Cognitive testing of elderly Chinese people in Singapore: influence of education and age on normative scores

SURESH SAHADEVAN, NOELLYN J. L. TAN, TAHCHEW TAN\textsuperscript{1}, SHIRLEY TAN\textsuperscript{2}

Department of Geriatric Medicine, Tan Tock Seng Hospital, Moulmein Road, Singapore 1130
\textsuperscript{1}Department of Community, Occupational and Family Medicine, National University of Singapore
\textsuperscript{2}Psyhealth Centre, Gleneagles Medical Centre, Singapore

Correspondence to: S. Sahadevan. Fax: (+65) 252 7282

Abstract

Aim: to establish age- and education-adjusted mean psychometric test scores of cognitively normal elderly Chinese people in Singapore.

Method: 155 community-living, cognitively intact elderly Chinese subjects, stratified according to age and education, were selected from the membership lists of senior citizens' clubs and were administered a neuropsychological assessment battery. The mean test scores were then analysed for correlations with demographic characteristics.

Results: only education and age were independent variables; gender and language spoken were strongly correlated with education by the $\chi^2$ method ($P<10^{-5}$). Subsequent statistical analysis of education and age on the psychological test scores by the general linear model procedure enabled the formulation of a table of age- and education-adjusted normative cognitive scores.

Conclusion: the derivation of age- and education-adjusted normative scores of a selected neuropsychological assessment battery on elderly Chinese people enables a more accurate diagnosis of dementia to be made and may facilitate better international comparisons of this condition. As our subjects are representative of many others from China, these results may serve not only as a local, but also as a regional, reference source.

Keywords: Chinese, elderly subjects, normative data, psychometric testing

Introduction

Dementia affects 2-4% of elderly people (defined as being 65 years of age or older) in Singapore \cite{1,2}, where the size of the elderly population is expected to triple over the next three decades, reaching 20% by 2030 \cite{3}. By then the numbers with cognitive impairment are estimated to reach 20,000. Furthermore, epidemiological studies of the local prevalence rates \cite{1,2} may have only picked up cases of established dementia; if those in the earlier stages are also included, numbers will be far larger.

The diagnosis of dementia has been made easier by employing well-defined clinical criteria \cite{4-6} and by psychometric testing. Cultural, educational and ethnic factors must however be considered when drawing conclusions from psychological tests obtained in populations differing from the ones in which they were validated. This prompted our study. We have already described the methodology and presented the overall mean scores of the neuropsychological assessment battery \cite{7}; here, we focus upon the influence of the subjects' age and educational attainments on their test performances and describe the rationale of the methods employed to derive our table of normative values.

Method

Design

The main aim of this community study was to identify cognitively intact elderly subjects based upon predefined criteria and to apply selected psychological as well as cognitive screening tests to them. The master list of subjects was obtained from the membership records of seven randomly chosen senior citizens' clubs as well as the retirees' club; these clubs are affiliates...
of a statutory body and organize recreational and community activities for the well elderly of Singapore.

Singapore is a cosmopolitan country and its 1990 census shows the ethnic composition to be Chinese (78%), Malay (14%) and Indian (7%) [8]. Amongst those who were 50 years of age and above in 1990, 72.3% had no formal or an incomplete primary education. Women made up 49.6% of the population. To avoid the confounding influence of different cultures, only Chinese subjects of the major dialect groups (Mandarin, Hokkien, Teochew) were enrolled into the study. The subjects were stratified according to gender, age (60–64 years, 65–74 years and 75–84 years) and educational exposure (0 years of education, 1–6 years of education and >6 years of education).

From the master list we intended to study 180 elderly subjects from all the age-, education- and gender-specified categories mentioned above. A telephone interview was first conducted with the subjects, whereby their response to a 10-question elderly cognitive assessment questionnaire—a locally validated cognitive screening instrument [9]—was elicited, as well as information on medication, functional status and alcohol consumption. A separate telephone interview with the subject’s principal care-giver was also conducted to pick up features of dementia or depression. A home visit, with permission, was then made by the research psychologist to administer the neuropsychological assessment battery.

All tests and procedures were approved by the hospital’s ethics committee and were conducted with the informed consent of both subjects and principal care-givers.

Definitions
To identify cognitive normality, we used the following criteria: (i) scoring >7 on the elderly cognitive assessment questionnaire (earlier validation work [9] had shown that locally a score of 5 and below identified dementia); (ii) not having any of the features indicative of dementia as set out by the American Psychiatric Association’s Diagnostic and Statistical Manual IV criteria [4]; (iii) not having a history or current features of depression or other psychiatric conditions or stroke; (iv) not having a regular intake of sedatives or a history of heavy alcohol intake (defined as more than one bottle of spirits at each drinking session and/or a frequency of more than one drinking session per week); (v) not being housebound or dependent for any of the basic activities of daily living (the aim being to exclude those with serious physical disability).

We wished to ensure that those described as ‘normal’ were reasonably so, while at the same time avoiding the formation of a ‘super-normal’ group (whose performance scores could not be usefully extrapolated to ‘average’ patients).

Neuropsychological assessment battery
The two main groups of tests in the neuropsychological assessment were brief cognitive screening tests and the more detailed psychological tests, which probed different aspects of cognition.

Two cognitive screening instruments were analysed. The Abbreviated Mental Test [10] assesses memory, concentration and orientation. The Chinese Mini-Mental Status [11] also assesses these functions, along with language, calculations and visuo-spatial abilities.

The following eight psychological tests were selected because of their brevity (the entire battery was completed within 45 min for each subject) and ease of administration, as well as their ability to assess the characteristic impairments accompanying dementia in the domains of memory, language, visuo-spatial abilities and executive functioning:

1. Word list memory (immediate recall) [12]: a 10-word list was presented and free recall assessed across three trials with the words being presented in a different order for each trial.
2. Word list (delayed recall): delayed memory for the 10 words in the word list (immediate recall) test was assessed 10 min after the third trial on immediate memory.
3. Word list (recognition): recognition of the 10 words from the word list (immediate recall) test was tested when they were presented together with 10 distractor words.
4. Set test (animal category) [13]: this assesses language, verbal fluency and executive functioning by asking the subject to name as many animals as possible within 1 min.
5. Modified Boston naming test: 15 line drawings were chosen from the original 60 drawings in the Boston naming test [14] and were presented for subjects to name.
6. Constructional praxis [15]: visuo-spatial skills were assessed by presenting four line drawings of generic figures (a circle, intersecting rectangles, a diamond and a cube) for the subject to copy.
7. Block design: this performance subtest of the Wechsler Adult Intelligence Scale—Revised (WAIS-R) [16] measured visuo-spatial organisation by asking the subject to construct replicas of designs using red and white blocks.
8. Object assembly: in this performance subtest of the WAIS-R the subject was presented with cut-up cardboard figures of familiar objects to assemble.

The neuropsychological assessment battery was administered by a research psychologist who underwent training in interview techniques and procedures as well as familiarization with the various languages and dialects.

All the test materials and the instructions were translated into Mandarin and the two main dialects in...
Cognitive testing of elderly Chinese

Singapore—Hokkien and Teochew. Back-translation was performed to ensure accuracy and appropriateness in the phrasing of the items. Adaptations were made to the Chinese Mini-Mental Status, word list (immediate recall) and the modified Boston naming test so as to make them more locally suitable.

Statistical analysis

Frequencies and $\chi^2$ tests were carried out using SPSS for Windows (Statistical Package for Social Sciences, Version 6.0, 1993). The $\chi^2$ test was used to look at the associations between the demographic variables to establish which were independently exerting a significant effect on the test performances.

Our observational study had produced unbalanced data (different number of observations in each cell of the data table) and this precluded the use of analysis of variance (ANOVA) [17]. Therefore, to study in greater detail the relationship between the mean test scores and the various demographic variables, the type III sum of squares of the general linear model (GLM) in the SAS proprietary software (Release 6.04, 1987) was used.

Results

Subjects

From the original master list, 155 subjects met the entry criteria and gave consent to participate in the study. Their stratified distribution according to age, education and gender is shown in Table 1. Fifty-one percent were women and approximately half of the subjects had more than 6 years of education (47%). The subjects were generally spread across the three age categories: 35% aged 60–64 years, 39% aged 65–74 years and 26% aged 75–84 years. The two most common languages used for testing the subjects were English (39%) and Mandarin (41%); Hokkien and Teochew dialects were used in only 15 and 5% of cases respectively.

Influence of variables

Four demographic variables (age, education, gender and language tested) were assessed to analyse their influence on the mean scores of the neuropsychological assessment battery. We anticipated that education would have a marked effect; the associations between education and the remaining demographic variables were thus looked at by the $\chi^2$ method. It was then found that gender and language were strongly associated with education ($P < 10^{-5}$) while age was not ($P = 0.63$). Hence age and education remained as the independent variables requiring further analysis of their effects on the test performances and this was accomplished by the GLM procedure.

Despite our original intentions to age- and sex-match the subjects, the observational design of this study could not totally avoid discrepancies in the various stratified cells (see Table 1). There were no male subjects between 65 and 74 who were uneducated, but gender was not independently exerting any influence on the psychological tests scores so we pooled the numbers from both genders for the GLM analysis and the results are shown Table 2.

For all the tests in the neuropsychological assessment battery, the main effect of education was statistically significant ($P < 0.01$). The main effect of age was only significant for the Chinese Mini-Mental Status, the set test (animal category), the three word-list memory tests (immediate recall, delayed recall and recognition) and the modified Boston naming test. In addition, the presence of interaction effects between age and education were found in the abbreviated mental test and the word-list memory (delayed recall) test.

The GLM procedure also enabled a post hoc analysis to determine, in those instances where age and education were exerting an influence, which of the categories of these independent variables were important. The mean scores across all three categories of education were statistically significant but those for the age variable were not, allowing us, for several tests, to...

Table 1. Number of subjects by education, age and gender

<table>
<thead>
<tr>
<th>Education level (years)</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60–64</td>
<td>65–74</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1–6</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>&gt;6</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>31</td>
</tr>
</tbody>
</table>
Table 2. General linear model results: education by age—statistically significant effects

<table>
<thead>
<tr>
<th>Test</th>
<th>Main effect (P &lt;)</th>
<th>Interaction effect (P &lt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviated Mental Test</td>
<td>0.001</td>
<td>0.013</td>
</tr>
<tr>
<td>Chinese Mini-Mental Status</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>Set test (animal category)</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>Word list memory (immediate recall)</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>Word list memory (delayed recall)</td>
<td>0.008</td>
<td>0.003</td>
</tr>
<tr>
<td>Word list (recognition)</td>
<td>0.010</td>
<td>-</td>
</tr>
<tr>
<td>Modified Boston naming test</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>WAIS-R block design</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>WAIS-R object assembly</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>Constructional praxis</td>
<td>0.001</td>
<td>-</td>
</tr>
</tbody>
</table>

WAIS-R, Wechsler Adult Intelligence Scale—Revised.

collapse and thus reduce the number of age categories. These results are shown in Table 3.

For the constructional praxis test and the WAIS-R block design and object assembly tests where the main effect of age was not statistically significant, the age categories were collapsed and the mean scores were reflected only across the categories of educational attainment. For the Chinese Mini-Mental Status, word list (recognition) and the modified Boston naming test, the post hoc analysis allowed for the two age categories of 60-64 years and 65-74 years to be condensed. However, for the set test (animal category) and the word-list memory (immediate recall) test, only the differences between the youngest and oldest age

Table 3. Mean scores and standard deviations by age and education

<table>
<thead>
<tr>
<th>Test</th>
<th>Age (years)</th>
<th>Education (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviated Mental Test*</td>
<td>60-64</td>
<td>8.6 (1.2)</td>
</tr>
<tr>
<td></td>
<td>65-74</td>
<td>7.5 (1.1)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>7.6 (1.3)</td>
</tr>
<tr>
<td>Chinese Mini-Mental Status</td>
<td>60-74</td>
<td>22.9 (2.4)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>20.7 (2.5)</td>
</tr>
<tr>
<td>Animal category</td>
<td>60-74</td>
<td>11.6 (3.1)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>9.6 (2.0)</td>
</tr>
<tr>
<td>Immediate memory</td>
<td>60-74</td>
<td>15.3 (4.3)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>12.9 (3.9)</td>
</tr>
<tr>
<td>Delayed memory*</td>
<td>60-64</td>
<td>5.3 (1.7)</td>
</tr>
<tr>
<td></td>
<td>65-74</td>
<td>4.1 (2.0)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>4.4 (1.1)</td>
</tr>
<tr>
<td>Recognition</td>
<td>60-74</td>
<td>7.9 (1.3)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>6.3 (2.7)</td>
</tr>
<tr>
<td>Modified Boston naming test</td>
<td>60-74</td>
<td>13.1 (1.4)</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>12.1 (1.5)</td>
</tr>
<tr>
<td>Block design</td>
<td>60-84</td>
<td>9.5 (3.9)</td>
</tr>
<tr>
<td>Object assembly</td>
<td>60-84</td>
<td>14.6 (6.8)</td>
</tr>
<tr>
<td>Constructional praxis</td>
<td>60-84</td>
<td>1.6 (0.8)</td>
</tr>
</tbody>
</table>

*Significant interaction effect between age and education.
groups yielded statistical significance and thus the collapsing of the age categories could be between
either 60–64 years and 65–74 years or 65–74 years and 75–84 years. To maintain a consistent mode of
presentation, we decided to merge the age categories of 60–64 years and 65–74 years. Where interaction
effects were observed, however, the original age and education categories were retained.

Discussion
This study shows the primacy of the roles of education and age in cognitive testing: the interpretation of
test scores can only be meaningfully accomplished after these variables have been considered. The
language and gender of our subjects were strongly associated with their educational attainments: the
more highly educated tended to be male and English-
or Mandarin-speaking. Differences in the mean test
scores for language and sex were thus largely, if not totally, explained by this relationship. We believe that
this applies to other elderly Chinese subjects in Singapore.

Our results allow a more accurate identification of
cognitive impairment in this part of the world. Many
elderly Chinese people in Singapore migrated here
from China in their younger days, at the same time as
others left to settle in other parts of the South-east
Asia and the Far East. Culturally, as well as education-
ally, these migrant workers have many similarities. We
believe that our results can serve not only as local, but
also as a regional, reference source and pave the way
for epidemiological and clinical comparisons of
patients with dementia from the East and the West.

Conclusion
This community study of cognitively intact elderly
Chinese people in Singapore establishes the need to
define the subject’s age as well as past educational
exposure before interpreting his or her cognitive test
performance. The age- and education-adjusted norma-
tive scores can enable the diagnosis of dementia to be
made more accurately and this can promote further research work of this devastating condition in this part
of the world.

Acknowledgements
This study was supported by a grant from Laboratoires
Servier. We are also grateful for the co-operation of the
retirees’ club and the senior citizens’ clubs of
Alexandra Community Centre, Ayer Rajah Zone 1
Residents’ Committee, Bukit Batok Community Club,
Cheng San Community Club, Kaki Bukit Zone E
Residents’ Committee, Kampong Kembangan Com-
munity Club and Rochor Community Centre. Ng Tze
Pin and Goh Bee Choo from the National University of
Singapore and Ayliana Phe from Tan Tock Seng
Hospital also provided helpful comments on statistical
analyses.

Key points
- Psychometric tests can be useful in the diagnosis of
dementia, especially in the early stages.
- Psychometric test performances are significantly
influenced by variables such as education, age and
cultural background.
- Cognitively intact elderly Chinese subjects were
administered a selected neuropsychological assess-
ment battery and their education- and age-adjusted
normative scores are presented.

References
2. Kua EH. Dementia in elderly Malays—preliminary findings
3. Choo PWJ, Sahadevan S, Chee YC, Jayaratnam FJ. Health
care services for the elderly—a Singapore perspective.
4. American Psychiatric Association, Committee on Nomen-
clature and Statistics. Diagnostic and Statistical Manual of
Mental Disorders, 4th ed. Washington DC: American Psychi-
5. McKhann G, Drachman D, Folstein M, Katzman R, Price D,
Stadlan EM. Clinical diagnosis of Alzheimer’s disease: report
of the NINCDS–ADRDA work group under the auspices of the
Department of Health and Human Services Task Force on
Validation of clinical diagnostic criteria for Alzheimer’s
7. Sahadevan S, Tan NJL, Tan TC, Tan S. Cognitive testing of
elderly Chinese from selected community clubs in Singapore.
8. Lau KE. Singapore Census of Population 1990: literacy,
languages spoken and education. Singapore: Department of
9. Kua EH, Ko SM. A questionnaire to screen for cognitive
impairment among elderly people in developing countries.
10. Hodkinson HM. Evaluation of a mental test score for
assessment of mental impairment in the elderly. Age Ageing
version of the Mini-Mental State Examination: impact of
illiteracy in a Shanghai dementia survey. J Clin Epidemiol
12. Morris JC, Heyman A, Mohs RC et al. The consortium to


Received 6 January 1997