Screening for cognitive impairment among older people in black and minority ethnic groups

CHRIS PARKER* AND IAN PHILP

Sheffield Institute for Studies on Ageing, Sheffield, UK

Abstract

There is a well-documented tendency for cognitive tests to underestimate the abilities of older people in black and minority ethnic groups. This gives rise to a substantially higher risk of mistaken diagnosis of dementia. Reasons include differences in extent or focus of formal education, lack of familiarity with English, lack of literacy in own first language, and culture-specific factors related to individual test items. Attempts to improve the accuracy of screening for these groups have included adaptation of existing tests, including adjustment of cut-points, translation and replacement of culture-specific items. So-called ‘culture-free’ tests have also been developed, which are less dependent on language, literacy and other skills developed during formal education. Cultural modifications and evidence of cross-cultural performance are summarized here for traditional tests (Mini-Mental State Examination, Short Portable Mental Status Questionnaire, Short Orientation-Memory-Concentration Test, Abbreviated Mental Test Score, Clifton Assessment Procedures for the Elderly), and for culture-free tests (Clock Drawing Test, Mini-Cog, 7-minute screening battery, Time and Change Test). The evidence on unadapted traditional tests shows that short ones perform at least as well as longer ones, and are more consistent across cultural and educational groups. Cut-point adjustments have not been universally found successful in improving accuracy, and do not address issues of acceptability. Translated and/or culturally adapted versions exist for a number of tests: it is important to establish cut-points appropriate to the target populations. There are promising results on culture-free tests, which are seen as less threatening and require little language interpretation, but they require further evaluation.

Keywords: cognitive screening, dementia, minority ethnic, elderly

Introduction

Cognitive screening is an integral part of assessing the health or care needs of an older person. Cognitive decline has profound implications for the ability to live independently, and detecting dementia at an early stage can improve management decisions and support for carers. Drug therapies now have the potential to ameliorate the early to middle stages of the disease. However, screening which is sufficiently sensitive to pick up early cases will inevitably classify some people with no disease as impaired. This risk is substantially higher for people in black and minority ethnic (BME) groups: one study found that 6% of non-impaired white people and 42% of non-impaired black people were wrongly classified by the Mini-Mental State Examination (MMSE) [1]. A mistaken diagnosis of dementia can lead to unnecessary tests or treatment and cause emotional suffering for older people and their families.

In this article we provide an overview of cognitive screening among minority ethnic older people, with particular reference to the UK. The focus is on short tests suitable for brief screening in the primary care setting, rather than the longer tests predominantly used by specialists for clinical diagnosis. We outline the reasons for cultural bias in testing and the efforts made to reduce it. Evidence on the cross-cultural performance of standard screening tests is summarised, with recommendations on current best practice and future development.

UK minority ethnic groups: language and literacy

Over half of the minority ethnic people in the UK aged over 65 have their origins in South Asia [see Appendix 1 in the supplementary data on the journal website (http://www.ageing.oupjournals.org/)]. Gujarati-speaking Indians arrived either directly from the Indian sub-continent or from East Africa: the latter are more numerous and generally received more education. Many Gujarati women from rural areas of the sub-continent are unable to read or write in their own language. Indian Punjabi speakers comprise Sikhs from Punjab state, and Hindus from Punjab or East
Africa. Pakistanis from provincial districts generally speak a Punjabi dialect and may be unable to read or write, while those from the cities speak, read and write Urdu. Immigration from Bangladesh has been more recent: the most common language spoken is Bengali (also known as Bangla or Bangadeshi), with some speaking Urdu. There are high levels of illiteracy both in Bengali and in English, particularly among women. English is the first language of UK African-Caribbeans, who predominantly came from Jamaica, and there are smaller numbers in other black groups. Most older people of Chinese origin arrived from Hong Kong, and many do not speak, read or write English. The smaller groups from Eastern European countries have been in the UK for longer, and literacy in English is relatively high.

**Cognitive screening tests**

Traditional screening tests for cognitive impairment comprise a set of items to detect some of the principal intellectual deficits associated with dementia. Memory for previously learned information is tested by orientation items (including current date and place) and by asking for names of prominent people and dates of significant events. The ability to learn new information is tested by the presentation and later recall of names of objects, phrases or stories, with intervening items forming a distraction task. Items to test attention or concentration include counting backwards, saying the months in reverse order, and carrying out successive subtractions. Slightly longer screening tests may include language items such as naming objects, responding to a written or spoken command, or category fluency (naming as many examples as possible within a given time). Other commonly used items assess visuo-spatial skills, for example in copying a drawing.

The ideal screening test for non-specialist use would be brief and easy to score, require no information from a third party, no special equipment and little assessor training. It would be reliable, acceptable to the older person and the assessor, and its performance would be unaffected by the education, language or culture of the older person. High specificity would minimise incorrect diagnoses of dementia, and high sensitivity ensure that early cases were not missed. Increasing the number of test items might be expected to lead to better performance, but shorter tests with simple scoring systems have been found at least as accurate as longer ones [1, 2]. Orientation and memory items appear to discriminate better between impaired and non-impaired than the additional items in longer tests (see Appendices 2 and 3 in the supplementary data). A few tests exist which are not composites of the type of items described: these are considered later.

**Why do screening tests misclassify minority ethnic older people?**

The formal education of people growing up in developing countries is likely to have been relatively short, and less related to the tasks involved in cognitive tests [3, 4]. Educational level has a well-documented association with cognitive test scores: in some studies this accounts for all the apparent effect of ethnic group [e.g. 5], but in others there is a remaining effect (see Appendices 4 and 5). Most of the current generation of minority ethnic elders in the UK moved here as adults, and relatively few have acquired fluency and literacy in English. In one study, 35% of older Asians in a UK city could speak English, 21% could read and write in English, and 73% could read and write in their first language [6]. Any test items requiring numeracy, or literacy either in English or in their first language, may therefore disadvantage these older people. They will be further disadvantaged by any items assuming familiarity with a different culture to the one they experienced while growing up. Table 1 indicates potential sources of bias in the common types of test item, and provides references to studies which have identified educational or cultural item bias.

Adults migrating from a developing country therefore combine two potential sources of under-performance: lower educational attainment resulting from their childhood setting, and a lack of familiarity with the language and culture of the country they now live in.

<table>
<thead>
<tr>
<th>Item type</th>
<th>Examples</th>
<th>Potential bias</th>
<th>Observed bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientation to time</td>
<td>Date/season</td>
<td>Requires habitual use of western calendar; some groups use own cultural calendar [11].</td>
<td>Educational [20], cultural [12].</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>Address/names of nearby streets</td>
<td>Requires familiarity with UK-style addresses</td>
<td>Cultural [29, 12].</td>
</tr>
<tr>
<td>Attention/concentration</td>
<td>Serial subtraction/backward</td>
<td>Requires literacy/numeracy in language of test</td>
<td>Educational [20], cultural [29, 30].</td>
</tr>
<tr>
<td>Memory for previously learned</td>
<td>Date of World War I</td>
<td>Requires familiarity with culture, e.g. prominent people/events, or accurate knowledge of own date of birth (not recorded for some) [11]</td>
<td>Ethnic group [29].</td>
</tr>
<tr>
<td>information</td>
<td>Own date of birth/age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory for new information</td>
<td>Recall names of objects/recall</td>
<td>Requires familiarity with type of material learned e.g. UK-style address</td>
<td>Educational [20], cultural [30].</td>
</tr>
<tr>
<td>phrase</td>
<td>phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Follow written/verbal command</td>
<td>Requires literacy/fluency in language of test</td>
<td>Educational [20, 29], cultural [28, 30].</td>
</tr>
<tr>
<td>Visuo-spatial skills</td>
<td>Copy complex figure</td>
<td>May require familiarity with geometric shapes</td>
<td>Educational [20], cultural [30].</td>
</tr>
</tbody>
</table>
This combination may also make cognitive testing a more threatening experience for them. One study conducted in the majority population found high levels of anxiety among working class people when asked to write a sentence, spell a word backwards or perform calculations (see Appendix 6). This anxiety can only be exacerbated when the test is in an unfamiliar language and makes reference to unfamiliar cultural information, and in these circumstances the testing process cannot accurately reflect ability.

How have test developers addressed the problem?

Efforts to develop culturally appropriate cognitive screening have proceeded in two ways: by adapting existing culturally biased tests, and by developing new, relatively culture-free tests.

1. Adaptation

The effect of educational level on scores is commonly dealt with by adjustment of cut-points [e.g. 7]. The calculation is usually based on number of years of education, which addresses differences in duration but not in quality or focus: a suggested alternative is reading level [8]. Tests may have additional adjustments based on ethnic group [e.g. 9]. However, some authors have found adjustments of negligible benefit in improving accuracy [2, and also see Appendix 7], and cut-points calculated in ethnically homogeneous populations may not be relevant in mixed populations [10].

Translating tests into the older person’s preferred language should improve acceptability and performance, but items requiring literacy will still be a problem for those unable to read or write in their first language. Some items in standard tests pose particular problems in translation. Choosing a five-letter word to replace ‘world’ in the backward-spelling task poses technical difficulties in some languages [11]. A direct translation of the repetition phrase ‘no ifs, ands or buts’ would not be equivalent: the substitute ‘neither this nor that’ was chosen for the Gujarati and Hindi versions of the MMSE [12, 13]. Cultural bias is not necessarily removed by translation: asking for the date of a prominent event in the MMSE [12, 13]. Cultural bias is not necessarily removed by translation: asking for the date of a prominent event in the MMSE [12, 13]. Cultural bias is not necessarily removed by translation: asking for the date of a prominent event in the MMSE [12, 13]. Cultural bias is not necessarily removed by translation: asking for the date of a prominent event in the MMSE [12, 13]. Cultural bias is not necessarily removed by translation: asking for the date of a prominent event in the MMSE [12, 13].

Cognitive screening among BME older people

Predictive power has been found significantly related to language [16]. Translated versions include Gujarati [12, 14]; Punjabi, Urdu [14]; Hindi [13, 14]; Bengali [14]; Chinese (see Appendices 14 and 15). A culturally-adapted version exists for African-Caribbeans [21], and adjusted norms have also been provided for this group [see Appendix 16 in the supplementary data on the journal website (http://www.ageing.oupjournals.org/)].

How do existing screening tests perform for ethnic minority groups?

In this section a number of short screening tests are described, with a summary of their cultural adaptations and reported performance across ethnic groups. Caution is needed in interpreting evidence of cross-cultural performance: studies vary widely according to the population sampled and the test cut-point [19], and there are particular difficulties in achieving a ‘gold standard’ diagnosis which is consistent across groups with different language and culture (see Appendices 9 and 10). The majority of cross-cultural studies have been conducted in the US, where the history of immigration and mix of minority groups is very different to that found in the UK. The fact that a screening test shows different dementia prevalences across ethnic groups does not necessarily indicate bias. For example, higher rates of hypertension, diabetes and stroke in some groups may elevate the risk of cognitive impairment (see Appendix 11).

MMSE (see Appendix 12)

The MMSE is the most widely used and extensively studied measure of cognitive function. Nineteen items cover orientation to time and place, registration and recall, attention, language and figure copying. MMSE score is strongly related to education, with a high false positive rate for those with little education [20], and there are suggested cut-point adjustments for this [e.g. 7, and also see Appendix 13]. Specificity is lower among non-whites: some studies have found the effect of ethnicity is eliminated when adjusting for education and social class [e.g. 5]; in some it persists [e.g. 4]. Predictive power has been found significantly related to language [16]. Translated versions include Gujarati [12, 14]; Punjabi, Urdu [14]; Hindi [13, 14]; Bengali [14]; Chinese (see Appendices 14 and 15). A culturally-adapted version exists for African-Caribbeans [21], and adjusted norms have also been provided for this group [see Appendix 16 in the supplementary data on the journal website (http://www.ageing.oupjournals.org/)].

Short Portable Mental Status Questionnaire (SPMSQ) [9]

Ten items cover orientation to time and place, remote and recent memory, and attention. The scoring has adjusted cut-points, derived in a US population, for education and for ethnic group (one more error being allowed for black elders). After these adjustments, the SPMSQ has been found less culturally biased than other short tests [1].

Short Orientation-Memory-Concentration Test (OMC) (see Appendix 17)

Six items cover orientation to time, new learning and attention. OMC scores are related to education [2], and specificity in black groups is reported to be low [1].
Abbreviated Mental Test Score (AMTS)
(Appendix 18)
Ten items cover orientation to time and place, remote and recent memory, new learning, attention, and naming of two people personally known to the older person. Translated versions with good sensitivity and specificity exist in Gujarati, Punjabi, Urdu, Hindi, Bengali [14]. A culturally-adapted version exists for African-Caribbeans [21].

Clifton Assessment Procedures for the Elderly (CAPE) (see Appendix 19)
The Information-Orientation sub-test is a 12 item cognitive screen covering orientation to time and place, remote and recent memory. There is relatively little evidence on test performance, and none was found on cross-cultural validity.

Clock Drawing Test (CDT) (see Appendix 8)
The task is to draw a clock-face with all the numbers, and draw the hands to show a specified time. It is thought to tap a wide range of functions including executive control, long-term and working memory, visuospatial representation and global attention [22]. The CDT has been extensively studied and many scoring systems devised ranging from multi-point scales (see Appendix 20) to binary scoring as normal/impaired [10]. The binary system has been found effective even when used by untrained raters [23]. Some scoring systems have poor specificity in low-education or non-English speaking groups (see Appendices 21 and 22 in the supplementary data). In a study of Chinese older people, the legibility of numbers was left out of the scoring to help account for lack of formal education (see Appendix 23). Reported sensitivity is lowest for better educated English speakers [10]. Administration needs little language interpretation, and the test is highly acceptable to patients and assessors.

Mini-Cog [16]
Devised to improve the sensitivity of the CDT by adding a test of new learning: the names of three objects are presented before the clock drawing and recalled afterwards. If all three names are recalled, the test indicates no impairment; if no names are recalled the indication is impairment. Only if one or two names are recalled is the CDT scored and used to provide the test result, on binary scoring. The Mini-Cog is relatively new and un-tested, although it has been recommended by two recent overviews of screening tests [24, and also see Appendix 24]. The test authors found no evidence of language bias.

7-minute screening battery [17]
Another test incorporating the CDT: four sub-tests cover memory, verbal fluency, orientation to time and clock drawing. Orientation errors are weighted, and the CDT scored on seven attributes; an equation produces a total score from the subscale scores. The test is hampered by the need for special equipment and training, and no evidence was found on cross-cultural validity.

Time and Change Test (T&C) [18]
A ‘real-life’ assessment incorporating two activities: reading the time from a clock-face and making up one dollar from a selection of change. Both tasks must be completed correctly to score as non-impaired. An alternative scoring system takes into account the time to completion of each task. The tasks tap a wide range of cognitive functions such as orientation, concentration, conceptualization and visuospatial skills, and are said to be highly acceptable to older person and assessor. One study found no effect of age, education, disability or depression on scores [25]. Another found that while specificity was high and approximately equal across ethnic groups, sensitivity was better for black than white groups on the basic scoring: the alternative timed cut-points increased sensitivity at the expense of specificity [26].

Discussion and recommendations

The role of short screening tests in a clinical setting is to provide an initial rapid indication of possible cognitive impairment. Individuals should screen positive because they have cognitive impairment, not because the tasks or the way they are presented are inappropriate. It is clear that standard cognitive screening tests currently lead to substantial false positive results among older people in minority ethnic groups. Referral for further assessment is distressing for the individuals and wasteful of resources, although diagnosis on the basis of functional abilities, possibly based on an informant assessment, appears to be much less prone to educational or cultural bias (see Appendices 25 and 26).

Of the existing multi-item screening tests, used without adaptation, short tests perform at least as well as longer ones and are more consistent across cultural and educational groups [27]. The adaptation which is the least labour-intensive is to establish adjusted cut-points or norms. This has not been found universally successful in improving accuracy, and it fails to deal with the lack of acceptability to the older person of a testing process tailored to the majority population. Some progress has been made towards developing versions which take into account differences in education, lack of familiarity with English, lack of literacy in own first language, and cultural background. These versions need to be studied within their target populations to establish appropriate cut-points, since adaptation can radically change the performance of the test.

There are promising results on the ‘culture-free’ tests: they are seen as non-threatening, with less reliance on skills learned during formal education and greater comparability between minority groups. Little translation is needed for non-English speakers. The T&C test is highly acceptable but there are reported problems in achieving good sensitivity and specificity: it may be possible to develop the scoring to address this. The CDT has considerable cross-cultural evidence, and the more recent Mini-Cog is thought to improve the sensitivity of the CDT while maintaining its acceptability and lack of language bias.
bias. It is worthy of further evaluation among a range of cultural and language groups.

The advantages of culture-free screening methods over traditional tests are not limited to minority groups. They provide a less threatening situation in which older people can fully demonstrate their cognitive abilities, whatever their language, culture and education. There are parallels with the concept of inclusive design, which originated in their language, culture and education. There are parallels in minority ethnic groups could lead to better screening for all older people.

**Key points**

- Cognitive screening tests under-estimate the abilities of older people in BME groups.
- Contributing factors are education, language, literacy and culture-specific references.
- Translated and/or adapted versions, or adjusted cut-points, exist for several standard tests.
- There are promising results on ‘culture-free’ tests but more evidence is needed.

**Acknowledgements**

We are grateful for the help of Manjinder Sarai and Akhlak Rauf, Ambassador Training and Development Consultants in appraising this review.

**References**

C. Parker, I. Philp


Received 22 October 2003; accepted in revised form 11 February 2004