Cognitive assessment in the elderly: a review of clinical methods

H.J. WOODFORD and J. GEORGE

From the Department of Medicine for the Elderly, Cumberland Infirmary, Carlisle, UK

Introduction

Doctors are poor at predicting patients’ cognitive function based on a routine, non-cognitive evaluation alone.\textsuperscript{1,2} Cognitive assessment is a valuable clinical skill. It facilitates the diagnosis of disorders that impair thinking, and allows for more accurate estimates of functional ability to be made. Cognition also predicts mortality during hospital admissions.\textsuperscript{3} These benefits are clearly of practical value, and cognitive assessment may also be a skill that could be used in the Directly Observed Procedural Skills (DOPS) framework for on-going evaluation of training geriatricians and neurologists.

As people age, changes occur within the brain that lead to differences in thinking and behaviour.\textsuperscript{4} Distinguishing these from the early stages of an abnormal (disease) process is fairly arbitrary; definitions usually depend upon an impact on social, functional or occupational activities. To further cloud the picture, there is a diagnosis of ‘cognitive impairment not dementia’ (CIND) also available to the clinician, which sits somewhere between normal ageing and dementia. Dementia rises in prevalence from <1\% of people aged <65 years, to an estimated 3–11\% of those aged >65 years, and to \sim33\% of those aged >85 years.\textsuperscript{5–8} CIND is even more common, with an estimated prevalence of around 17\% in people aged >65 years.\textsuperscript{9} Of course, changes in cognition are not specific to either CIND or dementia: other common causes in the elderly include delirium and depression. Two or more of these are frequently found within a single patient. Other, psychiatric (or ‘non-organic’), cognitive disorders (e.g. schizophrenia) are beyond the scope of this text.

Cognitive assessment is commonly used for the following reasons: (i) screening for cognitive impairment; (ii) differential diagnosis of cause; (iii) rating of severity of disorder, or monitoring disease progression. These three factors form the basis of the structure of this review.

A wide range of tools has been developed to aid the clinician in this process. These vary from brief screening tools that take <1 min to complete, to formal neuropsychological assessments that take several hours; appropriate choices depend both on the time available and the purpose of assessment. This article aims to review the literature regarding the more common techniques used by physicians for evaluating cognition. It also attempts to establish a practical framework for their usage by busy clinicians.

Methods of review

cognitive impairment. Additional articles were identified by hand searches of relevant reference lists. Only articles available in English were considered. The discussion of every assessment technique is beyond the scope of this review; instead, particular focus has been given to those tests in which there are comparative data with a similar, commonly used instrument, in the hope of permitting meaningful conclusions.

Screening for cognitive impairment

Surveys of psychogeriatricians in the USA, Canada and the UK have consistently found the Mini Mental State Examination (MMSE) to be the most commonly used cognitive screening tool by some distance. However, the most frequent reason for poor acceptance of this test by clinicians is the time taken to perform it. There is thus a practical need for a briefer test than the MMSE that retains reasonable sensitivity and specificity.

This section discusses first the MMSE, and then some of the other tools more commonly used in screening for cognitive impairment, which typically take <5 min to complete. The value of all of these methods is harder to establish in the presence of illiteracy, dysphasia, and sensory loss, or in people who do not speak English as a first language (translated versions are available for some of them but they have not all been validated). Differences in the use and interpretation of such tests have been reported and may result in large inter-user variability.

Mini Mental State Examination (MMSE)

The MMSE is a 30-point assessment tool. It was initially developed as a screening test to distinguish ‘organic’ from ‘non-organic’ (e.g. schizophrenia) cognitive disorders. More recently, it has become a common method of screening for, and monitoring the progression of, dementia and delirium. Generally the MMSE correlates well with other cognitive screening test scores, and reasonably well with a number of neuropsychological tests.

It takes ~8 min to perform in hospitalized elderly patients (range 4–21 min). There is some inter-user difference in scoring tests, and some variation in the questions; for example, the calculation task can use either serial 7s (subtracting 7s from 100) or spelling of the word ‘WORLD’ backwards—in some cases both are performed, and the highest result is used. Additionally, orientation to place is dependent on the location in which the test is performed: a familiar (home) vs. unfamiliar (hospital) environment. As these items are not identical, they may result in score variance.

Various cut-off values have been advocated for the maximum sensitivity and specificity in differing populations. Scores can be biased by baseline educational level, language and cultural barriers; thus patients with lower levels of education may be wrongly classified as demented, and those with higher levels may be missed. Values of 23 or less for those with education up to high school, and 25 or less for those who underwent higher education are commonly used to indicate significant impairment. As MMSE scores generally decline with advancing age, other authors have recommended lower cut-offs for the elderly, perhaps as low as 20 or less to indicate impairment. However, average lower scores in the elderly may simply reflect the higher prevalence of dementia in this age group.

It has both a ‘ceiling’ and ‘floor’ effect: a score of 30 does not always mean normal cognitive function, and a score of zero does not mean an absolute absence of cognition. It does not contain much capacity to test frontal/executive or visuospatial (typically right parietal) functions. The pentagon task of the MMSE simply requires the patient to copy the image, and does not assess planning skills. As a consequence it may have a limited ability to detect non-Alzheimer’s dementias, such as post-stroke cognitive impairment, frontotemporal or subcortical dementias in their early phases.

Standardized Mini-Mental State Examination (SMMSE)

The SMMSE has been introduced in an attempt to reduce inter-rater variability in scores. It incorporates the same questions, but with clear guidance on the administration, scoring and time allowed for each of the components. The result is better correlation between testers, and a reduction in the time taken to administer the test.

Abbreviated Mental Test (AMT)

The Abbreviated Mental Test (AMT) is a brief, 10-item scale used to screen for impairments. It was derived by selecting 10 questions with the most discriminatory value from the longer Mental Test Score (rated out of 34). It includes components requiring intact short and long term memory, attention and orientation. A score of <8 is the usual cut-off suggesting a significant cognitive deficit. It may quickly provide a severity assessment comparable to that obtained by longer tests. It may have sufficient discriminatory ability to detect changes in cognition associated with the
post-operative development of delirium.\(^2^9\) It takes approximately 3 min to administer in elderly patients.\(^1^6\)

There is also a four-question version of the AMT (the AMT4), using the questions age, date of birth, place and year only. Scores achieved have been found to correlate reasonably well with those from the longer form of the AMT.\(^3^0\) It is even quicker to perform and easier for the examiner to remember, which may increase the chance of some form of cognitive assessment occurring routinely in elderly patients in busy hospital settings such as emergency departments.

**Six-Item Screener (SIS)**

The SIS is composed of three orientation questions (day, month and year) and a three-item recall task derived from the MMSE.\(^3^1\) Each item scores one point, with a lower score signifying more cognitive impairment. A cut-off of \(\leq 3\) was used in the original study; however, a more recent study advocated a cut-off of \(\leq 4\).\(^3^2\)

**Six-Item Cognitive Impairment Test (6CIT)**

The 6CIT is also known as the Short Orientation-Memory-Concentration Test, or the Blessed Orientation-Memory-Concentration Test. It is constructed from six items of the Blessed mental status test (initially 26 items).\(^3^3\) It includes one memory, two calculation and three orientation questions. The components are given a weighting when scored that leads to a value between 0 and 28, with higher numbers representing more significant cognitive impairment. This requirement for some mathematics may make this test less suitable for use in busy clinical settings, such as during ward rounds.

**Clock Drawing Test (CDT)**

The Clock Drawing Test (CDT) is a screen for visuospatial, constructional praxis and frontal/executive impairment. The patient is first asked to draw a circle, and then put on numbers as though it were a clock face. Inability to correctly space the numbers around the circle can be due to a visuospatial impairment, neglect, or a planning deficit. Finally they are asked to draw on the hands to represent a specific time. The time ‘10 past 11’ is typically used, as this tests the patients capacity to compute that the minute hand should be pointing to the number 2, rather than 10 (a frontal/executive function). This task also requires the use of both sides of the visual fields. Alternative versions of the test do exist, and in some the circle is pre-drawn. At least 15 different scoring systems have been tested, with some allocating over 30 marks.\(^3^4,3^5\) Opinion is divided as to which method is the best.\(^3^6,3^7\) The simplest is a three-point scale, with one mark for each of: a correctly drawn circle; appropriately spaced numbers; and hands that show the right time. Typically it takes just 1 to 2 min to perform. An additional advantage is the relative independence from bias due to intellect, language or cultural factors.\(^3^8\)

Opinion is divided over the merits of the CDT to detect early, mild changes in cognition and, therefore, its usefulness alone as a screening test.\(^3^9\)–\(^4^3\) Like other brief assessment tools, it is poor at distinguishing various subtypes of dementia.\(^4^4,4^5\) Although it may detect cognitive impairment, the CDT has no additional benefit in either the diagnosis or monitoring of delirium.\(^4^6\) Scores are not significantly effected in the presence of depression alone.\(^4^7,4^8\) A review of studies assessing the CDT suggested that the mean sensitivity and specificity were both \(\sim 85\%\),\(^4^9\) although the value obtained will depend on the population studied and the scoring system adopted. In an unselected community sample, values of \(83\%\) for sensitivity and \(72\%\) for specificity have been obtained.\(^5^0\)

**Mini-Cog**

The Mini-Cog adds a three-word recall test to the CDT, thus improving memory testing.\(^5^1\) It takes around 3 min to perform. Subjects are classified as having cognitive impairment if they are unable to recall any of three words (after performing the clock drawing) or if they recall only 1 or 2 words and draw an abnormal clock (i.e. any of the circle, numbers and hands are incorrect). The result obtained is thus that cognitive impairment is absent or present, rather than a numerical scale. This adds to its simplicity as a screening test, but means the test has no value in either monitoring disease progression or rating severity.

**The General Practitioner Assessment of Cognition (GPCOG)**

The GPCOG has similarities with the Mini-Cog, in that it combines a recall task (a name and address) with the CDT.\(^5^2\) There are also brief components testing memory of recent events and orientation. Additionally, there is a short informant questionnaire (see later) that is performed in those patients with intermediate scores from the first part. It is estimated to take 5–6 min to perform.\(^5^3\) However, this estimation will vary according to the number of patients going on to complete the informant component.
Comparing brief assessment scales

There is considerable difficulty in reliably comparing cognitive assessment scales. For example, sensitivities and specificities of the test will vary according to the population studied, the cut-off score considered abnormal, and the comparison 'gold standard' adopted. Evaluations have mainly been performed in selected populations with a high prevalence of dementia, rather than random population samples of elderly people. Other important population variables include levels of education, language spoken, and mean age. The estimated time taken to perform the evaluations should be interpreted with caution. Generally speaking, assessment tools take longer to perform in patients with cognitive impairment than in those without. Therefore, the proportion of cognitively impaired individuals within the study population affects the average times taken. Some aspects of selected tests are compared in Table 1. A summary of studies comparing brief screening tools is shown in Table 2.

The AMT has a lower sensitivity and specificity to detect cognitive impairment than the MMSE. The AMT4 appears to perform even less favourably, although it is particularly quick and easy to administer. The CDT correlates with the MMSE, but the sensitivity and specificity to detect dementia appear to be a little lower than those of the MMSE. The CDT does have some capacity to detect frontal/executive cognitive abnormalities in people who score normally on the MMSE. Generally, the 6CIT and the MMSE correlate well. In a selected sample of older people the 6CIT did better than the MMSE, especially in the detection of milder cases of dementia. This suggests that it may make a better screening test than the MMSE. The Mini-Cog has outperformed the MMSE in its ability to correctly classify ethnically diverse groups of elderly subjects, many of whom did not speak English, with suspected dementia. It has also performed similarly to the MMSE in an English-speaking community sample. However, in a study of older patients attending an emergency department, the Mini-Cog performed less well than the SIS when compared to the MMSE. The SIS also performed only a little worse than the MMSE in a community sample. The GPCOG appeared to outperform the MMSE in a general practice-based sample. However, the composition of the test was changed post hoc, and the available data are incomplete.

Two previous reviews of the use of brief screening tests to detect dementia in primary care settings have both concluded that the most suitable tests are the Mini-Cog, GPCOG or the Memory Impairment Screen (MIS). The basis for these conclusions is mainly due to reported performances in the

<table>
<thead>
<tr>
<th>Cognitive domain</th>
<th>AMT4</th>
<th>CDT</th>
<th>SIS</th>
<th>Mini-Cog</th>
<th>AMT</th>
<th>6CIT</th>
<th>GPCOG</th>
<th>MMSE</th>
<th>ACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>++</td>
</tr>
<tr>
<td>STM</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Remote</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Visuospatial/visuospatial</td>
<td>–</td>
<td>++</td>
<td>–</td>
<td>++</td>
<td>–</td>
<td>–</td>
<td>++</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>constructional praxis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal/executive</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>++</td>
</tr>
<tr>
<td>Orientation</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Attention/calculation</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>++</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td>Other aspects</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Informant component</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Equipment required</td>
<td>–</td>
<td>Pen and paper</td>
<td>–</td>
<td>Pen and paper</td>
<td>–</td>
<td>Pen and paper</td>
<td>Pen and paper</td>
<td>Pen, paper, watch and series of specialized pictures</td>
<td></td>
</tr>
<tr>
<td>Average time needed (min)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>20</td>
</tr>
</tbody>
</table>

–, Not specifically tested; +, minimal assessment; ++, moderate assessment; ++++, relatively extensive assessment. STM, short term memory.
### Table 2  Comparing brief screening tools

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Study design</th>
<th>Mean age (years)</th>
<th>Age criteria</th>
<th>Setting</th>
<th>Comparator test</th>
<th>Test(s) used</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swain 1999, 2000&lt;sup&gt;34&lt;/sup&gt;</td>
<td>276</td>
<td>Consecutive medical admissions</td>
<td>81</td>
<td>&gt;60</td>
<td>Medical admissions unit</td>
<td>MMSE &lt;24</td>
<td>AMT &lt;8</td>
<td>70</td>
<td>92</td>
<td>59% prevalence of cognitive impairment</td>
</tr>
<tr>
<td>MacKenzie 1996&lt;sup&gt;21&lt;/sup&gt;</td>
<td>145</td>
<td>Community sample</td>
<td>N/A</td>
<td>&gt;75</td>
<td>Community</td>
<td>CAMCOG &lt;70</td>
<td>MMSE &lt;21</td>
<td>AMT &lt;9</td>
<td>80</td>
<td>98</td>
</tr>
<tr>
<td>Nishiwaki 2004&lt;sup&gt;57&lt;/sup&gt;</td>
<td>3369</td>
<td>Community sample</td>
<td>80</td>
<td>&gt;75</td>
<td>Community</td>
<td>MMSE &lt;24, MMSE &lt;18</td>
<td>CDT &lt;3</td>
<td>45</td>
<td>91</td>
<td>CDT scale 0–4, 29% prevalence of dementia, 23% delirium</td>
</tr>
<tr>
<td>Kirby 2001&lt;sup&gt;48&lt;/sup&gt;</td>
<td>648</td>
<td>Community sample</td>
<td>75</td>
<td>&gt;65</td>
<td>Community</td>
<td>GMS-AGECAT</td>
<td>MMSE &lt;24, CDT &lt;6</td>
<td>88</td>
<td>88</td>
<td>CDT scale 0–10, 6% prevalence of dementia, 13% depression</td>
</tr>
<tr>
<td>Brooke 1999&lt;sup&gt;61&lt;/sup&gt;</td>
<td>287</td>
<td>135 volunteers vs. 152 patients with dementia known to a POAS</td>
<td>73</td>
<td>NA</td>
<td>Community</td>
<td>Global Deterioration Scale &gt;2</td>
<td>6CIT &lt;8, MMSE &lt;24</td>
<td>90</td>
<td>100</td>
<td>53% prevalence of dementia</td>
</tr>
<tr>
<td>Borson 2000&lt;sup&gt;51&lt;/sup&gt;</td>
<td>249</td>
<td>Community sample of people known to social services</td>
<td>74</td>
<td>NA</td>
<td>Community</td>
<td>CERAD, DSM-IV and NINCDS-ADRA criteria</td>
<td>MMSE &lt;24, Mini-Cog</td>
<td>91</td>
<td>92</td>
<td>50% non-English speaking, 52% prevalence of dementia</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Study design</th>
<th>Mean age</th>
<th>Age criteria (years)</th>
<th>Setting</th>
<th>Comparator test</th>
<th>Test(s) used</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borson 2003</td>
<td>1179</td>
<td>Random community sample</td>
<td>73</td>
<td>&gt;65</td>
<td>Community</td>
<td>CERAD, DSM-III, NINCDS-ADRD and CDR criteria</td>
<td>MMSE &lt;24 Mini-Cog</td>
<td>71</td>
<td>94</td>
<td>6.4% prevalence of dementia</td>
</tr>
<tr>
<td>Borson 2005</td>
<td>371</td>
<td>Community sample of people known to social services</td>
<td>75</td>
<td>NA</td>
<td>Community</td>
<td>CASI</td>
<td>MMSE &lt;24 Mini-Cog</td>
<td>81</td>
<td>81</td>
<td>64% non-English speaking, 62% prevalence of cognitive impairment</td>
</tr>
<tr>
<td>Callahan 2002</td>
<td>344</td>
<td>Random sample of community dwelling black people</td>
<td>74</td>
<td>&gt;65</td>
<td>Community</td>
<td>BDRS, Word List Recall</td>
<td>MMSE &lt;24 SIS &lt;4</td>
<td>95</td>
<td>87</td>
<td>4.3% prevalence of dementia</td>
</tr>
<tr>
<td>Wilber 2005</td>
<td>149</td>
<td>Randomized: 74 SIS</td>
<td>75</td>
<td>&gt;65</td>
<td>Emergency department</td>
<td>MMSE ≤23 SIS ≤5 Mini-Cog</td>
<td></td>
<td>94</td>
<td>86</td>
<td>23% prevalence of cognitive impairment</td>
</tr>
<tr>
<td>Brodaty 2000</td>
<td>283</td>
<td>Community sample</td>
<td>80</td>
<td>50–74 with symptoms, or &gt;75</td>
<td>General practice</td>
<td>CAMDEX, DSM-IV</td>
<td>GPCOG AMT &lt;8 MMSE &lt;25</td>
<td>85</td>
<td>86</td>
<td>29% prevalence of dementia</td>
</tr>
</tbody>
</table>

BDRS, Blessed Dementia Rating Scale; CDR, Clinical Dementia Rating; CERAD, Consortium to Establish a Registry for Alzheimer’s Disease; GMS-AGECAT, Geriatric Mental State-Automated Geriatric Examination for Computer Assisted Taxonomy; POAS, Psychiatry of Old Age Service.
of Neurological and Communicative Disorders and Stroke-Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria require autopsy proof for ‘definite’ Alzheimer’s disease (AD),67 which is clearly unfeasible in most studies. The Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition) (DSM-IV) criteria provide limited differentiating ability between subtypes of dementia.68 For example, AD and vascular dementia (VaD) are chiefly discriminated by the former having a gradual onset and progressive nature, and the latter being associated with clinical or laboratory evidence of cerebrovascular disease ‘judged to be aetiologically related’. Clarity of diagnosis is further hampered by the high prevalence of mixed dementias and overlap between neurodegenerative conditions.69,70

Longer tests of cognitive function may provide sufficient information to define areas of cognitive impairment and so help in distinguishing between causes. The alternative is to perform an individually modified qualitative evaluation of the main cognitive domains. Assessment batteries will be discussed first, followed by alternative methods of evaluating specific cognitive domains.

**Assessment batteries**

**Depression**

The focus of this article is towards the evaluation of dementia and delirium. However, a reliable cognitive assessment should try to identify any significant component of depression. The gold standard for the diagnosis of depression is the psychiatric interview. Brief screening tools include the Geriatric Depression Score (GDS) and the Hospital Anxiety and Depression Scale (HADS), which contain 30 and 14 questions, respectively (although the GDS is also available in 5- and 15-question versions), and take around 10 min to complete.71–73 However, the validity of such screening instruments may be impaired in patients with significant cognitive impairment.74

**Delirium**

The Confusion Assessment Method (CAM) has been developed as a screening test to detect delirium.75 It has a very simple four-question format (based on DSM-III-R criteria) and is designed to be used by a wide range of health-care staff to allow early patient identification and appropriate management. It is deemed to be positive when items A and B plus C and/or D, from the following list, are present: (A) an acute confusional state with fluctuating...
Evidence suggests that the test can reliably be administered by individuals with only limited training. However, in a study of unselected elderly patients that compared it to the DSM-IV standardized diagnostic criteria, the sensitivity and specificity for delirium detection were only 81% and 84%, respectively. The CAM gives a ‘present’ or ‘absent’ result and so has no role in rating delirium severity. The longer Delirium Rating Scale (DRS) may be used for this purpose.

Dementia

Some assessment scales have been developed to try to distinguish between causes of dementia in a systematic manner. These include the Hachinski Ischaemic score for the detection of VaD and the Dementia of the Alzheimer Type (DAT) inventory for AD compared to other causes of dementia. Such scales may be more appropriate for usage in the setting of clinical trials.
Informant questionnaires

A relative, friend or carer who knows the patient well completes an informant questionnaire. They are not always feasible, due to the need for a suitable informant to be available. Their advantage is being able to look at more than just a snapshot in time, as they ask for an impression of change. For example, the history of onset and progression is extremely important when distinguishing between delirium and dementia. However, this information would usually be gathered by informal interview with a suitable source during standard assessment. Informant questionnaires usually give an impression of general decline rather than specific domains of cognitive impairment. They are not biased by the patient’s baseline educational level, but may be influenced by factors regarding the informant’s state of mind and relationship with the patient. Informant depression or poor relationship with the patient tends to cause an over-estimation of cognitive changes, whereas informants who do not live with the patient tend to underestimate changes.81 A range of different tests is available, and scores obtained generally correlate well with other cognitive screening tests.82 A number of tools also exist that incorporate both patient and informant questioning.52,83 In addition, some authors have proposed methods of adding informant rating scales to standard tools such as the MMSE to improve screening accuracy.84,85

An example is the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE), which asks a person who knows the patient well to answer 26 questions based on change in cognitive function over a 10-year period.86 This results in a score between 1 and 5, with higher values representing greater degrees of decline in cognitive function. It takes an estimated 10–15 min to complete.53 The MMSE has been found to be marginally better than the IQCODE in distinguishing cases of dementia from normal controls.87 Also, the AMT and MMSE outperformed the IQCODE in a population of geriatric patients.88 However, the reverse appeared to be true in a study of patients attending a geriatric day hospital.89

In another study, the MMSE scored similarly to an informant questionnaire.90 When compared to DSM-IIIR criteria in elderly people admitted as emergencies to a geriatric unit, sensitivities and specificities of 100% and 86% were obtained for the IQCODE, compared to 96% and 73% for the AMT (<8).91 The current data do not seem to indicate a particular advantage for informant questionnaires, and their place in the assessment of cognition is unclear.

Frontal lobe testing

The individual tests that can be used to assess frontal/executive functioning will be discussed later, but several brief batteries of tests of frontal function have also been developed.

The Frontal/Subcortical Assessment Battery (FSAB) uses a combination of a verbal fluency test, the Luria sequencing task and a ‘go/no go’ test. It was found to be useful in discriminating patients with subcortical dementia, who scored well on the MMSE, from normal controls.92 However, patients with cortical dementia (AD) also scored poorly in the FSAB and the additional use of the MMSE (on which they scored badly) was required to distinguish these patients.

The Frontal Lobe Score (FLS) was derived by performing a series of cognitive tests in 118 patients below the age of 70 years, in comparison to lesions seen on brain imaging studies.93 Compared to normal controls, it had a sensitivity of 78–92% and a specificity of 100% to detect frontal lobe lesions. However, compared to those with lesions outside the frontal lobes, these values fell to 78–92% and 75–84%, respectively.93,94 It has been estimated to take between 20 and 45 min to perform in this age group.94

The Frontal Assessment Battery (FAB) is a briefer screening tool of frontal/executive function, which is estimated to take around 10 min to complete.95 Subjects are allocated a score between 0 and 18, with lower scores indicating more severe impairment. It is effective at distinguishing patients with frontal lobe impairment from normal controls, but its ability to distinguish frontal impairment from other cognitive deficits (e.g. AD) is unknown.95

The Executive Interview (EXIT25) is a 25-item screening test of executive function that is estimated to take 15 min to complete, with higher scores representing greater impairment.96 As with the FSAB, patients with AD also score poorly on the EXIT25 and its combination with the MMSE (on which the AD patients score less well) is of greater discriminatory value.96,97

Extended versions of the MMSE

Recognising the limited cognitive domains that the MMSE assesses, there have been attempts to make it more comprehensive. Some authors advocate the combination of the CDT and MMSE for dementia screening.36,98,99 This approach is incorporated into several of the instruments discussed below.
Addenbrooke’s Cognitive Examination (ACE)

The ACE is a 100-point scale that was initially developed as a way to distinguish between AD and frontotemporal dementia (FTD). It includes the questions from both the MMSE and the CDT. There are also more detailed components for memory and frontal/executive functioning. A specialized series of pictures is required for its use. A score of <87 is typically used to indicate a significant impairment. Low scores in individuals who do not have a clinical diagnosis of dementia are predictive of the future development of dementia. The ACE is reported to have a better ability than the MMSE to detect subcortical dementia syndromes. However, its ability to reliably distinguish between AD and FTD (based on a ratio of scores in different sub-sections) has been questioned.

Modified Mini Mental Status Examination (3MS)

The 3MS shares much in common with the ACE. It is an extension of the MMSE, with additional components of verbal fluency and extended memory testing, to make the overall score out of 100. It had a sensitivity of 88% and a specificity of 90% to detect dementia in a community sample of people aged ≥65 years, using a cut-off of <78. The most appropriate cut-off point used may need to be adjusted according to the age and education levels within the population being studied. In a large community sample of elderly people (n=1600, mean age 80 years) the 3MS had sensitivity and specificity of 86% and 87% respectively (cut-off <78) compared to values of 86% and 77% for the MMSE (cut-off <26).

Cognitive Abilities Screening Instrument (CASI)

The CASI contains questions from both the MMSE and 3MS. It is also scored out of 100, and is estimated to take 15–20 min to complete. The CASI had only a marginally better sensitivity and specificity in a selected population compared to the shorter MMSE.

Cambridge Cognitive Examination (CAMCOG)

The Cambridge Mental Disorders of the Elderly Examination (CAMDEX) is a structured schedule for the assessment of cognition in the elderly, including structured history-taking from the patient and an informant, a structured examination, and a mental state assessment (the CAMCOG). The full evaluation is estimated to take around 80 min to perform. The overall aim of the CAMDEX is not only to detect a problem, but also to identify the cause and rate the severity.

The CAMCOG shares the components of the MMSE, but also tests additional cognitive aspects (mainly praxis, abstract thinking and perception). The total score is up to 107, with a cut-off of <80 being typical to diagnose cognitive impairment. It is estimated to take ~30 min to complete. In a selected sample of patients with suspected cognitive impairment, the CAMCOG had moderately better sensitivity and specificity to detect dementia, compared to standardized criteria, than either the AMT or MMSE. In addition, it had a sensitivity of 92% and specificity of 96% to detect ‘organic’ mental impairment in a cohort of elderly people, compared to a sensitivity of 94% and specificity of 85% for the MMSE (cut off of <24) in the same sample. However, the combination of the CDT with the MMSE offered equivalent sensitivity and specificity in detecting dementia to the longer CAMCOG in a psychogeriatric out-patient setting.

Middlesex Elderly Assessment Memory Score (MEAMS)

The MEAMS is currently more commonly used by occupational therapists than doctors. It is scored out of 47, with lower scores indicating more significant cognitive impairment. In addition to areas covered by the MMSE, it has components designed to assess frontal and right parietal lobe function (verbal fluency, motor perseveration and fragmented letter perception). It requires the use of a specialized series of picture cards to complete. Subsections are scored in an ‘all or none’ fashion, which may account for some of the observed problems with test-retest reliability. The MEAMS has been compared to the MMSE in a small series of psychiatric in-patients with a diverse range of diagnoses. Generally, the two tests correlated reasonably well, but there appeared to be a benefit in the MEAMS for detecting non-dementia, isolated cognitive impairments.

Neuropsychological testing

Trained neuropsychologists are needed for formal neuropsychological testing. The data obtained are compared to normal population values, and can be adjusted for the patient’s baseline intelligence and previous educational level. It is composed of different tests, the exact constituents of which tend to vary between individual neuropsychologists and
depending on the specific population or clinical question to be addressed. Typically, it will take 1–3 h to perform. Although useful for selected patients, it is beyond the scope of standard geriatric care, and will not be discussed further in this review.

Qualitative assessments

A mental state examination targeted towards specific cognitive processes may define problems more efficiently and clearly than quantitative assessment scales. The anatomic location of impairments may be described, but the functions of cortical regions are complex, and much overlap exists. The key components to be evaluated are outlined below.

Attention

Attention is a term for the ability to focus on a task: when reduced, patients are easily distracted. A variety of bedside tests can be used to evaluate it. The simplest is to ask the patient to recite reverse sequences (e.g. to count down from 20 to 1, or list the months of the year backwards). Slightly more elaborate tests include serial 7s (the patient is asked to subtract 7 from 100, and then subtract 7 from the remaining number, and so on) and digit span testing (the patient is asked to repeat back sequences of numbers of increasing length). Attention is a basic requirement for being able to perform other elements of the cognitive assessment. When it is impaired, the results of subsequent tests may be hard to interpret. It is characteristically impaired in patients with delirium.

Memory

There are a number of components to memory. The crudest divide is into long- and short-term elements. Short-term memory is more dependent on an intact limbic system (mainly in the temporal lobes) than long-term memory, which is dependent upon other cortical processes. A more descriptive distinction divides memory into four subtypes: episodic, semantic, procedural, and working.112 Episodic memory is related to personal experiences (e.g. what you had for dinner last night), semantic memory to impersonal facts (e.g. the capital of France), procedural memory to performing actions (e.g. riding a bike), and working memory is the capacity to briefly ‘hold it in your head’ (e.g. for the period of time between reading and then dialling a previously unknown phone number).

Questions about orientation test both attention and short-term memory; new information, such as the time of day, must be continually learned. In the MMSE, patients are given three words to remember after a few minutes of distraction with a different task. In the ACE, there is also a name and address to recall. Patients with subcortical dementias are more likely to recall information when given clues (e.g. ‘it was a type of fruit’) than those with cortical deficits (reflecting the problem with memory retrieval rather than memory formation or storage). Visuospatial memories are usually encoded in the non-dominant parietal lobe: this is tested by asking the patient to draw from memory images that were shown to them some minutes previously. Remote memories are harder to test, as the information cannot always be validated (e.g. the name of the patient’s old headmaster). Historically important dates that are known by most people might be used. For example, the AMT asks for the dates of World War I.

Language

The presence of a language disturbance suggests a problem within the dominant hemisphere. The assessment of speech should have begun while taking the history. Dysarthria may be associated with subcortical dementias. When subtle, it can be exaggerated by asking the patient to repeat complex phrases (e.g. ‘West Register Street’). Other language problems seen with subcortical dementias include the loss of prosody (rhythmic and melodic quality) and reduced verbal fluency (see below),22 however, dysphasia is rare.

In testing comprehension, the patient is asked to obey first one-stage, and then more complex instructions (e.g. point to the window, then the ceiling, and then the door). This may detect a receptive dysphasia. Expressive dysphasia is identified by asking the patient to name objects such as a watch, and then name the smaller components (e.g. winder, strap, hands, etc.). Paraphrasia is a term for the incorrect insertion of words into sentences. This may include semantic errors (the insertion of incorrect, but related words), for example saying ‘dog’ instead of ‘cat’. There is also a rare disorder called ‘conductive aphasia’, which causes particular impairment with the repetition of phrases (e.g. ‘no ifs, ands or buts’) in the MMSE.

Visuospatial skills

Visuospatial disturbances can be caused by lesions in either hemisphere, but tend to be more severe when the non-dominant hemisphere is involved. They are common in both dementia and delirium, but are rarely seen with ‘non-organic’ cognitive disorders. The usual method of detecting such a deficit is to ask the patient to copy diagrams,
the commonest of which is the interlocking pentagons (the patient is judged to have succeeded when they have drawn two five-sided objects that overlap). More complex and three-dimensional objects may also be used. Visual memory can be tested, as mentioned earlier.

Frontal/executive function

Executive function is a term for the higher cerebral functions, mainly derived from the frontal lobes, but also involving subcortical connections with the basal ganglia and thalamus. They include components such as planning, abstract thought, and judgement. They are required to complete complex tasks.

There is a wide range of techniques available to the clinician to assess frontal lobe functions. They include the trail-making tests. In the simplest form, this involves joining a sequence of numbers scattered across a page, in the correct order (e.g. 1 → 2 → 3…). A more challenging test may alternate between numbers and letters (e.g. 1 → A → 2 → B → 3…). Asking the patient to interpret proverbs can assess abstract thought (e.g. what is meant by ‘a rolling stone gathers no moss?’). Alternatively they can be asked to describe similarities and differences between words (e.g. love and hate).

Verbal fluency requires both language and executive skills. Common assessment techniques include asking the patient to list as many animals as they can in a 1-min period, or list as many words as they can beginning with a specific letter (e.g. ‘P’). Scoring scales have been developed for the number of words listed. A small number of words recalled in the category suggests impairment. Normal individuals should generally manage at least 15 words within 1 min.

Tests of mental flexibility include the Wisconsin Card Sort Test113 (where the subject has to correctly categorize specialized cards in response to changing criteria) and ‘go/no go’ tests. In this latter category, the patient is typically asked to tap the desk in response to the examiner’s taps—one or twice. This is then varied so as the patient is asked to tap twice in response to one tap and once in response to two taps. They are then asked to not tap when the examiner taps once.

Rating severity of disorder, and monitoring disease progression

Some assessment scales provide a numerical rating of severity that may be useful for disease progression, rather than a binary ‘present’ or ‘absent’ conclusion, which may influence the choice of test for a specific task. AD is typically associated with an annual decline on the MMSE of ~3–4 points,15,114 making it a reasonable tool for monitoring disease progression in this condition. It has also been proposed as a tool to define severity of cognitive impairment, with scores between 23 and 18 being classed as mild, and scores of 17 and below being classed as severe.15 Longer qualitative scales may provide a greater sensitivity to detect cognitive change. However, the occurrence of functional impairment is more likely to be relevant to the patient and their carers than simple numerical scores. Performance in instrumental activities of daily living (IADLs), especially telephone and transportation use, taking own medication and handling finances, correlates well with cognitive impairment,115–117 although in people with other co-morbidities, it may be difficult to distinguish the specific effects of cognitive losses.

Summary

Screening tests are available that detect the presence of cognitive impairment but do not distinguish between causes. Brief tests with a reasonable sensitivity and specificity include the 6CIT, Mini-Cog and SIS. No test has been shown to have excellent discriminatory value in random community samples. Their use as screening tests for unselected populations is likely to result in more false positives than true positive cases. They may have a role in screening selected populations, such as geriatric patients seen in hospital emergency departments and clinics. Typically the shorter tests do not assess frontal/executive function, and this can mean that diagnoses such as FTD and subcortical dementias are missed. This may lead to patients functioning less well in their home environment than predicted.

Longer tests may have a small additional benefit in sensitivity and specificity to detect cognitive impairment, but their main roles may be to help define patterns of cognitive loss and to rate disease severity. However, the best method of classifying causes of cognitive impairment remains a comprehensive clinical evaluation. The most suitable technique for a given situation will be judged by the time available and the diagnostic accuracy required. Geriatricians and neurologists who see older people should be able to perform and teach cognitive assessment, and be familiar with the advantages and disadvantages of various techniques. All doctors who treat older people should be able to administer a short cognitive screening test...
References
43. Lee H, Swanwick GR, Coen RF, et al.
52. Brodaty H, Pond D, Kemp NM, et al.
54. Swain DG, O’Brien AG, Nightingale PG.
55. Koch HJ, Guttler K, Szecsey A.
56. Richardson HE, Glass JN.
59. Juby A, Tench S, Baker V.
65. Burns A, Lawlor B, Craig S.
69. Zekry D, Hauw J, Gold G.
79. Cummings JL, Benson F. Dementia of the Alzheimer Type: an inventory of diagnostic clinical features. JAGS 1986; 34:12–19.


