A validation study of the Chinese-Cantonese Addenbrooke’s Cognitive Examination Revised (C-ACER)

LL Wong
CC Chan
JL Leung
CY Yung
KK Wu
SY Cheung
CLM Lam

1Department of Psychiatry,
2Department of Medicine and Geriatrics, United Christian Hospital, Kwan Tong, Kowloon, Hong Kong;
3Clinical Psychology Services, Kwai Chung Hospital, Kwai Chung, New Territories, Hong Kong;
4Hong Kong Federation of Youth Groups, Department of Clinical Psychology, Tseung Kwan O Hospital, Tseung Kwan O, New Territories, Hong Kong

Correspondence: CY Yung
Department of Medicine and Geriatrics, United Christian Hospital, 130 Hip Wo Street, Kwun Tong, Kowloon, Hong Kong
Tel +852 3513 4828
Fax +852 3513 5781
Email yungcy@ha.org.hk

Background: There is no valid instrument for multidomain cognitive assessment to aid the detection of mild cognitive impairment (MCI) and mild dementia in Hong Kong. This study aimed to validate the Cantonese Addenbrooke’s Cognitive Examination Revised (C-ACER) in the identification of MCI and dementia.

Methods: 147 participants (Dementia, n = 54; MCI, n = 50; controls, n = 43) aged 60 or above were assessed by a psychiatrist using C-ACER. The C-ACER scores were validated against the expert diagnosis according to DSM-IV criteria for dementia and Petersen criteria for MCI. Statistical analysis was performed using the receiver operating characteristic method and regression analyses.

Results: The optimal cut-off score for the C-ACER to differentiate MCI from normal controls was 79/80, giving the sensitivity of 0.74, specificity of 0.84 and area under curve (AUC) of 0.84. At the optimal cut-off of 73/74, C-ACER had satisfactory sensitivity (0.93), specificity (0.95) and AUC (0.98) to identify dementia from controls. Performance of C-ACER, as reflected by AUC, was not affected after adjustment of the effect of education level. Total C-ACER scores were significantly correlated with scores of global deterioration scale (Spearman’s rho = -0.73, P < 0.01).

Conclusion: C-ACER is a sensitive and specific bedside test to assess a broad spectrum of cognitive abilities, and to detect MCI and dementia of different severity. It can be used and interpreted with ease, without the need to adjust for education level in persons aged 60 or above.

Keywords: dementia, mild cognitive impairment, multidomain, bedside, assessment, Hong Kong

Introduction
Dementia is conventionally a clinical diagnosis, characterized by progressive cognitive decline and a gradual loss of ability to carry out activities of daily living. The brain pathology that will eventually lead to symptoms of dementia develops well in advance of the time at which the person’s symptoms would first have been noticed. Mild cognitive impairment (MCI) is a transitional phase between healthy ageing and dementia, carrying an annual conversion rate to dementia of 10%–15% in clinic-based studies. As there is growing evidence supporting the cost-effectiveness and benefits of early recognition and intervention of dementia in patients and their carers, screening for MCI and early dementia is therefore important in clinical practice.

The Mini-Mental State Examination (MMSE) is the most widely used assessment tool for global cognitive ability in many countries. However, the MMSE has several well-known drawbacks, including likelihood of ceiling effects due to low level of task difficulty, absence of tasks measuring frontal executive function, limited range of...
possible scores for individual items, low sensitivity for the
detection of patients with MCI, and its diagnostic accuracy
is age and education level dependent. The Cantonese version
of MMSE (CMMSE) has been validated and adopted as the
standard measure to screen for dementia in Hong Kong for
two decades. As duly addressed by the authors, the results
might not be generalized to persons with MCI or early stage
dementia because the study sample consisted only of
moderate-to-severe dementia (CMMSE mean score = 9)
and normal controls (CMMSE mean score = 25). Moreover,
the cut-off scores for different education level were only
arbitrarily set due to the insufficient number of subjects with
higher education level for analysis.

The Addenbrooke’s Cognitive Examination Revised
(ACER) was developed as a brief bedside cognitive assess-
ment instrument. It can be administered without trained
personnel, and be completed in approximately 15 minutes.
It consists of five cognitive domains: attention/orientation
(18 points), episodic and semantic memory (26 points), verbal
fluency (14 points), language (26 points) and visuospatial
ability (16 points), summing up to a maximum total score
of 100. The original English ACER has shown very good
reliability (alpha coefficient = 0.8), and high sensitivity (0.94)
and specificity (0.89) at the cut-off of 88 for identifying
dementia. It has also been translated and validated in different
countries. ACER is consistently shown to be capable of
providing information on a range of cognitive domains and to
differentiate individuals with or without dementia. However,
the psychometric properties of ACER to distinguish MCI
from normal ageing were not well demonstrated. This study
aimed to validate the Cantonese Addenbrooke’s Cognitive
Examination Revised (C-ACER) in the detection of MCI
and dementia in people aged 60 or above, and to compare
its diagnostic accuracy with that of the CMMSE.

Methods

Participants and procedure

This was a cross-sectional validation study conducted at the
cognitive clinic of the Medicine and Geriatric Department
and the memory clinic of the Department of Psychiatry in
a general hospital (United Christian Hospital, Kwun Tong,
Kowloon, Hong Kong). Both clinics provide a specialist
service to those presenting primarily with memory problems.
Potential subjects were recruited from the two clinics during the
study period from August 2011 to March 2012, and that
included patients who were 60 years or older, Cantonese
speaking with reasonable vision and hearing ability, and the
ability to communicate verbally. Patients who had a history,
as documented in medical records, of major depression,
schizophrenia, epilepsy, significant head injury, substance
abuse and alcoholism were excluded. Healthy participants
were recruited among the patients’ spouses or those who
attended an elderly vaccination program in the hospital. They
shared the same inclusion and exclusion criteria.

We examined a total of 147 participants, divided into
three groups: dementia (n = 54), mild cognitive impairment
(MCI, n = 50), and normal control (NC, n = 43). All patient
participants were examined by either the research geriatrician
(CC or psychogeriatrician (JLL) with a standard assess-
ment battery, which included an interview with patients and
informants, general and neurological examinations, mental
state examination, laboratory tests (complete blood count,
liver and renal function tests, thyroid function tests, serum
vitamin B12 and folate level, and syphilis serology) and
brain imaging (MRI or CT). Diagnoses of dementia and
MCI were made according to DSM-IV criteria for dementia
and Peterson’s criteria for MCI respectively. Each patient
was staged functionally by the Global Deterioration Scale
(GDS). Control subjects were assessed with the CMMSE,
GDS and brief mental state examination by one of the
researchers (CYY, CC or JLL), and participants in GDS
stage 2 or below were confirmed to be normal controls. An
independent research psychiatrist (LLW), who was blind to
patients’ clinical findings, would then administer C-ACER
to all participants within one week after the standard clinical
assessment and obtain their demographic data (age, year of
education). Written informed consent was obtained from all
participants and this study was approved by the Hong Kong
Hospital Authority Ethical Committee.

Sample size calculation

Based on the estimated prevalence rates of dementia and
MCI at 12.5% and 8.5% respectively, it was determined that
the overall sample size was 138 (NC:MCI:dementia = 1:1:1)
with n = 46 in each group, by using a sample size calculator
exe. These sample sizes would give a power of 0.8 and a Type
I error of 0.05.

Measurements

C-ACER

The original English ACER was translated into
Chinese-Cantonese version by an expert panel with cultural
and linguistic adaptation specific to the Hong Kong population.
Modifications were made in the following subtests: name and
address recall and recognition, semantic memory, verbal
fluency, and reading and perceptual ability (fragmented
letters). Based on the original criteria, a common Chinese
name and local address (district, place, street and number) were constructed in the anterograde memory part. In the semantic memory subtest, the four questions were replaced by “the current Chief Executive of Hong Kong SAR”, “the year of Hong Kong’ reversion to China”, “the current president of the People’s Republic of China”, and “the first president of the People’s Republic of China”. In the verbal fluency subtest, since there is no equivalent substitution for letter fluency in Cantonese, the subtest was replaced by a second category fluency of naming vegetables in one minute. In the first part of the repetition subtest, the words “雙休日 (weekend)”, “顯影劑 (contrast medium)”, “搖滾樂隊 (rock and roll band)” and “掩耳盜鈴 (hiding one’s head in the sand)” were selected according to the criteria used in the original English ACER: multisyllabic words that are perceived as relatively difficult and rare yet known to most speakers. The relative frequencies of the four Chinese multisyllabic terms were chosen with comparable level of relative frequencies as that of the original English words in American word frequencies. For the second part of the repetition subtest, we substituted with short sentences used in MMSE in Taiwan or the People’s Republic of China. In the comprehension subtest, the instruction “point to the one which is a marsupial” was replaced by “point to the one which is a reptile” (the correct answer is “crocodile”) as the Chinese translation of marsupial (有袋類動物) and kangaroo (袋鼠) are quite similar. In the reading subtest, as the words used in this reading task have to be “exception words” or “irregular words”, ie, words which do not obey the normal sound-to-print rules of English, five Chinese “exceptional” characters, (氵, 纟, 息, 資, 笛) were selected. In the original version of the perceptual ability subtest, four incomplete English alphabets were used to assess perceptive visual agnosia as in the Visual Object and Space Perception Battery (VOSP). Therefore, we picked up four Chinese characters (also Chinese radicals) instead of four alphabets, and converted them into incomplete words by degrading them by 70% as in the development of VOSP. Finally, in the recognition subtest, two sets of name and local addresses (district, place, street and number) spreading across the territory were composed as recognition stimuli.

Global deterioration scale (GDS)
The GDS was developed to provide caregivers with an overview of the stages of functional decline for those suffering from a primary degenerative dementia. The GDS is comprised of 7 stages defined by a set of clinical characteristics. The higher the stage, the more severe the cognitive decline: Stage 1 = no cognitive decline; Stage 2 = subjective cognitive impairment; 3 = mild cognitive impairment; Stage 4 = mild dementia; Stage 5 = moderate dementia; Stage 6 = moderately severe dementia; Stage 7 = severe dementia. Inter-rater reliability for the GDS was found to be high (0.87–0.97) and scores of GDS showed high correlation with MMSE scores. We adopted GDS in this study to stage the cognitive level from a functional perspective.

Statistical analysis
All statistical analyses were performed using the Statistical Package for Social Sciences (SPSS) version 19 (IBM Corporation, Armonk, NY, USA), and a P value of <0.05 was considered statistically significant. For comparisons of continuous variables (age, years of education, total and sub-scores of C-ACER and score of CMMSE) among groups, one-way ANOVA was used. Chi-square tests were performed for categorical data. If differences attained statistical significance, multivariate linear regression analysis was carried out to examine the influence of demographic variables on the participants’ performance in the total and subdomain C-ACER scores. Logistic regression analysis was performed to assess the extent to which C-ACER and CMMSE scores could predict the clinical diagnosis, with age and years of education covaried. A threshold analysis and receiver operating characteristic (ROC) analysis were adopted to measure the performance of C-ACER. The performance of C-ACER was compared with that of CMMSE by comparing their AUCs, using a calculator available at http://www.medcalc.org/download/medcalcsetup.exe. The optimal cutoff was identified by calculating the Youden’s Index. Spearman’s correlation was used to investigate the convergent validity between total C-ACER scores and GDS scores.

Results
A total of 147 participants were recruited (43 normal control, 50 MCI and 54 dementia) and none of the eligible subjects refused to participate in the study.

Clinical characteristics
Demographic data, CMMSE and C-ACER total and component scores of the participants within the three groups are shown in Table 1. There was no significant difference in the years of education among groups. Subjects in the normal control group (NC) were significantly younger than those in the patient groups (NC = 72.8 years, MCI = 76.9 years, dementia = 79.2 years, CI = 78.1 years; P < 0.01). Post-hoc analysis showed that there were significant differences between the three groups (NC, MCI and dementia) regarding their performance in (1) CMMSE, (2) C-ACER total, (3) orientation/attention subtest, (4) memory subtest,
Table 1: Demographic characteristics, mean scores ± SD of MMSE and ACER total and subdomain scores of study sample (ANOVA, Post Hoc-Bonferroni)

<table>
<thead>
<tr>
<th></th>
<th>NC (n = 43)</th>
<th>MCI (n = 50)</th>
<th>Dementia (n = 54)</th>
<th>CI (104)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types, % (n)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD/amnestic</td>
<td>–</td>
<td>70% (35)</td>
<td>59.2% (32)</td>
<td>–</td>
</tr>
<tr>
<td>Mixed</td>
<td>–</td>
<td>20% (10)</td>
<td>33.3% (18)</td>
<td>–</td>
</tr>
<tr>
<td>Vascular</td>
<td>–</td>
<td>8% (4)</td>
<td>7.4% (4)</td>
<td>–</td>
</tr>
<tr>
<td>Age</td>
<td>72.8 ± 7.5</td>
<td>76.9 ± 7.3a</td>
<td>79.2 ± 6.0a</td>
<td>78.1 ± 6.7a</td>
</tr>
<tr>
<td>Gender (Female ratio, %)</td>
<td>29 (67.4%)</td>
<td>29 (58%)</td>
<td>35 (64.8%)</td>
<td>64 (61.5%)</td>
</tr>
<tr>
<td>Education (yr)</td>
<td>5.6 ± 4.3</td>
<td>4.2 ± 4.2</td>
<td>3.7 ± 4.2</td>
<td>3.9 ± 4.2a</td>
</tr>
<tr>
<td>MMSE</td>
<td>27.3 ± 2.6a</td>
<td>21.6 ± 5.0a</td>
<td>16.4 ± 5.4a</td>
<td>18.9 ± 5.8a</td>
</tr>
<tr>
<td>ACER</td>
<td>86.4 ± 8.9a</td>
<td>68.2 ± 15.7a</td>
<td>50.8 ± 15.4a</td>
<td>59.1 ± 17.8a</td>
</tr>
<tr>
<td>Orientation and attention</td>
<td>16.5 ± 2.0a</td>
<td>12.7 ± 3.6a</td>
<td>9.2 ± 3.8a</td>
<td>10.9 ± 4.0a</td>
</tr>
<tr>
<td>Memory</td>
<td>21.5 ± 3.3a</td>
<td>14.2 ± 5.2a</td>
<td>8.8 ± 3.6a</td>
<td>11.4 ± 5.2a</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>13.7 ± 0.8a</td>
<td>11.5 ± 3.1a</td>
<td>9.0 ± 3.7a</td>
<td>10.2 ± 3.6a</td>
</tr>
<tr>
<td>Language</td>
<td>21.0 ± 3.4a</td>
<td>17.6 ± 4.9a</td>
<td>14.2 ± 4.9a</td>
<td>15.8 ± 5.1a</td>
</tr>
<tr>
<td>Visuospatial ability</td>
<td>13.8 ± 2.3a</td>
<td>12.3 ± 3.1a</td>
<td>9.6 ± 3.5</td>
<td>10.9 ± 3.5a</td>
</tr>
</tbody>
</table>

Notes: Significant differences from NC, P < 0.05; *significant differences among NC, MCI and dementia groups, P < 0.05; †significant differences from dementia, P < 0.05; CI was only compared with NC.

Abbreviations: SD, standard deviation; NC, normal controls; CI, cognitively-impaired, includes “MCI” and “dementia”; AD, Alzheimer’s disease; MMSE, Mini-Mental State Examination; ACER, Addenbrooke’s Cognitive Examination Revised; MCI, mild cognitive impairment; ANOVA, analysis of variance.

(5) verbal fluency subtest, and (6) language subtest (P < 0.01). Concerning the visuospatial ability subtest, subjects with dementia scored significantly lower than those in the NC and MCI groups (P < 0.01), however, such difference was not observed between NC group and MCI group (P > 0.05).

In multivariate linear regression analysis (Table 2), with C-ACER total scores as the dependent factor while the age and years of education and the clinical diagnosis (NC vs MCI vs dementia) as the independent variables, only the latter two (ie, years of education and clinical diagnosis) were found to be affecting the C-ACER total scores significantly (P < 0.01). Same significance was demonstrated in four out of five C-ACER subdomains including orientation/attention, memory, verbal fluency and language (P < 0.01). Regarding visuospatial ability, the clinical diagnosis was shown to be the single significant factor.

Psychometric properties of C-ACER

In the first and second threshold analyses, the normal control (NC) was compared with the MCI group and dementia group respectively. In the third threshold analysis, patients with MCI or dementia were categorized in the “cognitively-impaired” group (CI, n = 104) for comparison against NC.

Table 3 shows the sensitivity, specificity and the areas under the ROC curve (AUC) at the optimal cut-off scores of C-ACER and CMMSE. The optimal cut-off was determined as the score where Youden’s index was maximized. The AUCs did not have significant differences between C-ACER and CMMSE regarding all three threshold analyses (P > 0.05). No differences in AUC were identified between ACER after adjustment of education level.

With regard to C-ACER, AUCs were 0.84, 0.98 and 0.91 for “NC vs MCI”, “NC vs dementia” and “NC vs CI” respectively. The optimal cut-off scores to distinguish NC

Table 2: Effect of education level, age and diagnostic groups (NC, MCI and dementia) on ACER total scores (multivariable linear regression)

<table>
<thead>
<tr>
<th></th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficients</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Standard error (β)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>83.54</td>
<td>12.31</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Years of education</td>
<td>1.40</td>
<td>0.26</td>
<td>0.30</td>
<td>5.37</td>
</tr>
<tr>
<td>Age</td>
<td>–0.07</td>
<td>0.16</td>
<td>–0.03</td>
<td>–0.43</td>
</tr>
<tr>
<td>MCI (reference)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>–15.95</td>
<td>2.70</td>
<td>–0.38</td>
<td>–5.90</td>
</tr>
<tr>
<td>Dementia</td>
<td>–32.52</td>
<td>2.76</td>
<td>–0.79</td>
<td>–11.78</td>
</tr>
</tbody>
</table>

Note: *P < 0.05.

Abbreviations: NC, Normal controls; MCI, mild cognitive impairment; ACER, Addenbrooke’s Cognitive Examination Revised.
from MCI and NC from the cognitively-impaired group were the same at 79/80, giving a sensitivity of 0.74 (NC vs MCI) and 0.88 (NC vs CI), and the same specificity of 0.84. At the optimal cutoff score of 73/74, ACER had high sensitivity (0.93) and specificity (0.95) to differentiate NC from dementia.

Logistic regression analysis reveals that (1) none of the covariates (age, years of education, ACER total and CMMSE scores; P > 0.05) contributed significantly to the differentiation between NC and MCI; (2) C-ACER total scores (regression coefficient = −0.29, P = 0.03) and years of education (regression coefficient = 0.59, P = 0.02) were independent factors to discriminate between NC and dementia while; (3) C-ACER total score (regression coefficient = −0.10, P = 0.04) was the single significant factor to differentiate NC from CI, independent of other covariates.

Reliability and convergent validity of ACER
The test-retest and inter-rater reliability of ACER, using Cronbach’s alpha coefficient, were both 1.00. The total C-ACER scores were significantly correlated with the scores of global deterioration scale (GDS, Spearman’s rho = −0.73, P < 0.01) and CMMSE (Spearman’s rho = 0.94, P < 0.01).

Discussion
This study verifies that the Chinese-Cantonese ACER has a high diagnostic accuracy in the detection of dementia (sensitivity = 0.93, specificity = 0.95) and is reasonably good to identify MCI (sensitivity = 0.74, specificity = 0.84). Besides, it has good convergent validity with CMMSE and GDS in people aged 60 or above.

In concurrence with the findings in the study by Alexopoulos et al., our study shows that C-ACER was not superior to the CMMSE in identifying MCI (AUC of C-ACER = 0.84, CMMSE = 0.85) and dementia (AUC of C-ACER = 0.98, CMMSE = 0.98). When MCI and AD were categorized in the “cognitively-impaired” group (CI), C-ACER and CMMSE had comparable performance (AUC of C-ACER = 0.91, CMMSE = 0.92; P > 0.05) to differentiate CI from normal subjects. The clinical usefulness of C-ACER, which takes a longer time than CMMSE to administer, is therefore not supported by the threshold analysis. However, such results have to be interpreted with caution and findings derived from logistic regression analysis should also be taken into consideration: (1) C-ACER total scores (P = 0.03) but not CMMSE scores (P = 0.15) were shown to be a significant contributor to the discrimination between normal control and dementia, and (2) C-ACER total scores (P = 0.04) stood out to be the single factor significantly contributing to the detection of patients with cognitive impairment (either having MCI or dementia), independent of CMMSE scores and education level. In addition, with a closer examination of the optimal cut-off scores, C-ACER had the same cut-off score (79/80) for the detection of MCI and CI, which was distinct from that for dementia (73/74). In contrast, the optimal cut-off scores of MMSE to discriminate dementia (25/26), MCI (26/27) and CI (25/26) from normal control were very close to each other or the same. In real clinical practice, we hope to identify the earliest stage of cognitive impairment so to provide intervention in a timely fashion. In this regard, C-ACER is shown to be more clinically useful than CMMSE to detect cognitive impairment by having a larger effect size.

The effect of age on ACER total and subdomain scores was shown to be inconsistent. In the original study by Mioshi et al., age had relatively little effect on ACER scores. On the other hand, in the validation study of the Japanese version of ACER, age had a significant effect on the total score (P = 0.01) and memory subscale (P = 0.01). Our study accords with the finding that age is not a significant factor affecting C-ACER total and subdomain scores according to linear regression analysis (P > 0.05). This may be due to
the relatively narrow range of age of the participants in the study, thus limiting the power to detect, if any, significant effect of age.

An unexpected finding in this study is the absence of significant effect of education level on the performance of C-ACER in the detection of dementia, MCI and CI. In logistic regression analysis, neither years of education nor age is an independent contributor to the discrimination between CI and normal control. We, thus, propose to adopt the cut-off score of C-ACER total at 79/80 (sensitivity = 0.88, specificity = 0.84) for the detection of cognitive impairment in persons aged 60 or above, regardless of their years of education.

This study demonstrates a good correlation between C-ACER total scores and GDS scores ($P < 0.01$), which indicates that C-ACER is a potential tool for assessing the severity of dementia and to monitor the disease progression.

Our study has two limitations. Firstly, the majority of our patient subjects received diagnosis of AD, mixed type dementia or amnesic MCI, reflecting a recruitment bias of the cognitive ability. Secondly, the predictive values could not be ascertained in this study because the patient group and normal controls were recruited independently and not from a designated population, leaving the true prevalence unknown.

In conclusion, the Chinese-Cantonese ACER (C-ACER) is a sensitive, specific and reliable bedside test to assess a broad spectrum of cognitive abilities, and to detect MCI and dementia of different severity. It can be used and interpreted with ease, without adjustment for age and education level in persons aged 60 or above.

**Authors’ roles**

LL Wong, CC Chan and JL Leung were responsible for designing the study, data collection, data analysis and writing the paper. CY Yung was responsible for designing the study, data collection and writing the paper. KK Wu, SYY Cheung and CLM Lam were responsible for translation of ACER into Chinese-Cantonese, conception and design, drafting and revising the article. All the authors approved the final version.

**Acknowledgments**

The authors thank Prof John Hodges for his permission to use and translate the ACER; Ms Mandy Yu for statistical advice; Dr Oliver Chan for proof-reading; and all participants for their kind support.

**Disclosure**

The authors declare no conflicts of interest in this work.

**References**
