,18 such as anti-amyloid beta monoclonal antibody, the identification of MCI is becoming more important.

In practice, health-care providers frequently have to deal with complaints of memory problems.9 The prevalence of memory complaints varied widely across several studies, ranging from 11% to 70%.20-25 In a busy daily practice where time is an important and limiting factor, the complaints needed to be expressed by the patient to try improving differentiation between the two conditions.

Materials and Methods

This study involved volunteers living within a community environment aged over 60 years who complained of memory problems, recruited mostly from community centers for the elderly. The complaints needed to be expressed spontaneously by the subjects and were also evaluated by the Memory Complaint Questionnaire.24 Study subjects were submitted to a clinical evaluation and screening cognitive battery which included the Mini Mental State Examination (MMSE)27,28 and Brief Cognitive Battery (BCB).29,30 The BCB is a visual learning test in which subjects have 3 attempts to learn 10 simple line drawings: shoe, house, comb, key, airplane, turtle, book, spoon, tree, and bucket. Initially, the 10 figures are shown and the subjects are instructed to name them. After that, the pictures are withdrawn and they are asked to remember the name of the pictures spontaneously. The pictures remembered are scored as incidental memory. The pictures are then shown again for 30 seconds (s) and subjects are asked to try
to memorize the figures on them. The pictures are removed and subjects have one minute to remember the name of the objects. This is repeated once (2 attempts of 30 s to learn the 10 pictures). The number of pictures learned after the second attempt is scored as learned memory. After, functioning also as a distractor activity, subjects perform a categorical fluency test (number of animals in 1 min) and the clock drawing test (CDT). The next step is a delayed recall of drawings where subjects have one minute to try to remember the pictures (called delayed recall) and a recognition of the previous learned drawings combined with 10 other drawings. This battery has shown good accuracy in diagnosing early dementia, with little influence of level of education. In an attempt to evaluate an easy and fast way to improve the accuracy of both tests for this population, we added two delayed recalls of the three words of the MMSE, one after the MMSE itself, which we called delayed recall (DR), and another after the BCB (carried out in sequence after the MMSE), called late-delayed recall (LDR). We also included a phonemic verbal fluency test (number of words in 1 min) with the letter P (LPF) in the BCB after the animal fluency test (also increasing the distraction interval of the test). We used the Geriatric Depression Scale (GDS) to evaluate the intensity of depressive and anxiety symptoms, respectively. The Functional Activities Questionnaire (FAQ) was used to evaluate functional status. We excluded individuals with FAQ over 4 and/or GDS over 5 or those considered to have dementia, or other active neuropsychiatric conditions through consensus discussion. Subjects with a past history of or neurological evidence of stroke, neurodegenerative disorders, head injury, serious non-compensated medical illness, drug abuse, hearing, visual or motor impairment that could have affected their cognitive performance were not included. For this analysis, we included only patients with over eight years of formal education. After the initial evaluation, all eligible patients were submitted to NT with a neuropsychologist with experience in cognitive assessment, patients were submitted to NT with a neuropsychologist with experience in cognitive assessment and behavioral neurology. Patients were classified as MCI or SMC according to the presence or absence of cognitive deficits in NT (considered the gold standard in this study). We considered a cognitive function to be impaired if the score on that function were lower than −1.5 standard deviation (SD) in one test or if in more than one test of the same function, the scores were between −1 and −1.5 SD.

Statistical analysis was carried out using SPSS software, version 17.0 (SPSS, Inc., Chicago, IL, USA). We used the Kolmogorov-Smirnov test to verify the normality of the data. Student’s t-test scores for comparative differences between MCI and SMC, and ROC curve analysis to determine accuracy, sensitivity and specificity of different tests. After that, we created a proposal score with the tests with greater area under the curve (AUC). For all analyses 0.05 was considered significant. The study was approved by our institutional ethics committee. All patients taking part in the study gave written informed consent before evaluation.

**Results**

A total of 106 patients (32 SMC and 74 MCI) were included in our study sample. Demographic data, scores of GAI, GDS, and screening cognitive tests are shown in Table 1.

**Discussion**

In everyday clinical practice, it is sometimes difficult to distinguish between pathological cognitive decline and aging-related cognitive decline. Complaints such as subtle forgetfulness, problems in remembering names, misplacing objects, and a lack of attention are very common among elderly people and may not necessarily be a sign of cognitive disorders. Furthermore, the relationship between memo-
ry complaints and cognitive performance may not be straightforward. Most studies have found that subjective memory complaints have a better relationship with psychiatric symptoms rather than cognitive performance on NT,20,21,23,25 although worse cognitive performance has been demonstrated.24,26 Some authors have even suggested that memory complaints should be removed from the diagnostic criteria of MCI because of its poor relationship with cognitive performance.4,6 Even though NT is not essential to the diagnostic criteria for MCI,14 it is very useful in assessing people with memory complaints, especially in borderline cases (the so-called early MCI).14 Furthermore, NT may be helpful in differentiating amnestic from non-amnestic patients and in grading the cognitive decline. Although it is not to be taken as a rule, on average the suggested intensity of cognitive decline in MCI is between −1 and −2 SD. However, NT is not available for every patient, mainly in the primary care setting where most patients are initially seen. A rapid, sensitive and easy to apply screening test is useful for referring suspect patients to a specialized center.47

In this study, we used two quick screening tests (MMSE and BCB), and added two delayed recalls of the words of the MMSE and one phonemic fluency test (letter P). From these tests, a score was created in an attempt to improve diagnostic accuracy. The approximate duration of the tests was less than 15 min. Our results showed that when compared with NT, screening cognitive tests had only a moderate accuracy in differentiating between MCI and SMC. AUC in ROC ranged from 0.62 to 0.75, sensitivity ranged from 56% to 71% and specificity from 62% to 82%. A score derived from these tests could improve the accuracy (AUC ~0.81). Sensitivity ranged from 76% to 88% and specificity from 62% to 75%, depending on the cut-off value. These results are comparable to the accuracy of other brief neuropsychological batteries while having the advantage of being quicker and saving time. The CAMCOG battery has been shown to have a sensitivity of 64%, specificity of 88% and AUC of 0.83 when used to compare controls and MCI patients.44 In a study of screening tests for MCI, a combination of the MMSE, a categorical fluency test (animals) and the CDT, the authors have concluded that the combination of tests does not have a good diagnostic accuracy for identifying cases of MCI, in spite of their usefulness in the diagnostic screening for dementia.49 Others attempted to combine MMSE and CDT, showing a sensitivity of approximately only 50%.30 Combination of MMSE with an informant-based functionality scale, the Informant Questionnaire of Cognitive Decline in the Elderly (IQCODE), also showed only moderate accuracy when attempting to discriminate MCI from controls (AUC 0.7, sensitivity 73.7%, specificity 62.7%).31

Despite being the most known, most used, and most cited tool for cognitive screening, the MMSE has been shown to have little sensitivity for anterograde amnesia and executive function, especially in borderline cases.22 Other authors have already proposed delayed recalls of the words of the MMSE to improve its accuracy for detecting memory impairment. Loewenstein et al., adding three delayed recalls of the words, have shown a sensitivity of 83.3% and specificity of 90.4% in differentiating MCI from control.32 In our sample, a delayed recall of the words of the MMSE could improve the detection of early memory impairment. In the MMSE itself, there was no statistically significant difference between groups in the recall of the three words (P=0.457). The strategy of prolonging the time available and including more tasks of distraction between exposure and recall probably makes more demands on the anterograde memory systems, making this a more sensitive task.

As discussed above, the MMSE has low sensitivity to detect mild executive function impairment, so letter P fluency was added to the BCB (that already includes a categorical fluency test) to try to improve detection of executive dysfunction. Even though both categorical and phonemic fluency tests involve word production (language), the former is dependent on semantic memory systems and the latter on executive control, to implement mental planning and searching strategies.42,53 It has been shown to represent a surrogate of dorsolateral pre-frontal cortex function, bilaterally.44 Letter P fluency had the biggest isolated AUC in our sample. Interestingly, our results showed that categorical fluency did not differentiate between SMC and MCI: M (SD); 17.4 (4.4) and 16.5 (3.8); P=0.338. This is the opposite to that found in some reports in which categorical verbal fluency (a marker of semantic knowledge) was found to be more impaired in MCI in relation to control than phonological fluency,35 which maintains the same pattern as patients with Alzheimer's disease when compared with healthy control.36 On the other hand, some authors found no such dissociation, reporting impairment in both fluency tests.37,38. In our sample, phonemic fluency (letter P) was shown to be superior than a categorical fluency (animals) test in differentiating early cognitive impairment from subjective memory complain- ers. This could be due to an early executive dys- function in patients with MCI. Taking into account that we excluded patients with GDS over 5 or with other active psychiatric disor- ders, in our sample no differences were seen between scores of depression and anxiety scales assessed by GDS and GAL, respectively. Although the majority of authors have correlat- ed the presence of memory complaints to psy- chological symptoms,20,21,23-25 our results cannot confirm this because in our sample all subjects studied had memory complaints. However, poorer cognitive performance in the MCI group could not be explained by more intense depressive or anxiety symptoms.

### Conclusions

In conclusion, our results demonstrated that a score derived from screening tests can discrimi- nate between MCI and SMC with moderate to good accuracy. Simple modifications may improve accuracy and make such tests better able to identify those who will be diagnos- ed to have MCI in the NT. Phonological flu- ency test (letter P) and a delayed recall of the words of the MMSE were shown to improve the accuracy of screening tests. Although screening tests may be appropriate in primary care or at first visit, these findings underline the importance of NT for better evaluation of those patients in whom there is a suspicion of early cognitive impairment.

Our study has some limitations. Firstly, we included only patients with more than eight years of formal education and excluded patients with GDS over 5. This could reduce the relevance of our results in developing countries with low levels of education and also for patients with depressive symptoms. It is possible that when

### Table 2. Sensitivity, specificity, area under dc curve of screening test and proposed score.

<table>
<thead>
<tr>
<th></th>
<th>SENS*</th>
<th>SPEC*</th>
<th>MCI=74</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=106</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>&lt;29</td>
<td>64</td>
<td>79</td>
<td>0.75</td>
</tr>
<tr>
<td>LDR</td>
<td>&lt;3</td>
<td>56</td>
<td>62</td>
<td>0.62</td>
</tr>
<tr>
<td>DR-BCB</td>
<td>&lt;9</td>
<td>56</td>
<td>82</td>
<td>0.68</td>
</tr>
<tr>
<td>LPF</td>
<td>&lt;14</td>
<td>71</td>
<td>71</td>
<td>0.71</td>
</tr>
<tr>
<td>FSC</td>
<td>&gt;1</td>
<td>88</td>
<td>62</td>
<td>0.81</td>
</tr>
<tr>
<td>FSC</td>
<td>&gt;2</td>
<td>76</td>
<td>75</td>
<td></td>
</tr>
</tbody>
</table>

AUC, area under the curve; DR-BCB, delayed recall in brief cognitive battery; FSC, final score; LDR, late-delayed recall of mini mental state examination’s words; LPF, letter P fluency; MCI, mild cognitive impairment; MMSE, mini mental state examination; SENS, sensitivity; SMC, subjective memory complaint; SPEC, specificity. Percentage.
used in subjects with less than eight years of formal education, these screening tests and the proposed score might have lower sensitivity and specificity. We chose to set this education level as part of inclusion criteria to avoid differences between groups in this variable. Also, schooling showed a trend for non-statistically significant difference between SMC and MCI. In our study, both groups were highly educated so level of education cannot be responsible for the differences observed between groups. Besides this, all NT scores were adjusted according to schooling. Second, we used NT as gold standard of the diagnosis. Even though there is no consensus about the use of formal assessment as gold standard to diagnose MCI, it is a well-accepted practice. In our opinion, NT is important for early diagnosis of MCI. Also, we did not differentiate between the subtypes of MCI. Since the main objective of the study was to evaluate the usefulness of screening tests before NT, such a subdivision would be unfair. Most of the patients were amnestic MCI. Lastly, our results need to be prospectively tested in a different population with a large number of subjects and wider levels of education in order to further confirm the overall accuracy of the proposed score.

References

49. Ladeira RB, Diniz BS, Nunes PV, Forlenza OV. Combining cognitive screening tests for the evaluation of mild cognitive impairment in the elderly. Clinics (Sao Paulo) 2009;64:967-73.