American Journal of Alzheimer's Disease & Other Dementias®
Volume 24 Number 2
April/May 2009 129-135

10.1177/1533317508330222
http://ajadd.sagepub.com
hosted at
http://online.sagepub.com

Development and Evaluation of a Computerized Test Battery for Alzheimer's Disease Screening in Community-based Settings

Masashi Inoue, PhD, Daiki Jinbo, MS, Yuka Nakamura, MS, Miyako Taniguchi, PhD, and Katsuya Urakami, MD

Aim. To evaluate the capability of a computerized test battery for Alzheimer's disease screening which has been newly developed to provide a standardized and efficient method for widespread use in routine clinical and community-based settings.
Methods. Participants were 72 individuals diagnosed with Alzheimer's disease and 102 healthy elderly individuals. Both groups were tested by the battery. Receiver operating characteristic analysis was used to examine the ability of the battery to differentiate between those with Alzheimer's disease and cognitively healthy elderly individuals.
Results. On a group level, the Alzheimer's disease group performed worse than the control group on each of the 4 computerized test tasks. Receiver operating characteristic analysis yielded maximum sensitivity and specificity values of 96% and 86% for total scores, respectively.
Conclusion. We believe the battery is very useful for routine clinical and community-based settings.

Keywords: Alzheimer's disease; computer; screening

In Japan, the distribution ratio of persons over 65 years of age is increasing and approaches 20% of the total Japanese population, which amounts to about 24 million people overall. Accordingly, the number of patients with dementia is also increasing and is estimated to reach more than 2 million by 2010. Preventing an increase in the number of persons with dementia is regarded as a very urgent problem in Japan. It is important to detect the early stages of dementia because intervention programs are more effective when useful cognitive function can still be preserved.[1,2] Moreover, at the present time some medication can improve symptoms and functioning and may slow the progression of the basic disease process.[3] However, because most patients with dementia visit the hospital after their symptoms have become worse, medication is too late in many cases.[4,5] Family members and clinicians do not identify many patients as being cognitively impaired in the early stages, despite such patients having functionally significant memory problems. Although certain symptoms occur, the patients and those close to the patients may misunderstand the symptoms as signs of normal aging.[6] Such delayed detection is partly due to the lack of efficient screening tools that can be easily incorporated into a physician's practice.[7] Therefore, implementing community-based screening for dementia and appropriate primary care is critically important.

A cognitive screening battery of tests that takes a long time or is too complicated to administer is normally not deemed appropriate for the primary care setting. A useful screening battery should be required not only to be highly sensitive and specific but also to take a minimal amount of time to administer.[8] It has also been pointed out that for a test to be acceptable for general routine screening, the administration time will probably have to be approximately 2 minutes.[9] For example, even though the Mini-Mental State Examination (MMSE) is relatively easy and quick to administer, physicians often complain

From the Information Media Center (MI), and Department of Biological Regulation, School of Health Science, Faculty of Medicine (DJ, YN, MT, KU), Tottori University, Yonago, Japan.

Address correspondence to: Masashi Inoue, Information Media Center, Tottori University, Yonago 683-8503, Japan; e-mail: masashi@med.totttori-u.ac.jp.

that it takes too much time.[10] Using a computerized cognitive test battery yields some useful features; for example, it can provide quick, objective, and precise results based on the same standards for all participants examined.[11] Computerized cognitive testing has the potential to effectively address the limitations posed by traditional paper-based methods.[12] Technical innovations can improve accurate measurement of reaction time as well as decrease the frequency of errors, enhancing overall sensitivity. Although several computer-based neurocognitive test batteries have been developed,[11-16] because of the length of their administrator time, most of these do not seem suitable for community-based screening. We have developed a computerized screening battery for Alzheimer's disease (AD). This battery was intended to provide a standardized and efficient method for widespread use in routine clinical and community-based settings. In this article, we present a description and evaluation of the battery.

Methods

Participants

Participants in this study were patients of memory clinics and community-dwelling elderly residents aged over 65 in K-town, Tottori Prefecture, Japan. A total of 72 patients with AD were diagnosed by a neurological specialist based on several medical examinations as well as neuroimaging examinations and National Institute of Neurological and Communicative Disorders and Stroke Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA)[17] criteria. The healthy control group was comprised of 102 participants with no neurological or psychiatric pathology. All participants gave informed consent to be a part of this study.

Description of Computerized Test Battery

The total hardware for the test battery was comprised of a 14-inch touch panel display and computer devices built into 1 case. This made the hardware a very compact (450 W × 300 L × 38 D mm), light (2 kg), and easy-to-carry unit. The test battery operated using Windows 2000 and was bundled with the custom screening program. The screening program was originally made with reference to the Hasegawa Dementia Rating Scale[18] and consisted of 4 test tasks. These tasks are explained below in detail. The test battery was designed for use by the elderly individuals. We adopted a touch panel display as an input device so that aged people could operate it easily by touching the icon shown on the display without using a keyboard or mouse. Throughout the whole process, users were guided not only by text prompts but also by voice instructions. The test battery for the Japanese version was produced by NIHON KOHDEN Corp (Tokyo, Japan) and is being sold commercially as model MSP-1000 (Figure 1).

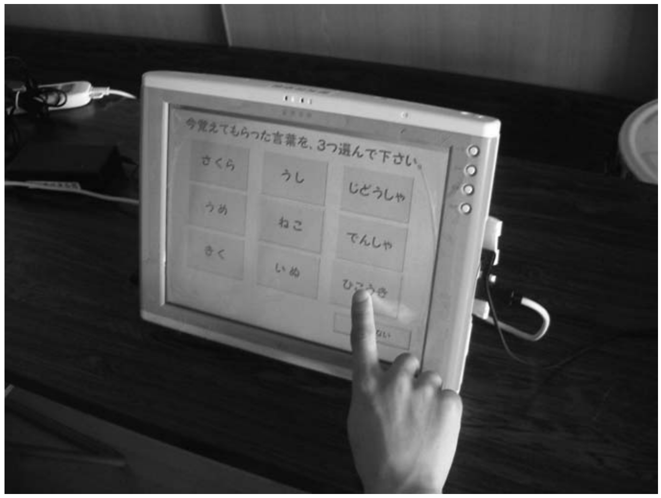

Figure 1. The Japanese version of the instrument (MSP-1000).

Details of Test Tasks

The battery is comprised of 4 test tasks and the tests are run in the same fixed order as follows:

Three-word memory test. This test examines the immediate memory of the participant. The voice prompt says 3 words (eg, cherry, dog, and train) and immediately asks the participant what the words were. Then the battery presents 9 choice icons on the screen and requires the participant to select 3 correct icons (Figure 2). The battery awards a maximum of 3 points, 1 point for each correct icon chosen. Before this test ends, the battery informs the participant that a delayed recall test will query them for the same 3 words again later on.

Temporal orientation test. For this test, the battery presents 4 screens in turn. On each screen, the battery shows icons displaying choices and asks the participant what year, month, day, and weekday it is, respectively (Figure 3). The battery awards a maximum of 4 points, 1 for each correct response.

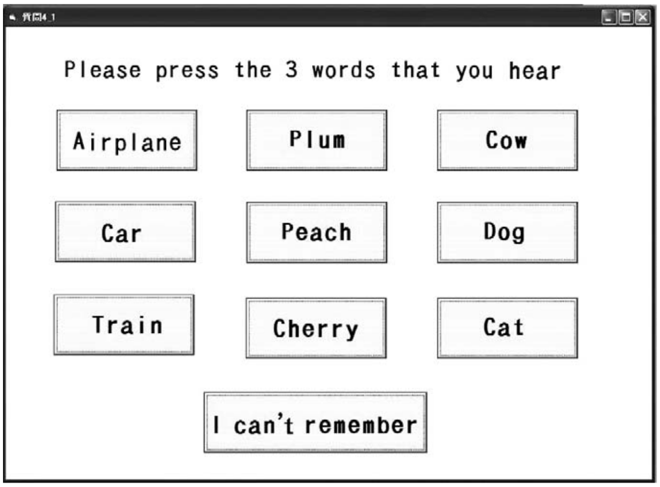

Figure 2. Screen for replay of the three-word memory test.

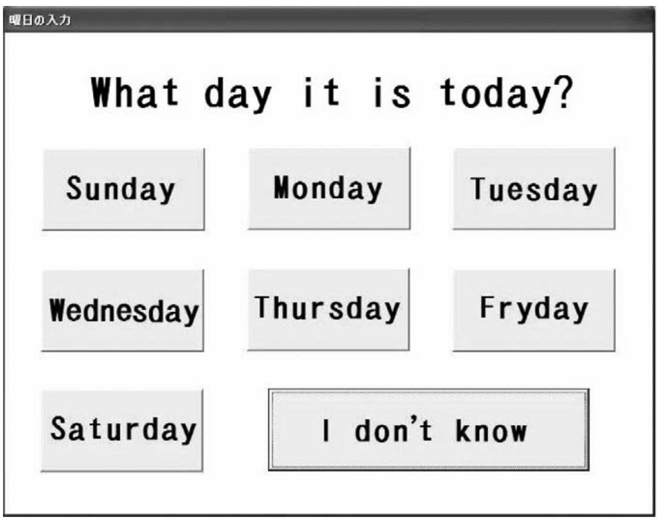

Figure 3. Screen for replay of temporal orientation for weekdays.

Three-dimensional visual-spatial perception test. This test inspects the visual-spatial perception of the participant. The battery uses two 3-dimensional figures for this test: a cube and a triangular prism. First, the battery presents a cube at the top of the screen and 5 shapes on the lower part of the screen (Figure 4). Among the 5 shapes, 1 is the same figure that is at the top (a cube), although rotated slightly. The battery requires the participant to select this 1 as the correct answer. When the participant picks out the correct figure, the battery records 1 point. Next, the battery presents a triangular prism as a target stimulus and gives 5 choices, performing the task in the same way (Figure 5). The battery awards

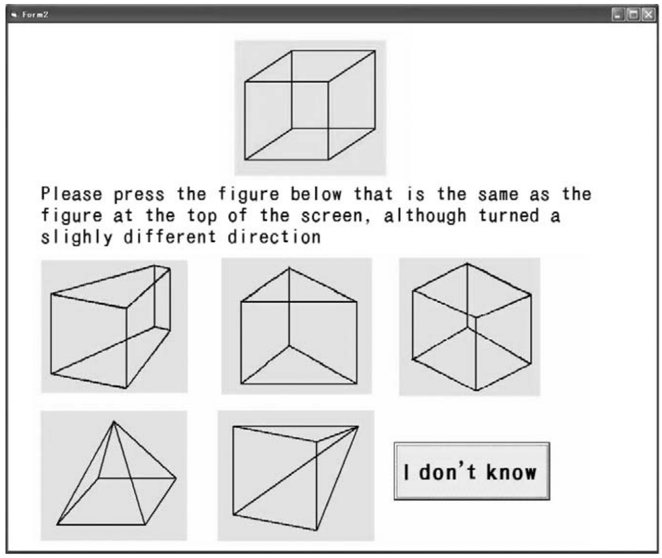

Figure 4. Screen for replay of the 3-dimensional visual-spatial perception test for cubes.

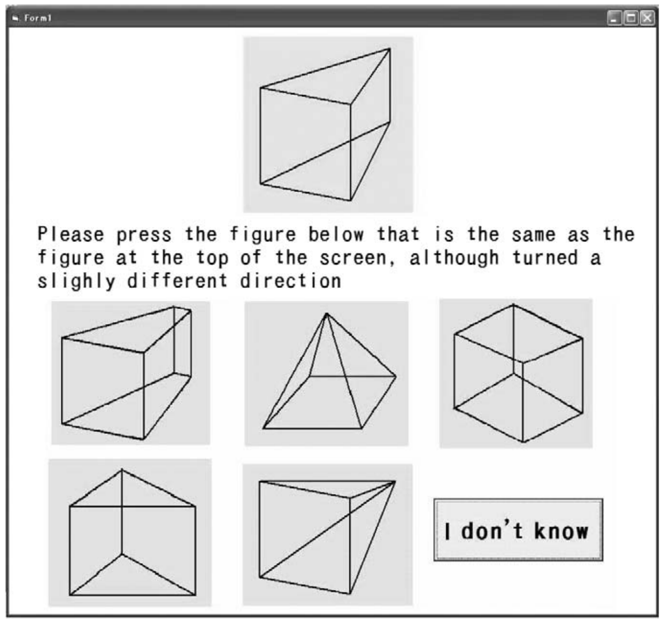

Figure 5. Screen for replay of the 3-dimensional visual-spatial perception test for triangular prism.

1 point for this one as well, for a maximum total of 2 points, 1 for each correct response.

Delayed recall test. The purpose of this test is to assess the short-term memory of the participant. The battery asks the participant to recall the same 3 words voiced in the "Three-word memory test." Icons for 9 choices are displayed as in the "Three-word

Table 1. The Demographic Characteristics of Samples and Results of Individual Test Tasks for the AD Group and Control Group[a]

	AD	Control	*t* Value	*F* Value
Sample size	72	102		
Age (years)	80.0 ± 5.1	77.1 ± 5.8	3.4[b]	
Sex M/F	12/68	37/65		
MMSE score	21.8 ± 3.1	28.2 ± 1.3		276.2[b]
Three-word memory test	2.57 ± 0.55	2.99 ± 0.01		27.4[b]
Temporal orientation test	2.56 ± 1.76	3.87 ± 0.11		86.3[b]
Visual-spatial recognition test (cube)	0.81 ± 0.15	0.92 ± 0.07		4.4[b]
Visual-spatial recognition test (triangular prism)	0.28 ± 0.20	0.4 ± 0.24		1.5
Delayed recall test	3.64 ± 2.02	5.92 ± 0.15		112.7[b]
Total score	9.87 ± 2.91	14.1 ± 0.53		178.7[b]

Abbreviations: AD, Alzheimer's disease; ANCOVA, analysis of covariance; F, female; M, male; MMSE, Mini-Mental State Examination.
[a] *F* values and *t* values are obtained by ANCOVA and Student *t* test, respectively.
[b] $P < .05$.

memory test," but the arrangement is different. The battery awards a maximum of 6 points, 2 points for each correct answer.

In addition to these tests, the battery provides a practice session to allow the participant to become familiar with the battery. Prior to the first test (the "Three-word memory test"), the battery presents a practice session for participants to prepare for the specific type of responses required for each test. All procedures are usually completed within about 4 minutes and the maximum possible score on this test is 15 points.

Statistical Analysis

Statistical analysis was performed using SPSS software (Version 11 for Windows). The level of statistical significance was set at 0.05 for all tests. The differences in mean results for individual tests for both groups were examined by analysis of covariance (ANCOVA). Analysis of covariance is a merger of analysis of variance (ANOVA) and regression for continuous variables. Analysis of covariance tests whether certain factors have an effect on the outcome variable after removing the variance for which quantitative predictors (covariates) account. Receiver operating characteristic (ROC) analysis was performed for each test to show the relation between sensitivity and specificity when the cutoff probability varies. The area under the curve (AUC), an index of effect size, was the primary result of the ROC analysis. Test accuracy is measured by the area under the ROC curve. An area of 1.0 represents a perfect test; an area of 0.5 represents an unusable test.

Results

Almost all the participants could understand how to interact with the battery and could complete the task by themselves. However, persons who had weak eyesight and/or bad hearing required the help of an assistant staff member who supported them when they were confused about how to respond. In such cases, the staff instructed the participants how to operate the battery without proposing any correct answers. The number of such persons was less than 20 in this study. Table 1 represents the demographic characteristics of the sample and results of individual test tasks of the AD and control groups. Unpaired Student *t* test revealed significant differences in mean ages for the 2 groups. Therefore, we used ANCOVA, which treated age as a covariate, to analyze differences of scores between the 2 groups. On a group level, the AD group performed worse than the control group on each of the 4 tasks. Analysis of covariance revealed that there were significant differences between the 2 groups for each individual test, except for the 3-dimensional visual-spatial perception test using the triangular prism, and for the total score.

A summary of the sensitivity and specificity parameters and cutoff points for the "delayed recall test" and the "temporal orientation test," and the total score, which revealed a high *F* value, comparing the AD group with the control group, are shown in Table 2. The data are represented graphically in Figure 6, which provides ROC curves for each test comparing the AD group with the control group and revealed that the total score was more specific and sensitive than each of the 4 tests. The ROC analysis also yielded

Table 2. Summary of the Sensitivity and Specificity of Delayed Recall, Temporal Orientation, and Total Score of the Instrument for the AD group Versus the Control Group

	Delayed Recall	Temporal Orientation	Total Score
Sensitivity	0.69	0.69	0.97
Specificity	0.95	0.87	0.85
Area under the ROC curve	0.84	0.80	0.93
Cutoff point	4	3	13

Abbreviations: AD, Alzheimer's disease; ROC, receiver operating characteristic.

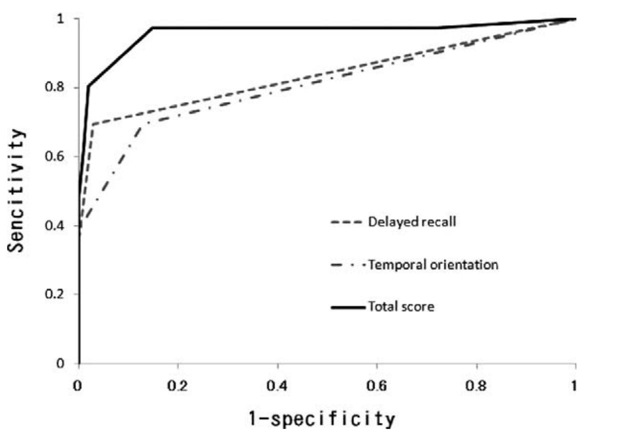

Figure 6. ROC curve for delayed recall, temporal orientation, and total score for the instrument between ADs and controls. AD = Alzheimer's disease; ROC = receiver operating characteristic.

maximum sensitivity and specificity values of 96% and 86% for the total score, respectively, with a cutoff point of 13.

Discussion

Community-based screening is a useful way to detect persons with AD in its early stages. A useful screening battery for AD needs to be brief, simple, and accurate. To satisfy these demands, the screening battery should combine the fewest sensitive test tasks possible and should yield the highest sensitivity and specificity possible. Because memory impairment is often one of the earliest features of AD, a very brief test for memory impairment such as a delayed recall test was found by previous research to be effective for detecting AD.[1] The "temporal orientation test" and "visual-spatial perception test" were also described in the literature as useful items for screening AD.[19] Based on such evidence, we assembled the aforementioned sensitive test tasks into the computerized screening battery. Currently, the MMSE and Hasegawa Dementia Scale are widely used to differentiate between normal aged people and cases of dementia in Japan. However, these tests were originally designed to diagnose dementia and are not suitable for the screening of AD because they take too long and are too cumbersome to administer. Several different batteries[20-22] among others were developed for the purpose of screening for AD. These tests are administered based mostly on face-to-face interviews.

Both the human screening method and the computerized screening method have advantages and disadvantages. Human testers can respond flexibly according to the condition of the participant. However, their treatment seems to be biased according to their experience or knowledge, and there seems to be a risk that the criteria could vary from tester to tester.[23] In the case of community-based screening, one of the necessary requirements which a test method must fulfill is to produce unbiased results even if the examiner is changed. Standardization is also required for screening tests to be accurate in judgment and to be widely used. Therefore, a screening test that requires the examiner to have special training or skills is not always feasible. In this way, computerized testing is good at getting the needed information based on the same standards, without the aforementioned problems. Our battery, which usually takes less than 4 minutes to administer, seems to be adequate for use in community settings. The overall sensitivity and specificity of the battery for AD were 96% and 86%, respectively, which are equivalent to other studies.[1,9,20,21] Of course, it is important to have follow ups or more precise diagnoses for people who screen with low scores. The fact that more false-positive than false-negative results will occur during screening is likely to be eliminated by accurate clinical diagnosis after screening.

Recently, computer speech recognition technologies have advanced and have been applied in some domains. However, these technologies often require users to be trained under controlled conditions to improve performance accuracy. Computers do not yet have the ability to understand correctly what an unspecified person might say. Therefore, we thought it best to design a battery of tests, which can give definite results through the recognition of responses by touching icons shown on a screen.

Displaying possible answer choices is like giving a small hint to the participant, because it does not require participants to generate answers spontaneously. On the delayed recall task, presenting cues during the retrieval phase improves the specificity in discriminating individuals with dementia from unimpaired older people.[24,25] Persons who are able to solve the question according to the hint seem to retain their cognitive ability and receive a high score. However, persons who get a low score even though hints are given are strongly suspected to be in cognitive decline. Selecting one of the choices seems to be adequate for the assessment of memory impairment, and that method can be easily adopted as a computerized procedure. However, computers are unsuitable for assessing dialogic ability such as verbal fluency, which is one disadvantage.

Additionally, examiners often feel awkward when patients are unable to answer questions correctly during face-to-face interviews. The patients often seem embarrassed as well. In such cases, even though it is in front of a doctor, the patients seem to feel their pride considerably wounded by revealing their cognitive decline. However, in the case of a computer testing battery, this feeling of unpleasantness rarely occurs.

These results were obtained from elderly Japanese persons. As was mentioned in another study,[26] performance on a screening test is affected by differences in population or ethnicity. Therefore, care should be used when conducting examinations with the battery in other populations. The predictive value of the battery must also be reevaluated for each different population.

In conclusion, we have developed a computerized screening battery for AD, which yields a sensitivity and specificity of 96% and 86%, respectively. This battery can provide quick, objective, and precise results based on solid standards proven by research. Although computerized testing cannot replace a human interviewer's methods in all settings, we believe it is useful for routine clinical or community settings.

References

1. Buschke H, Kuslansky G, Katz M, et al. Screening for dementia with the Memory Impairment Screen. *Neurology*. 1999;52:231-238.
2. Barker WW, Luis C, Harwood D, et al. The effect of a memory screening program on the early diagnosis of Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2005; 19:1-7.
3. Knopman DS. Current treatment of mild cognitive impairment and Alzheimer's disease. *Curr Neurol Neurosci Rep*. 2006;6:365-371.
4. Bennett DA, Schneider JA, Arvanitakis Z, et al. Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*. 2006;66: 1837-1844.
5. Ganguli M, Rodriguez E, Mulsant B, et al. Detection and management of cognitive impairment in primary care: the Steel valley Seniors survey. *J Am Geriatr Soc*. 2004;52: 1668-1675.
6. Ross GW, Abbott RD, Petrovitch H, et al. Frequency and characteristics of silent dementia among elderly Japanese-American men. *JAMA*. 1997;277:800-805.
7. Trenkle DL, Shankle WR, Azen SP. Detecting cognitive impairment in primary care: performance assessment of three screening instruments. *J Alzheimers Dis*. 2007;11: 323-335.
8. Solomon PR, Murphy CA. Should we screen for Alzheimer's disease? A review of the evidence for and against screening Alzheimer's disease in primary care practice. *Geriatrics*. 2005;60:26-31.
9. Mendiondo MS, Ashford JW, Kryscio RJ, Schmitt FA. Designing a Brief Alzheimer Screen (BAS). *J Alzheimers Dis*. 2003;5:391-398.
10. David RD, Jeffrey LC. Evaluating dementia screening tests. *Neurology*. 1999;52:224-227.
11. Letz R, Green RC, Woodard JL. Development of a computer-based battery designed to screen adults for neuropsychological impairment. *Neurotoxicol Teratol*. 1996;18:365-370.
12. Green RC, Green J, Harrison JM, Kutner MH. Screening for cognitive impairment in older individuals. Validation study of a computer-based test. *Arch Neurol*. 1994;51: 779-786.
13. Robbins TW, James M, Owen AM, Sahakian BJ, McInnes L, Rabbitt P. Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia*. 1994;5:266-281.
14. Dwolatzky T, Whitehead V, Doniger GM, et al. Validity of the Mindstreams computerized cognitive battery for mild cognitive impairment. *J Mol Neurosci*. 2004;24:33-44.
15. Doniger GM, Zucker DM, Schweiger A, et al. Towards practical cognitive assessment for detection of early dementia: a 30-minute computerized battery discriminates as well as longer testing. *Curr Alzheimer Res*. 2005;2: 117-124.
16. Gualtieri CT, Johnson LG. Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Arch Clin Neuropsychol*. 2006;21:623-643.
17. 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 Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34:939-944.

18. Katoh S, Shimogaki H, Onodera A, et al. Development of the revised version of Hasegawa's Dementia Scale. *Japanese J Geriatr Psychiatry*. 1991;2:1339-1347 (in Japanese).
19. Inoue M, Urakami K, Taniguchi M, Kimura Y, Saito J, Nakashima K. Evaluation of a computerized test system to screen for mild cognitive impairment. *Psychogeriatrics*. 2005;5:36-41.
20. Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry*. 2000;15:548-561.
21. Solomon PR, Hirschoff A, Kelly B, et al. A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease. *Arch Neurol*. 1998;55:349-355.
22. Cho A, Sugimura M, Nakano S, Yamada T. The Japanese MCI screen for early detection of Alzheimer's disease and related disorders. *Am J Alzheimers Dis Other Demen*. 2008;23:162-166.
23. Kobak KA, Taylor LH, Dottl SL, et al. A computer-administered telephone interview to identify mental disorders. *JAMA*. 1997;17:905-910.
24. Kuslansky G, Buschke H, Katz M, Sliwinski M, Lipton RB. Screening for Alzheimer's disease: the memory impairment screen versus the conventional three-word memory test. *J Am Geriatr Soc*. 2002;50: 1086-1091.
25. Ivanoiu A, Adam S, Van der Linden M, et al. Memory evaluation with a new cued recall test in patients with mild cognitive impairment and Alzheimer's disease. *J Neurol*. 2005;252:47-55.
26. Welsh KA, Fillenbaum G, Wilkinson W, et al. Neuroopsychological test performance in African-American and white patients with Alzheimer's disease. *Neurology*. 1995;45:2207-2211.