Validation of a new mass screening tool for cognitive impairment: Cognitive Assessment for Dementia, iPad version

Keiichi Onoda1
Tsuyoshi Hamano2
Yoko Nabika1
Atsuo Aoyama1
Hiroyuki Takayoshi1
Tomonori Nakagawa1
Masaki Ishihara1
Shingo Mitaki1
Takuya Yamaguchi1
Hiroaki Oguro1
Kuninori Shiwaku3
Shuhei Yamaguchi1

1Department of Neurology, 2Center for Community-Based Health Research and Education, Shimane University, Izumo, 3Shimane University, Matsue, Shimane, Japan

Background: We have developed a new screening test for dementia that runs on an iPad and can be used for mass screening, known as the Cognitive Assessment for Dementia, iPad version (CADi). The CADi consists of items involving immediate recognition memory for three words, semantic memory, categorization of six objects, subtraction, backward repetition of digits, cube rotation, pyramid rotation, trail making A, trail making B, and delayed recognition memory for three words. The present study examined the reliability and validity of the CADi.

Methods: CADi evaluations were conducted for patients with dementia, healthy subjects selected from a brain checkup system, and community-dwelling elderly people participating in health checkups.

Results: CADi scores were lower for dementia patients than for healthy elderly individuals and correlated significantly with Mini-Mental State Examination scores. Cronbach’s alpha values for the CADi were acceptable (over 0.7), and test-retest reliability was confirmed via a significant correlation between scores separated by a one-year interval.

Conclusion: These results suggest that the CADi is a useful tool for mass screening of dementia in Japanese populations.

Keywords: dementia, mass screening, early detection, iPad

Introduction
Dementia has been calculated to contribute 11.2% of years lived with disability in people aged 60 years and over, and the number of dementia patients worldwide is estimated at 24.3 million people, with 4.6 million new cases of dementia reported every year.1 Recent drug therapies can slow rates of cognitive decline and delay emergence and progression of behavioral and psychological symptoms of dementia. These therapies have been most successful when they have been initiated appropriately early and maintained over time.2 Therefore, early detection is crucial as a countermeasure against dementia. Furthermore, a screening test that can be used for mass screening in community health checks is urgently needed.

Many screening tests have been proposed for assessing the severity of dementia. For instance, the Mini-Mental State Examination (MMSE) is one of the most widely used, and is accepted not only for clinical use but also for use in epidemiological surveys. The MMSE has good sensitivity and specificity for detection of dementia.3 However, existing tests such as these are not necessarily appropriate for mass screening. This is because these tests require trained examiners, and their administration is time-consuming. Mass screening of dementia requires a short, simple test that does not necessitate use of expert examiners. We developed a new screening test that can be run on an iPad.
Using the iPad for this type of mass screening has several merits. First, the test is run on a tablet computer, which is operated primarily by touching the screen. Many elderly people are accustomed to touch screen technology through operation of automated teller machines, and manipulation of this type of screen may be of interest to them. Thus, the elderly people who would receive this type of screening can see the iPad as a nonthreatening and even interesting device.

Second, the iPad can easily record an individual’s reaction times to a series of items; in contrast, it is difficult to record such reaction times accurately during conventional face-to-face examinations. Third, the iPad can easily manage and store data without manual operation. Fourth, although the initial purchase of an iPad represents a significant startup cost, the overall cost of running a mass examination using iPad technology would be substantially lower than the cost of a comparable paper-based examination. Mobility and high penetration can also be seen as benefits of the iPad approach.

We developed a software program called the Cognitive Assessment for Dementia, iPad version (CADi), to serve as a mass screening test for dementia administered via an iPad. The CADi includes items either identical or similar to the ones used in the Hasegawa Dementia Screening Scale,4 MMSE,5 Frontal Assessment Battery (FAB),6 and Trail Making Test (Army Individual Test Battery, 1944).7 The CADi consists of ten separate items, including immediate recognition of three words, questions assessing semantic memory and categorizing six objects, subtraction, backward repetition of three digits, cube rotation, pyramid rotation, trail making test A, trail making test B, and delayed recognition of three words. The final CADi score ranges from 0 to 10 correct responses. The present study examined the reliability and validity of the CADi.

Materials and methods

Participants

The CADi was administered to three groups of participants. The first group was selected from a hospital patient database. The CADi was performed consecutively on inpatients and outpatients at the Neurology Section of Shimane University Hospital. A total of 222 patients completed the CADi. Fifty patients of mean age 75.9 ± 5.6 years comprised the dementia group. The diagnosis was made by experienced neurologists based on medical history, a functional assessment, clinical examination, neuroimaging, and neuropsychological tests, including the MMSE, FAB, and Clinical Dementia Rating scale. The clinical definition of dementia used was based on the following procedures: comprehensive interview with the patient conducted by a reliable informant and involving a neurological examination, evaluation of DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) criteria for dementia, and Clinical Dementia Rating scale scores of ≥1. Patients aged 85 years or older were excluded from further analysis because matched control participants were not found. Thirty of the patients exhibited signs of Alzheimer’s disease, six had dementia with Lewy bodies, seven had frontotemporal dementia, four had vascular dementia, and three had mixed-type dementia. Data from the dementia group along with those from healthy controls were used to assess construct validity.

The second group of participants was selected from the brain checkup system database at the Shimane Institute of Health Science. This checkup system includes collection of medical, neurological, and psychiatric history, formal neurological examinations conducted by an experienced neurologist, neuropsychological assessment including the MMSE, FAB, and modified Wechsler Adult Intelligence Scale-Revised, magnetic resonance imaging of the head, echosonography of the carotid arteries, and blood tests. A total of 121 participants who voluntarily participated in this checkup system completed the CADi, and 34 of these individuals were included in the present study as a healthy control group (mean age 70.7 ± 5.0 years). The criteria for healthy controls were MMSE ≥ 27, FAB ≥ 16, modified Wechsler Adult Intelligence Scale-Revised score ≥ age-adjusted mean score, no self-awareness of memory loss, normal magnetic resonance imaging, and no history of neurological or psychiatric disease. Participants with probable early dementia or mild cognitive impairment were excluded from the healthy control group using these criteria.

The third group consisted of 2435 community-dwelling participants (mean age 68.6 ± 7.9 years) who took the CADi test as a part of voluntary routine health checkups arranged by municipalities. Four to six iPads were operated in parallel during the checkup. These health checkups were held annually at 10 districts in four rural municipalities, for a total of 33 days over the course of a year. Data from 1791 participants aged 65 years or older were used to examine the frequency distribution of test scores (mean age 72.2 ± 4.7 years; 724 males, 1067 females). Test–retest reliability was examined using data from participants from this sample who received the CADi test in consecutive years (n = 396).

CADi test

The CADi test procedure was explained to all participants before administration of the test. CADi administration occurs
in the following order: ID input, touching panel practice, check for hearing disturbance, two practice questions, and the ten test items. The examiner controlled the software only for the ID input process, after which the participant self-administered the remaining procedures. All questions and instructions were presented as text on the iPad screen as well as via audio through headphones. Stimuli that had to be memorized or reversed (see descriptions for questions 1 and 5) were presented only in audio format. All instructions and stimuli were presented in Japanese. The examiner gave any required assistance to participants, with the exception of answers to the test questions. Most participants completed the CADi test within about ten minutes. Similarly, it required less than ten minutes to train examiners.

The CADi consists of the following 10 items:

- Question 1, three words (cat, bus, and orange) are presented via audio slowly, one at a time, and the list is presented twice; the participant is asked to select the three study words from a set of six candidates (cat, dog, bus, train, apple, and orange); following this, the participant is asked to remember these three words because they will be asked to recall them later
- Question 2, the participant is asked to provide the date of termination of hostilities in World War II; the participant chooses the correct answer from a list of months (July, August, September, and October) and days (6th, 9th, 15th, and 18th)
- Question 3, six objects (car, strawberry, dog, house, computer, and watermelon) are presented; the participant is asked to choose the two items that belong to the same category
- Question 4, the participant is asked to choose the answer to the problem 93 minus 7 from four candidates (84, 85, 86, and 87)
- Question 5, three digits (5, 1, and 8) are presented via audio slowly, one at a time; the participant is then asked to key in the digits one at a time in reverse order
- Question 6, four three-dimensional shapes (two cubes depicted from different viewpoints, a rectangular parallelepiped, and a trapezoid corpus) are presented; the participant is asked to choose the pair of matching objects
- Question 7, four three-dimensional shapes (two quadrangular pyramids depicted from different viewpoints, a triangle pyramid, and a pentagonal pyramid) are presented; the subject is asked to choose the pair of matching objects
- Question 8, six digits (1, 2, 3, 4, 5, and 6) are presented at random positions on the screen; the participant is asked to touch the digits on the screen from 1 to 6 in sequential order
- Question 9, three digits (1, 2, and 3) and three Japanese hiragana characters (a, i, and u) are presented at random positions on the screen; the participant is asked to touch the digits and hiragana characters on the screen, alternating between the two in sequential order (ie, 1, a, 2, i, 3, u)
- Question 10, the participant is again asked to select the three words presented in Question 1 from amongst the six candidates.

**Statistical analysis**

We assessed the reliability and validity of the CADi using five metrics, ie, construct validity, concurrent validity, internal consistency, test–retest reliability, and frequency distribution of test scores. For each analysis, we identified appropriate participants from the stored dataset and performed analyses accordingly (see Figure 1).

**Construct validity**

We applied receiver-operating characteristic analysis to the CADi scores for the dementia and healthy control groups. Because the CADi contains only ten items, the total score is a rather rough measure and each item may differentially contribute to scores for each type of dementia. Thus, we calculated a weighted CADi score as follows:

\[
\text{Weighted CADi score} = \frac{10 \sum_{i=1}^{10} \log(\beta_i) X_i}{\sum_{i=1}^{10} \log(\beta_i)}
\]

where \(X\) is correct (1) or incorrect (0) for each item, and \(\beta\) is a logistic regression coefficient that uses the categorical value of dementia versus control and \(X\) for the dependent and independent variables. Logistic regression analysis was performed for each item, with age as a covariate. We applied receiver-operating characteristic analysis to both weighted CADi scores and total reaction time, calculated as the sum of reaction times for each item. In addition, we performed cross-validation analysis based on the “leave-one-out”
classification method. Because we found significant group differences for age, we calculated age-adjusted CADi scores that were also subjected to receiver-operating characteristic and cross-validation analyses.

Concurrent validity
For this analysis, a total of 255 participants who completed the CADi, MMSE, and FAB were selected from a pool of 343 participants in the hospital and brain checkup system database. One hundred and thirty-four patients in the hospital and 121 participants in the brain checkup system were included. All subjects were aged 65 years and over. We performed Spearman’s rank correlation analyses to examine the relationship between CADi scores and those on the MMSE and FAB.

Internal consistency
Participant data for this analysis were the same as those used to assess construct validity. We first calculated Cronbach’s alpha for all the items as a whole, and then calculated this value again for remaining items after one was removed.

Test–retest reliability
Data from 396 participants in the third (checkup) group were used to assess test–retest reliability. These individuals performed the CADi twice, with a one-year interval between administrations. We computed a Spearman’s rank correlation coefficient between the first and second CADi scores.

Frequency distribution
Finally, a frequency distribution for the CADi scores was computed using data from 1791 participants aged 65 years or older from the third checkup group. These are summarized in the form of histograms.

Results

Construct validity
Demographic characteristics of the dementia and control groups are presented in Table 1. The dementia group was older than the control group, but gender ratios were not significantly different between the groups. Years of education were equivalent across the groups. MMSE, FAB, and CADi scores were lower in the dementia group than for the control group. Furthermore, the correct response rates for all CADi items were significantly lower in the dementia group. We calculated weighted scores for all items based on logistic regression analyses (see Table 2). The weighted CADi scores were represented by the following formula:

\[
\text{Weighted CADi} = 0.69X_1 + 2.46X_2 + 0.54X_3 + 0.54X_4 + 1.05X_5 + 2.43X_6 + 0.08X_7 + 0.68X_8 + 0.80X_9 + 0.72X_{10}
\]

As was the case for the other cognitive indices, weighted CADi scores were lower for the dementia group than for the control group. Total CADi reaction times were also longer for the dementia group than for the control group.

We also computed age-adjusted CADi scores by adding and averaging residual scores with the effects of age removed, given the aforementioned age difference between the dementia and control groups. Even after adjusting for the effects of age, both CADi and weighted CADi scores were lower and total reaction times longer for the dementia group.

| Table 1 Demographic characteristics and cognitive scores of the participants |
|---------------------------------|----------------|----------------|----------------|----------------|
|                                | Dementia patients | Controls | P       | Criteria | Sensitivity | Specificity | Cross-validation |
| n                               | 50             | 34         | <0.001 |          |            |            |                |
| Age (years)                     | 75.9 (5.6)     | 70.7 (5.0) |        | <0.001  |            |            |                |
| Gender (M/F)                    | 29/21          | 17/17      | 0.27 |          |            |            |                |
| Education years (years)         | 13.0 (2.4)     | 14.6 (2.3) | 0.14 |          |            |            |                |
| MMSE score (out of 30)          | 18.6 (4.9)     | 29.0 (1.0) | <0.001|          |            |            |                |
| FAB score (out of 18)           | 9.7 (3.4)      | 16.7 (0.8) | <0.001|          |            |            |                |
| CADi score (out of 10)          | 4.6 (2.3)      | 8.4 (1.3)  | <0.001|          | 7           | 0.96        | 0.77           | 81.0%          |
| Age-adjusted CADi score (out of 10) | 4.7 (2.5) | 8.2 (1.3)  | <0.001|          | 7.6         | 0.90        | 0.82           | 78.6%          |
| wCADi score (out of 10)         | 5.1 (2.8)      | 9.2 (1.0)  | <0.001|          | 8.4         | 0.96        | 0.82           | 82.1%          |
| Age-adjusted wCADi score (out of 10) | 4.4 (2.5) | 8.3 (1.5)  | <0.001|          | 7.6         | 0.92        | 0.82           | 78.6%          |
| CADi RT (seconds)               | 313 (164)      | 115 (31)   | <0.001|          | 160         | 0.90        | 0.94           | 86.9%          |
| Age-adjusted CADi RT (seconds)  | 301 (171)      | 133 (35)   | <0.001|          | 170         | 0.82        | 0.88           | 81.0%          |

Notes: RT shows total reaction time (seconds) for CADi. Age-adjusted scores are based on mean age of 73.8 years.
Abbreviations: CADi, Cognitive Assessment for Dementia, iPad version; MMSE, Mini-Mental State Examination; FAB, Frontal Assessment Battery; wCADi, weighted CADi.
We next performed receiver-operating characteristic analyses of CADi scores. Figure 2A–C shows the CADi score distributions. At a score of ≤7, sensitivity and specificity were 0.96 and 0.77, respectively (Figure 2A). Similarly, at a weighted CADi score of ≤7.6, sensitivity and specificity were 0.90 and 0.82 (Figure 2B). Corresponding values for total reaction time were 0.90 and 0.94 (Figure 2C). Age-adjusted scores showed near equivalent sensitivity and specificity to the nonadjusted scores. Cross-validation analyses showed that nonadjusted CADi scores classified over 80% of all participants correctly (see Table 1). Correcting for age, the percentages of participants classified correctly decreased somewhat, but remained close to 80%.

Concurrent validity

Figure 3 shows scatter plots for CADi and MMSE score distributions. CADi and MMSE scores were highly correlated \( r = 0.74 \), with weighted CADi scores showing a slightly higher correlation than unweighted scores \( r = 0.75 \). If the standard MMSE cutoff value of 23 points is applied for dementia, corresponding unweighted and weighted CADi score cut points are 5.6 and 5.3, respectively. As was the case for the MMSE scores, FAB scores were also significantly correlated with CADi \( r = 0.79 \) and weighted CADi scores \( r = 0.80 \).

Internal consistency

The right-most column of Table 2 presents Cronbach’s alpha values for the whole instrument and those calculated after each item was deleted from the total item pool. The Cronbach’s alpha values for the total scale and for each item were all >0.7; these are considered to be acceptable values.

Test–retest reliability

Figure 4 shows the correlations between the first and second CADi administrations separated by a one-year interval. The correlation coefficients for the CADi and weighted CADi scores were highly significant (CADi: \( r = 0.47, P < 0.001 \), weighted CADi: \( r = 0.55, P < 0.001 \)).

Frequency distribution

Figure 5 shows the frequency distribution of CADi scores for elderly participants. Mean, median, and mode CADi scores were 7.65, 8, and 8, respectively. The cumulative percentage for a score ≤ 5 points was 9.5%.

Table 2 Correct response rates, weighted coefficients, reaction times, and Cronbach’s alpha for each Cognitive Assessment for Dementia, iPad version item

<table>
<thead>
<tr>
<th>Correct response rate</th>
<th>Reaction time (seconds)</th>
<th>Cronbach’s alpha (total 0.81)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dementia</td>
<td>Control</td>
</tr>
<tr>
<td>Immediate recognition</td>
<td>66%</td>
<td>91%</td>
</tr>
<tr>
<td>Semantic memory</td>
<td>54%</td>
<td>100%</td>
</tr>
<tr>
<td>Categorization</td>
<td>62%</td>
<td>88%</td>
</tr>
<tr>
<td>Subtraction</td>
<td>76%</td>
<td>94%</td>
</tr>
<tr>
<td>Repeating backward</td>
<td>18%</td>
<td>88%</td>
</tr>
<tr>
<td>Cube rotation</td>
<td>74%</td>
<td>100%</td>
</tr>
<tr>
<td>Pyramid rotation</td>
<td>26%</td>
<td>65%</td>
</tr>
<tr>
<td>Trail making A</td>
<td>54%</td>
<td>91%</td>
</tr>
<tr>
<td>Trail making B</td>
<td>12%</td>
<td>71%</td>
</tr>
<tr>
<td>Delayed recognition</td>
<td>14%</td>
<td>71%</td>
</tr>
</tbody>
</table>

Figure 2 Comparisons of CADi scores (A and B) and total reaction times (C) for dementia patients and healthy elderly controls. (D) Receiver-operating characteristic curve for each score and time.

Abbreviations: CADi, Cognitive Assessment for Dementia, iPad version; wCADi, weighted CADi.
Discussion

We have developed and validated a cognitive test (the CADi) for mass screening of dementia. The CADi has high sensitivity and specificity, correlates well with the MMSE and FAB, and shows good concurrent validity, internal consistency, and test–retest reliability. These results indicate that the CADi is useful for detection of dementia in the Japanese population. Further, the CADi screening test can be completed within a relatively short time frame because it consists of only ten items. This is a considerable advantage for mass screening purposes.

Healthy elderly individuals perform well on this type of screening test, which is designed to be challenging for patients suffering from dementia. Both the total CADi score and its weighted version yielded significantly lower scores for dementia patients than for the healthy control group. The percent correct responses for each item also decreased for participants with dementia. CADi scores showed high sensitivity and specificity in receiver-operating characteristic analyses. In addition, sensitivity and specificity were improved by weighting the total score on the basis of each item. These results suggest that the CADi has adequate construct validity and that it can reliably discriminate dementia patients from cognitively normal elderly people.

CADi scores were highly correlated with MMSE scores. This result indicates good concurrent validity. The MMSE has been a worldwide standard for estimating general cognitive abilities and is also valuable in the detection of dementia. However, it has been suggested that the MMSE should not be used alone as a diagnostic tool, given that MMSE outcomes are affected by age, education, and cultural background.\(^8\) We should point out that the same concerns apply to use of the CADi, which is intended for use as a primary screening tool only.
The present study indicates that the CADi has moderate internal consistency and that all items are effective for estimation of cognitive abilities. In additional, there was a significant positive correlation between the two CADi scores obtained from the same individuals over a one-year time interval, indicating that the CADi has good test–retest reliability in cognitively healthy individuals. This is important in terms of being able to use the CADi as part of an annual community-based health checkup to detect any cognitive change in individuals.

CADi scores showed a typical normal distribution with a mode of 8 points; further, 9.5% of the community-dwelling participants had CADi scores ≤ 5 points. The age-adjusted and gender-adjusted prevalence of all-cause dementia in Japan was 8.3% in 2005, and this value is very close to the cumulative percentage of ≤ 5 points for the CADi found in the present study. According to the CADi frequency distribution, participants scoring ≤ 5 points are individuals who require further examination for possible dementia.

The most likely application of the CADi is as a mass-screening test for dementia in community-based medical checkup facilities. In a receiver-operating characteristic analysis using a cutoff point of 7/8, we obtained a sensitivity of 0.96 and a specificity of 0.77. However, because we used rigorous standards for selecting elderly people who served in our control group, these criteria are likely to be excessively strict from a practical standpoint, such that they could yield too many false positives. On the basis of CADi and MMSE score scatter plots (Figure 3), the standard MMSE cutoff score (23/24) corresponds to CADi scores of 5/6. The latter cutoff point (5/6) could be a better criterion than 7/8 for clinical application. However, even given a 5-point cutoff or lower on the CADi, some participants showed fairly high MMSE scores. To judge whether a person with a low CADi score (≤ 5) requires additional clinical assessment, other information should be considered, including total reaction time. This is because dementia patients typically show decreased mental processing speed relative to healthy controls. The elderly individual’s educational history also appears to have a significant influence on cognitive performance. For instance, in one study with a standardized clinical diagnosis as the criterion, all subjects with false positives on the MMSE had fewer than 9 years of education. This may mean that relatively little formal education adversely affects cognitive performance through unfamiliarity with test-taking, or that intellectually disadvantaged individuals have low educational levels. Adjusting scores according to educational level might improve the sensitivity of the CADi by reducing the number of false positives among those with less education, while also decreasing the number of false negatives among persons with more education.

We must acknowledge some limitations. One concerns cultural background. The semantic memory question asking about the date of termination of hostilities in World War II should be modified for use in other countries. Familiarity with computers may also be expected to influence CADi performance in countries where such devices are less widely available. Second, the test–retest reliability analysis may have included some individuals with dementia, given that we could not thoroughly assess these participants. This possibility would be expected to influence the magnitude of the test–retest results. Lastly, we need a follow-up examination of community-dwelling participants with low CADi scores to confirm whether they may have or develop dementia, given that people with low education levels may have performed poorly on the test. It might also be difficult to detect mild cognitive impairment or early dementia using the CADi in highly educated people. The current version of the CADi should be useful for detecting moderate to severe cases of dementia. Future versions of the CADi should take educational attainment into account when the test is scored and interpreted.

In conclusion, we have described a new and validated mass screening test for dementia that can be best administered during community-based medical examinations. We received a great amount of feedback, indicating that participants find this test to be nonthreatening and even an enjoyable experience, possibly because the touch panel device serves to reduce participant resistance. The CADi should receive general acceptance by participants and examiners because of its simplicity and convenience.

Acknowledgment
The authors thank the Strategic Information and Communications R&D Promotion Programme of the Ministry of Internal Affairs and Communications in Japan for supporting this research.

Disclosure
The authors report no conflicts of interest in this work.

References


