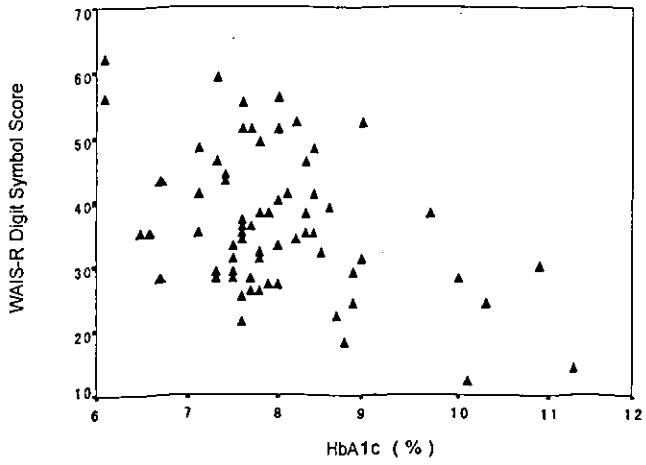


Fig. 1. Relationship between glycosylated hemoglobin (HbA1c) concentration and score of WAIS-R Digit Symbol on DM subjects (N = 69 r = −0.433, $P<0.001$).

DM, neuropathy was diagnosed as elevated vibratory perception thresholds or symptomatic neuropathy including parethesia; retinopathy was diagnosed as simple retinopathy and more advanced; and nephropathy was diagnosed as microalbuminuria (30 mg/g ≦ albumin-to-creatine ratio < 200 mg/g) and more advanced.

2.2. Assessment of cognitive function

Cognitive function was assessed by structured performance tests that were selected to represent a broad range of cognitive domains, including those measured in previous studies in type 2 DM. Strachan et al. (1997) summarized psychological tests used among the previous studies into six broad categories. In considering total administration time and elderly subjects' burden to complete the test battery, we decided to investigate four of these categories, and selected the following standardized psychological tests for measurement of each. (1) Mental Status: the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHigh, 1978) was used to evaluate this category. The MMSE assesses orientation, registration, attention, calculation, language, and recall with a score range from 0 to 30. (2) Verbal Memory: the Word List (a subtest of the Alzheimer's Disease Assessment Scale [ADAS] (Mohs, Rosen, & Davis, 1983)) was used. This test asks subjects to read aloud and remember 10 concrete words printed on individual cards. The subjects' immediate recall is evaluated directly after reading the words, and the delayed recall is assessed at 30 min after reading the words. The score range is 0–10. (3) Complex Psychomotor Skill: the Digit Symbol Test (a subtest of the Wechsler Adult Intelligence Scale—Revised [WAIS-R]; Shinagawa, Kobayashi, Fujita, & Maekawa, 1990) was used. This test consists of a sample line with nine pairs of numbers and meaningless symbols. Subjects are asked to fill in the blanks with the correctly paired symbols in 90 s. The score range is 0–93. (4) Attention: the Stroop Color Word Test (Stroop, 1935; Japanese version) employs a card with 24 colored dots and a card with 24 names of colors printed in different colored ink, e.g., the word "yellow" printed in blue ink. Subjects are asked to name the color of the dots as quickly as possible, and then to name the color of the ink that a color word was printed in as quickly as possible. Seconds to completion are recorded and the difference between the time required to read the word card and that required to read the dots card is calculated. Well-trained psychological testers administered all four tests in the same order for all subjects.

Table 4
Characteristics and performance on measures of cognitive function by diabetes subgroup

Variable	Insulin-treated diabetes	Noninsulin diabetes	Nondiabetic subjects	*P* value
N	13	56	27	
Age	72.8 ± 6.3	71.3 ± 5.4	73.4 ± 6.6	0.261
Education (year)	10.5 ± 2.6	10.4 ± 2.7	11.4 ± 3.0	0.386
MMSE	25.4 ± 2.0[††]	27.5 ± 2.0[¶]	28.3 ± 1.7[§]	$P<.01$
Word List (immediate)	5.2 ± 1.5	5.9 ± 1.7	6.2 ± 1.7	0.205
(delayed)	6.8 ± 1.3	7.2 ± 2.3	6.7 ± 2.0	0.567
WAIS-R Digit Symbol	29.2 ± 8.1[††]	37.9 ± 10.8[¶]	43.0 ± 12.1[§]	$P<.01$
Stroop Color Word Test	17.8 ± 7.6	19.5 ± 13.8	15.0 ± 6.7	0.252
HbA1c (%)	8.7 ± 1.5[†]	7.9 ± 0.8[¶¶]	5.7 ± 0.4[§§]	$P<.01$
Hypertension (%)	53.2	50.0	50.0	0.845
Hyperlipidemia (%)	36.4	36.6	60.0	0.074

Data are the mean ± S.D., unless otherwise indicated.
A higher score indicates better performance, except in the case of the Stroop Color Word Test.
ANOVA. Bonferroni's post-hoc test showed following;
[†] Significant difference with noninsulin diabetes ($P<.05$) and nondiabetic subjects ($P<.01$).
[††] Significant difference with noninsulin diabetes ($P<.05$) and nondiabetic subjects ($P<.05$).
[¶] Significant difference with insulin-treated diabetes ($P<.05$).
[¶¶] Significant difference with insulin-treated diabetes ($P<.05$) and nondiabetic subjects ($P<.05$).
[§] Significant difference with insulin-treated diabetes ($P<.01$).
[§§] Significant difference with insulin-treated diabetes ($P<.05$) and noninsulin subjects ($P<.05$).

2.3. Statistical analysis

All data are presented as the means ± S.D. Comparisons between two groups were made by using Student's t–test or the Kruskal–Wallis analysis. Pearson's correlation coefficients and Spearman's correlation coefficients were calculated for parametric and nonparametric variables, respectively. For the Spearman's correlation coefficients, nonparametric variables were coded as follows. The existence of DM complication or hypoglycemic episode was scored as 1, and the absence of these parameters was scored as 0. The use of insulin was scored as 1, and the nonuse as 0. Comparisons among three groups were made by using analysis of variance (ANOVA) followed by Bonferroni's post hoc test. In all analyses, values of $P<.05$ were considered to indicate statistical significance. All analysiswas performed with SPSS software for Windows (SPSS Inc., 2001).

3. Results

The characteristics of the subjects included in the current study are shown in Table 1. There were no significant differences between DM patients and non-DM subjects in any area except HbA1c ($P<.01$). Table 2 shows the means and S.D. for four measures of cognitive function in DM and control subjects. The DM group performed significantly worse in the MMSE ($P<.05$) and Digit Symbol Test ($P<.05$), and tended to perform more poorly in other tests, although these differences were not significant. The results of the correlation analysis between cognitive tests scores and diabetic characteristics are shown in Table 3. The scores of the Digit Symbol Test in DM subjects had a significant negative correlation with HbA1c ($r=-.433$, $P<.001$, Fig. 1). The insulin-use for DM treatment was also significantly correlated to the Digit Symbol Test ($r=-.304$, $P<.05$). Further, we divided the DM group into two subgroups: an insulin-treated and a non-insulin-treated group. The history of insulin treatment ranged from 1 to 30 years (mean = 10.85; S.D. = 11.10). The frequencies of daily insulin injection were scored as follows: once = 3, twice = 6, three times = 3, four times = 1. ANOVA showed that both the MMSE and Digit Symbol Test scores were significantly different among the three groups ($P<.01$, $P<.01$, respectively, Table 4). Bonferroni's post hoc test showed that the scores of the MMSE and Digit Symbol Test in the insulin-treated DM group was significantly lower than those in the non-insulin-treated DM ($P<.05$), and non-DM ($P<.01$) subjects. Insulin-treated DM subjects had significantly higher HbA1c. The prevalence of DM complications and hypoglycemic episodes by DM treatment were compared and are shown in Table 5. The frequency of hypoglycemic episodes was from three times a year to once a month. No subjects experienced hypoglycemic coma or hospitalization from hypoglycemia-related events. Only prevalence of retinopathy was significantly different between insulin-treated and non-insulin-treated DM subjects.

Table 5
The characteristics and presence of diabetic complication and hypoglycemia by diabetic treatment groups

Variable	Insulin-treated DM	Noninsulin DM	P value
Diabetes Duration (year)	19.2 ± 12.6	13.7 ± 11.2	.164
Neuropathy (%)	66.7	59.2	.637
Retinopathy (%)	63.6	31.2	$P<.05$
Nephropathy (%)	58.3	31.2	.084
Hypoglycemia (%)	36.4	11.4	.060

Data are [the] mean ± S.D., unless otherwise indicated.
Student's t–test. (diabetes history), Kruskal–Wallis analysis (other variables).

4. Discussion

In the current study we investigated cognitive function in the Japanese elderly with type 2 DM. Recently many studies—primarily from Western countries—have reported cognitive functional deficits in elderly DM subjects, and the term "diabetic encephalopathy" is seeing increasing use (Biessels, der Heide, Kamal, Bleys, & Gispen, 2002). Nonetheless, there have been few reports showing cognitive deficits in Japanese subjects with DM. Our results suggest that DM-related cognitive impairment does exist in Japanese subjects, and that such impairment is not dependent on cultural or genetic background.

In the present study the insulin-treated subjects had the worst cognitive function. Three possible mechanisms could be hypothesized for the poor cognitive function in insulin-treated subjects. In the present study, subjects receiving insulin-treatment had significantly higher HbA1c, and the scores of the Digit Symbol test were negatively correlated with HbA1c; therefore, the effects of hyperglycemia should be given primary consideration. Secondly, hypoglycemia-related neuronal damage may have been involved, since insulin-treated elderly DM subjects reportedly have more risks for hypoglycemia (Schorr, Ray, Daugherty, & Griffin, 1997). However, several studies have shown that subjects with impaired glucose tolerance who receive no drug treatment and are at minimum risk for hypoglycemia also sometimes show cognitive impairment (Convit, Wolf, Tarshish, & de Leon, 2003; Vanhanen et al., 1988). This suggests that hypoglycemia may not be a major factor for cognitive functional deficits. Our study failed to show a significant correlation between hypoglycemia and the scores of cognitive functional tests. Thirdly, direct effects of insulin on the neuronal system may have played a role in the poor cognitive function of insulin-treated subjects. Recently, insulin and its receptor have been shown to be present in the brain and appear to play a modulatory role in synaptic transmission (Schwartz, Figlewicz, Baskin, Woods, & Porte, 1992). Ott et al. (1996, 1999) reported that insulin-treated subjects are at greater risk for dementia. In the present study,

however, analysis of the correlation coefficient between fasting serum insulin levels in non-insulin-treated DM subjects and the performance on cognitive functional tests failed to show any significant differences ($P = .092–0.645$, data not shown).

As discussed above, hyperglycemia may be a major factor for cognitive impairment in elderly DM subjects. However, in the current study, microangiopathies, including neuropathy, were very weakly related to the scores of the MMSE and Digit Symbol test. The duration of DM also showed no significant relationship with the scores of the cognitive functional tests. These findings suggest that cognitive impairment in DM subjects is induced by mechanisms other than microangiopathy. The scores of the Digit Symbol test had a significant negative association with HbA1c, which reflects the status of glycemic control in a relatively short period of time. Interestingly, several studies have shown that improvement in glycemic control had beneficial effects on cognitive function in elderly DM subjects (Gradman et al., 1993; Meneilly et al., 1993). This would suggest that the cognitive function in DM subjects is affected at least partially by the blood glucose levels for a relatively short period of time. Further longitudinal studies with a larger number of subjects will be needed to follow the cognitive function in elderly DM subjects, as well as investigations into the mechanism by which hyperglycemia affects cognitive function.

5. Conclusion

In conclusion, we demonstrated that Japanese elderly with DM had mild cognitive impairment. Hyperglycemia may be a major factor for this cognitive impairment; however, further studies with increased number of subjects and longitudinal studies will be needed to clarify this relation.

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (DSM-III-R)* (3rd rev. ed.). Washington, DC: Author

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV)* (4 ed.). Washington, DC: Author

Atiea, J. A., Moses, J. L., & Sinclair, A. J. (1995). Neuropsychological function in order subjects with non-insulin-dependent diabetes mellitus. *Diabetic Medicine, 12*, 679–685.

Biessels, G. J., der Heide, L. P., Kamal, A., Bleys, R. L. A., & Gispen, W. H. (2002). Ageing and diabetes: implication for brain function. *European Journal of Pharmacology, 441*, 1–14.

Convit, A., Wolf, O. T., Tarshish, C., & de Leon, M. J. (2003). Reduced glucose tolerance is associated with poor memory performance and hippocampal atrophy among elderly. *Proceedings of the National Academy of Sciences, 100* (4), 2019–2022.

Folstein, M. F., Folstein, S. E., & McHigh, P. R. (1978). "Mini-mental state": a practical method of grading the cognitive function of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.

Gradman, T. J., Laws, A., Thompson, L. W., & Reaven, G. M. (1993). Verbal learning and/or memory improves with glycemic control in older subjects with non-insulin-dependent diabetes mellitus. *Journal of the American Geriatrics Society, 41*, 1305–1312.

Jagusch, W., Cramon, D. Y., Renner, R., & Hepp, K. D. (1992). Cognitive function and metabolic state in elderly diabetic patients. *Diabetes, Nutrition & Metabolism, 5*, 265–274.

Japan Atherosclerosis Society. (2002). *Japan atherosclerosis society guidelines for diagnosis and treatment of atherosclerotic cardiovascular disease* (Tokyo).

Mattlar, C. E., Falck, B., Ronnemaa, T., & Hyyppa, M. T. (1985). Neuropsychological cognitive performance of patients with type-2 diabetes. *Scandinavian Journal of Rehabilitation Medicine, 17*, 101–105.

Meneilly, G. S., Cheung, E., Tessier, D., Yakura, C., & Tuokko, H. (1993). The effect of improved glycemic control on cognitive functions in elderly patients with diabetes. *Journal of Gerontology, 48*, M117–M121.

Miles, W. R., & Root, H. F. (1922). Psychologic tests applied to diabetic patients. *Archives of Internal Medicine, 30*, 767–777.

Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: an instrument for assessing treatment efficacy. *Psychopharmacology Bulletin, 19*, 448–450.

Ott, A., Stolk, R. P., Harskamp, F. V., Pols, H. A., Hofman, A., & Breteler, M. M. (1999). Diabetes mellitus and the risk of dementia—the Rotterdam study. *Neurology, 53*, 1937–1942.

Ott, A., Stolk, R. P., Hofman, A., Harskamp, F. V., Grobbee, D. E., & Breteler, M. M. (1996). Association of diabetes mellitus and dementia: the Rotterdam study. *Diabetologia, 39*, 1392–1397.

Perlmuter, L. C., Hakami, M. K., Hodgson-Harrington, C., Ginsberg, J., Katz, J., Singer, D. E., & Nathan, D. M. (1984). Decreased cognitive function in aging non-insulin-dependent diabetic patients. *American Journal of Medicine, 77*, 1043–1048.

Reaven, G. M., Thompson, L. W., Nahum, D., & Haskins, E. (1990). Relationship between hyperglycemia and cognitive function in older NIDDM patients. *Diabetes Care, 13*, 16–21.

Schorr, R. I., Ray, W. A., Daugherty, J. R., & Griffin, M. R. (1997). Incidence and risk factors for serious hypoglycemia in older persons using insulin or sulfonylureas. *Archives of Internal Medicine, 157*, 1681–1686.

Schwartz, M. W., Figlewicz, D. P., Baskin, D. G., Woods, S. C., & Porte, D. (1992). Insulin in the brain: a hormonal regulator of energy balance. *Endocrine Reviews, 13* (3), 387–414.

Shinagawa, F., Kobayashi, S., Fujita, K., & Maekawa, H. (1990). *Japanese manual for the Wechsler Adult Intelligence Scale—Revised* (pp. 115–118). Tokyo: Nihon-bunka-kagaku-sya.

SPSS Inc. (2001). *SPSS Base 11.0J User's Guide* (Tokyo).

Strachan, M. W., Deary, I. J., Ewing, F. M., & Frier, B. M. (1997). Is type 2 diabetes associated with an increased risk of cognitive dysfunction? *Diabetes Care, 20*, 438–445.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Ushida, C., Umegaki, H., Hattori, A., Mogi, N., Aoki, S., & Iguchi, A. (2001). Assessment of brain atrophy in elderly subjects with diabetes mellitus by computed tomography. *Geriatrics and Gerontology International, 1*, 33–37.

Vanhanen, M., Kovisto, K., Kuusisto, J., Mykkanen, L., Helkala, E. L., Hanninen, T., Riekkinen, P. Sr, Soininen, H., & Laakso, M. (1988). Cognitive function in an elderly population with persistent impaired glucose tolerance. *Diabetes Care, 21* (3), 398–402.

Original Research Article

Dementia
and Geriatric
Cognitive Disorders

Dement Geriatr Cogn Disord 2004;18:172–179
DOI: 10.1159/000079198

Accepted: January 22, 2004
Published online: June 21, 2004

The Clock Drawing Test as a Valid Screening Method for Mild Cognitive Impairment

Sayaka Yamamoto[a] Nanaka Mogi[a] Hiroyuki Umegaki[a] Yusuke Suzuki[a]
Fujiko Ando[b] Hiroshi Shimokata[b] Akihisa Iguchi[a]

[a]Department of Geriatrics, Medicine in Growth and Aging, Program in Health and Community Medicine, Nagoya University Graduate School of Medicine, Nagoya, and [b]Department of Epidemiology, National Institute for Longevity Sciences 36–3, Gengo, Morioka, Obu, Japan

Key Words
Clock drawing · Early diagnosis of dementia · Screening test · Cutoff point · Cahn's scoring protocol

Abstract
To validate the Clock Drawing Test (CDT) as a screening method for detecting mild cognitive impairment (MCI) and to find the appropriate scoring protocol and its cutoff point, we compared the sensitivity and specificity of three CDT protocols. Subjects included 219 outpatients with memory complaints, who were attending the geriatric memory clinic. Cahn's protocol, with a cutoff point of 7, was more successful at differentiating clinically diagnosed MCI subjects from normal elderly individuals, with higher sensitivity (74.7%) and specificity (75.6%), than were the other protocols. The CDT, as a handy screening method, may be useful for clinicians to reliably identify subjects with MCI, and it may contribute to early detection of dementia.

Introduction

Early detection of dementia is an issue of growing concern because of improved clinical outcomes expected as a result of early therapeutic interventions or preventive approaches [1]. In terms of care, the diagnosis of cognitive deficits at an early stage, when the patient is still competent enough to make important decisions, can give the patients and their caregivers the opportunity to prepare for situations expected to occur as the symptoms progress (e.g. making environmental arrangements or educating the family), and also facilitate autonomic future planning (e.g. writing a living will, assigning durable power to an attorney or composing advanced directives) [2]. The term 'mild cognitive impairment' (MCI) was originally used to describe a transitional state between normal condition and Alzheimer's disease (AD) [3] and was first defined by Petersen et al. [4]. Recently, a revised and extended definition of MCI has been proposed that covers a broader range of cognitive impairment. It categorizes MCI into the following three subtypes: purely amnestic syndrome, impairment of a single nonmemory domain of cognition, and slight cognitive impairment in multiple domains of cognition [5]. It has also been suggested that enlarging the definition would allow the screening of more subjects at

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1420–8008/04/0182–0172$21.00/0

Accessible online at:
www.karger.com/dem

Sayaka Yamamoto, MD
Department of Geriatrics, Medicine in Growth and Aging, Program in Health and Community Medicine, Nagoya University Graduate School of Medicine
65 Tsurumai-cho, Showa-ku, Nagoya-shi, 466–8550 (Japan)
Tel. +81 52 744 2364, Fax +81 52 744 2371, E-Mail saaya@med.nagoya-u.ac.jp

risk of dementia [6]. Although many detailed neuropsychological tests evaluating executive functions are available as screening instruments to quantify the degree of cognitive impairment, most of them are impractical for general physicians to administer in their clinical settings [2, 7]. To our knowledge, none of these neuropsychometric tests can contribute to the accurate diagnosis of MCI with reliable sensitivity and specificity.

The Clock Drawing Test (CDT) has been arousing the interest of clinicians and researchers as a convenient screening instrument for dementia, either by itself or as a part of a brief neuropsychological test battery. The CDT takes less than 2 min to administer, and it is easy to comprehend the instructions, making it suitable for elderly patients who may not be able to maintain concentration [8]. Previous replication studies, which applied various scoring systems, demonstrated that the CDT is a reliable method for the detection of dementia [9–12]. The CDT is relatively less affected by the level of education, language, and cultural background than are the other cognitive tests such as the Mini-Mental State Examination (MMSE) [13, 14]. The protocols of the CDT in published studies are various. They differ not only in the instructions (with or without predrawn circle, different time setting) but also in their scoring criteria. To obtain acceptable reliability for screening MCI among a target population, the cutoff points of the CDT applied in previous studies must be reexamined, since they may not be appropriate in populations with lower prevalence rates of dementia [9]. To date, there has been a dearth of studies examining the utility of the CDT in detecting cognitive deficits in their early stages, in particular MCI. Previous studies yielded conflicting results with limitations in terms of the criteria used for group assignment, the sample size, and the optimal cutoff points for different types of cognitive status [15, 16]. Besides, most previous studies emphasized the use of quantitative analyses of the CDT results, leaving detailed qualitative analyses of the results somewhat neglected. Despite the existence of reports regarding qualitative analyses of the CDT results in AD and vascular dementia (VD) patients, those focusing on MCI subjects are still lacking [17, 18]. The purpose of this study was to warrant the validity of the CDT as a screening method for detecting MCI by determining an appropriate scoring protocol with an optimal cutoff point and qualitative features.

Table 1. The neuropsychological test battery

Function	Test
Global cognitive function	MMSE
Orientation	MMSE-1, 2
Memory	MMSE-5
	Verbal recall (ADAS, paragraphs)
Verbal fluency	Initial letter
	Category
Visuospatial praxis	CDT, MMSE-11
	ADAS-7
Psychomotor speed	Digit symbol
Attention	Stroop test
	Digit Span

ADAS = The Alzheimer's disease assessment scale; ADAS-7 = constructional ability.

Methods

Participants

Subjects were recruited from outpatients at the geriatric memory clinic in the Nagoya University Hospital. A total of 219 subjects (male: 75, female: 144) aged 60 years and older who had either subjective memory complaints or memory loss reported by their informants participated in this study. Informed consent was obtained from all the participants or their primary caregivers after complete description of the study. The age of the participants ranged from 60 to 93 years (mean = 75.1 years, SD = 6.7 years). Years of education ranged from 3 to 24 years (mean = 10.2 years, SD = 2.8 years). None of the participants had a history of neurological or psychiatric disorders, and none had been diagnosed as having reversible causes of cognitive impairment. Routine physical examinations and neurological examinations had been carried out in all subjects. Subjects with receptive aphasias or visual impairment and those who had abnormal thyroid functions or serum vitamin B_{12} or folate levels in laboratory studies were excluded from the study. Magnetic resonance imaging of the brain was performed on all subjects. The Geriatric Depression Scale (GDS)-15 was applied as a screening test for excluding subjects with possible depression at a cutoff point of 8 [19]. Subjects were administered a neuropsychological test battery including the CDT, as shown in table 1 [20–22]. General cognitive impairment was assessed by the MMSE with a score <24 [23]. Information derived from a series of diagnostic evaluations, except for the CDT in the neuropsychological test battery, was reviewed by a team of experienced geriatricians at a case conference, and all the participants were categorized into five groups: normal elderly (NE), MCI, AD/senile dementia of Alzheimer's type (AD/SDAT) and mixed dementia, VD, and unclassified demented. The distribution of subjects by diagnostic category and sex is shown in table 2. Consensus diagnosis was made at the conference, using the Diagnostic and Statistical Manual of Mental Disorders Revised Third Edition (DSM-III-R) for dementia [24] as well as the National Institute of Neurological and

Table 2. Participant characteristics

	Whole sample	Nondemented		Demented		
		NE	MCI	SDAT[1]	VD	others
n; M/F	219; 75/144	41; 10/31	48; 21/27	102; 33/69	14; 6/8	14; 5/9
Age	75.1 (6.7)	72.7 (6.3)	74.7 (6.2)	76.0 (6.7)	76.9 (6.0)	74.9 (8.3)
Education	10.2 (2.8)	10.2 (1.8)	11.5 (3.7)	9.7 (2.4)	10.3 (3.0)	8.5 (2.5)
MMSE	24.4 (5.0)	28.4 (1.8)	27.2 (2.1)	22.2 (5.1)	20.5 (5.8)	20.6 (3.7)
CDT (Sunderland)	7.1 (2.4)	9.2 (1.1)	7.8 (2.1)	6.3 (2.3)	5.6 (2.5)	5.4 (2.3)
CDT (Rouleau)	7.0 (2.3)	8.7 (1.0)	8.0 (1.2)	6.3 (2.5)	5.6 (2.2)	5.6 (2.2)
CDT (Cahn)	6.1 (2.8)	8.4 (1.4)	7.1 (2.0)	5.2 (2.8)	4.1 (2.5)	4.1 (2.4)

Figures indicate means, with SD in parentheses.
[1] SDAT includes SDAT, AD and mixed dementia.

Communicative Disorders and Stroke and the Alzheimer's Disease and Related Association Work Group (NINCDS-ADRDA) criteria for probable AD to determine patients with AD/SDAT [25], and using the National Institute of Neurological and Communicative Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neuroscience Work Group (NINCDS-AIREN) criteria for probable VD to determine VD patients [26]. Mixed dementia was diagnosed as probable mixed dementia, when there was a clinical indication that dementia was likely to be attributable to both conditions. In this study, patients with mixed dementia were incorporated into the AD/SDAT group as stated above. The diagnosis of MCI was made according to the following criteria: (1) not demented, (2) subjective memory complaint, (3) normal general cognitive functioning assessed by the MMSE (score, ≥24), (4) objective memory impairment and/or impairment in other cognitive domains as evidenced by scores >1.5 SD below the age-appropriate mean score of at least one or more neuropsychological tests examined, (5) autonomy in the basic activities of daily living [7, 27].

Measurements

The subjects were given a blank piece of paper and asked to follow a two-step instruction: 'First, draw a 10-cm diameter clock face with all numbers on it. Second, put hands on the clock to make it read 10:10.' The CDT was scored by a psychologist according to the rating scales of Sunderland et al. [28], Rouleau et al. [29] and Cahn et al. [30]. These three sets of scoring criteria were chosen because they are characterized by concrete scoring instructions with presentations of actual error types, unlike other scoring methods, which contain vague or equivocal expressions in their scoring criteria. The psychologist who was the CDT rater of the present study was not given any information about the participant, including performance on other cognitive tests or clinical diagnosis.

The scoring methods used in the present study are as follows. (1) The CDT by Sunderland et al. [28] (Appendix 1): this method is based on the assumption that the representation of the hands is the first and solely affected item (score 6–10 points), and additional errors in the representation of numbers and the clock face occur later (score 1–5 points), so that a 10-point scale is used, with higher numbers indicating better performance. (2) The CDT by Rouleau et al. [29] (Appendix 2): three components of the drawing (integrity of the clock face, 0–2 points; presence and sequencing of the numbers, 0–4 points, and presence and sequencing of the hands, 0–4 points) are independently assessed. The scoring method supplies 0–10 points, with higher numbers indicating better performance. (3) The CDT by Cahn et al. [30] (Appendix 3): this is considered to be a modified version of the method by Rouleau et al. [29]. The difference in the Cahn scoring method is that while Rouleau's scoring method is regarded as a quantitative scoring of 0–10 points, the administrator notes the presence of qualitative errors shown in Appendix 3 and adds the error numbers up as a qualitative score with a maximum number of 8. The global CDT score is calculated by subtracting the qualitative score from the quantitative score. A 10-point scale is used, with higher numbers indicating better performance.

Statistical Analysis

All statistical analyses were performed using SPSS 11.0J. for Windows. Differences in age and years of education among the diagnostic groups were tested using the Kruskal-Wallis test. To examine the relationships between the CDTs and other variables (age, years of education, GDS, MMSE score), correlations and their p values were calculated using the Spearman rank order correlation coefficients. Distributions of Cahn's qualitative errors were examined using χ^2 analyses and Ryan's procedure for multiple comparisons.

Results

The five diagnostic groups shown in table 2 did not differ in terms of age (p = 0.0811). As for the educational years, except for the unclassified demented group, the four definite diagnostic groups did not differ (p = 0.0183). Distributions of the MCI subtypes are shown in table 3. The three groups (amnestic, single nonmemory and multiple domains) did not differ in age and years of education (p = 0.8623, 0.3575, respectively).

Table 3. MCI characteristics

	MCI		
	amnestic	single nonmemory	multiple domains
n; M/F	10; 2/8	10; 4/6	28; 15/13
Age	74.6 (7.2)	74.0 (7.2)	75.0 (5.6)
Education	11.5 (3.0)	9.0 (3.5)	12.4 (3.0)
MMSE	26.8 (2.8)	28.9 (1.2)	26.9 (2.0)
CDT (Sunderland)	9.4 (0.5)	7.3 (1.8)	7.4 (2.3)
CDT (Rouleau)	9.3 (0.7)	7.5 (1.1)	7.7 (1.2)
CDT (Cahn)	9.3 (0.7)	6.1 (2.0)	6.7 (1.8)

Figures indicate means, with SD in parentheses.

Table 4. Correlation matrix

	Sunderland	Rouleau	Cahn
Age	0.146	0.182	0.180
Education, years	0.210	0.220	0.201
GDS	0.281	0.312	0.312
MMSE	0.459 *	0.492*	0.490*
Sunderland		0.836*	0.857*
Rouleau			0.979*

* $p < 0.001$.

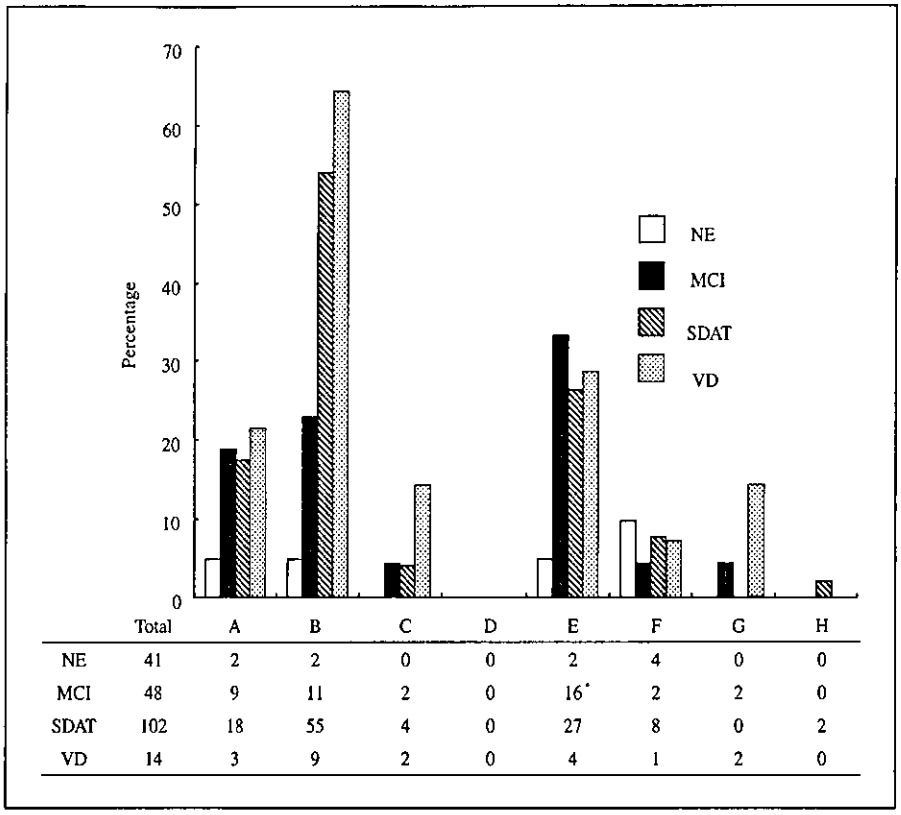
Correlations between the three CDT scores and other variables are presented in table 4. None of the CDTs correlated significantly with age, years of education, or the GDS score. However, all the CDTs correlated significantly with the MMSE score. Within the three CDTs, the scores correlated significantly with each other. In particular, the Cahn and Rouleau scores correlated with the highest correlation coefficient ($r = 0.979$, $p < 0.0001$).

We calculated the sensitivities and specificities with different CDT cutoff points to examine the discriminatory power of the CDTs for differentiating MCI and demented subjects from NE (table 5). The analyses demonstrated that Cahn's protocol had the highest discriminatory power at a cutoff point of 7, with a sensitivity of 74.7% and specificity of 75.6%.

Observed error types in the four definite diagnostic categories using Cahn's criteria for qualitative analysis are presented in figure 1. The four definite diagnostic groups did not differ in terms of age and years of education. The letters 'A' to 'H' represent Cahn's qualitative error types, which are described in Appendix 3. Regarding the distribution of error types in each group, only a few errors were noted in the NE group according to Cahn's criteria. In the MCI group, E (planning deficit) was the most frequent (16 of 48 cases), followed by B (conceptual deficit) and A (stimulus-bound response; 11 and 9 of 48 cases, respectively). In the AD group (SDAT/AD/MIXED), B was the most frequent (55 of 102 cases, 53.9%), followed by E (27 of 102 cases, 26.5%) and A (18 of 102 cases, 17.6%). In

Table 5. Sensitivities and specificities (%)

Cutoff point	Sunderland		Rouleau		Cahn	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
10	100.0	0.0	100.0	0.0	100.0	0.0
9	93.8	48.8	96.1	17.1	96.1	17.1
8	66.0	90.2	78.1	65.9	82.6	65.9
7	61.2	95.1	56.7	92.7	74.7	75.6
6	42.7	95.1	36.5	95.1	60.1	92.7
5	38.2	95.1	24.7	97.6	44.4	95.1
4	26.4	100.0	18.5	100.0	33.7	97.6
3	10.0	100.0	11.8	100.0	23.0	100.0
2	5.1	100.0	0.8	100.0	14.6	100.0
1	1.1	100.0	0.5	100.0	10.6	100.0



	Total	A	B	C	D	E	F	G	H
NE	41	2	2	0	0	2	4	0	0
MCI	48	9	11	2	0	16*	2	2	0
SDAT	102	18	55	4	0	27	8	0	2
VD	14	3	9	2	0	4	1	2	0

Fig. 1. Comparison of error types among the four diagnostic groups: NE, MCI, SDAT, (represents SDAT, AD and mixed dementia), and VD. The horizontal scale (A–H) represents Cahn's qualitative error types: A = Stimulus-bound response; B = conceptual deficit; C = perseveration; D = neglect of hemispace; E = planning deficit; F = nonspecific spatial error; G = numbers written on the outside of the clock, and H = numbers written counterclockwise. The bottom table is the matrix showing participants' actual number of errors in the four diagnostic groups.

the VD group, B was the most frequent (9 of 14 cases), followed by E and A (4 and 3 of 14 cases, respectively). The χ^2 analysis comparing the frequency of Cahn's error types made by the four diagnostic groups revealed that there was a significant effect of diagnosis in all error types ($p < 0.0001$).

Discussion

While various CDT scales are available for detecting dementia, few studies have examined the adequate scoring protocol and optimal cutoff point for screening MCI. With the aim of detecting dementia at an early stage, particularly MCI, the present study compares the three scoring methods, all of which were found to be independent of years of education and depression scale, which is in keeping with the findings by Shulman et al. [8]. In addition, the CDT scores determined using the three scoring methods correlated with MMSE scores with high statistical significance, as confirmed in previous studies [12, 31, 32]. In most of the former studies, the cutoff points of CDTs were provided with the aim of distinguishing a demented state or AD from normal cognition. Recently, Powlishta et al. [16], choosing 6 different scoring criteria of the CDT for comparing subjects without altering the cutoff points determined for dementia in each original CDT criterion, have reported that the CDT was a poor screening method for very mild dementia. The sensitivity and specificity for detecting MCI by the CDT obtained in the current study were satisfactory. The discrepancy between the results of the study by Powlishta et al. [16] and those of the current study may be due to different cutoff points. Comparison of sensitivities and specificities among the three CDT protocols revealed that the Cahn scale had the best discriminatory power at the cutoff point of 7. Thus, the results may indicate that Cahn's protocol is the most suitable method for screening MCI in general practice. As shown in table 3, the analyses based on MCI subtypes suggest that subjects with amnestic MCI cannot be screened by the cutoff point we consider optimal for differentiating MCI subjects from normal individuals. This may simply imply that MCI subjects without deficits in the cognitive domain do not lose scores on the CDT, but we need further investigation to warrant this notion, given the limited number of participants included in this study. However, as we acknowledge the significance of including MCI subtypes other than the amnestic type, we believe that the

present findings would provide useful information for clinicians for screening subjects at risk of dementia in earlier stages.

We also examined the error types in MCI subjects using Cahn's qualitative criteria and compared them with the results in the NE and subjects with dementia. In what follows, impairment underlying each type of frequent error is disclosed [11, 17, 30]:

(A) stimulus-bound response: disturbance of inhibition in executive control functioning, an aspect of the frontal cortical function;

(B) conceptual deficit: loss of semantic memory usually evoked by the word 'clock';

(C) perseveration: an aspect of frontal dysfunction, and

(E) planning deficit: suggested to be associated with visuospatial constructional/frontosubcortical dysfuntion.

As shown in figure 1, error type E (planning deficit) could be a distinctive feature of MCI, which is represented by imprecise gaps before 12, 3, 6, or 9 of the numbers arranged in the clock face, or by clock hands drawn not from the center of the clock face. This type of error is considered to represent the inability to form a strategy for drawing a clock, presumably due to frontosubcortical dysfunction. The frequency of conceptual deficit in the MCI group was significantly lower than that in the SDAT/AD/MIXED and VD group. The conceptual deficit reflects a loss or deficit in accessing knowledge of the attributes, features, and meaning of a clock, and this category includes misrepresentation of the clock itself and the time on the clock [30]. Eleven out of 41 MCI subjects made this type of error; the difference in frequency between the MCI and the other groups' subjects did reach statistical significance. The SDAT/AD/MIXED and VD groups made this type of error with a frequency of 50% or above, which was statistically higher than that in the MCI and NE group, and that in the MCI being again higher than that in the NE group. Further investigations with increased number of subjects may clarify detailed characteristics of the clock drawing in MCI or its subgroups.

Between SDAT and VD subjects, no significant difference in error types was identified. This might be because of the relatively small number of patients in the VD subgroup in the current study, which may thus have influenced the statistical analysis. The previous study showed that the frequency of spatial and/or planning deficit was significantly higher in patients with mild VD than mild AD, and in patients with moderate VD, the frequency of graphic difficulties was significantly higher than in moderate AD [17]. These assumptions derived from the observations in this study may help to guide and benefit from future studies with larger numbers of subjects.

Although the CDT cannot be used solely for clinical diagnoses, the CDT, as a simple screening method, provides objective and graphic documentation of cognitive deficits that can be shared by a wide range of clinicians. In conclusion, among the three scales examined in this study, the Cahn scoring method at a cutoff point of 7 is the most likely indicator for MCI. Petersen et al. [7] recommended the CDT using Cahn's protocol as an optional instrument for brief cognitive assessment, as an addition to a general cognitive screening test, e.g. MMSE. We believe that the results obtained in the current study provide important evidence of the validity of the CDT as one of the useful screening method for discriminating MCI from normal cognition.

Acknowledgments

The authors would like to thank all the individuals who helped with data collection and coordination of the study.

Appendix 1
Sunderland's protocol: a priori criteria for evaluating clock drawings

10–6	Drawing of clock face with circle and numbers is generally intact
10	Hands are in correct position
9	Slight errors in placement of the hands
8	More noticeable errors in the placement of hour and minute hands
7	Placement of hands is significantly off course
6	Inappropriate use of clock hands
5–1	Drawing of clock face with circle and numbers is not intact
5	Crowding of numbers at one end of the clock or reversal of numbers
4	Further distortion of numbers sequence; integrity of clock face is now gone
3	Numbers and clock face no longer obviously connected in the drawing; hands are not present
2	Drawing reveals some evidence of instructions being received but only a vague representation of a clock
1	Either no attempt or an uninterpretable effort is made

10 = Best, and 1 = worst.

Appendix 2
Rouleau's protocol: the score is calculated by a sum of three components (I, II, III)

I Integrity of the clock face (maximum: 2 points)
- 2 Present without gross distortion
- 1 Incomplete or some distortion
- 0 Absent or totally inappropriate

II Presence and sequencing of the numbers (maximum: 4 points)
- 4 All present in the right order and at most minimal error in the spatial arrangement
- 3 All present but errors in spatial arrangement
- 2 Numbers missing or added but no gross distortions of the remaining numbers; numbers placed in counterclockwise direction or all present but gross distortion in spatial layout (i.e. hemineglect, numbers outside the clock)
- 1 Missing or added numbers and gross distortions
- 0 Absence or poor representation of numbers

III Presence and placement of the hands (maximum: 4 points)
- 4 Hands are in correct position and the size difference is respected
- 3 Slight errors in the placement of the hands or no representation of size difference between the hands
- 2 Major errors in the placement of the hands
- 1 Only one hand or poor representation of two hands
- 0 No hands or perseveration on hands

Appendix 3
Cahn's protocol: the global score is calculated by subtracting qualitative score (II) from quantitative score (I)

I Quantitative CDT score = maximum 10 points: assesses the presence and correctness of the clock; the clock face (0–2 points), the placement of the hands (0–4 points) and the placement of the numbers (0–4 points)

II Qualitative CDT score = maximum 8 points: summary of the following errors
- 1 Stimulus-bound response: the tendency of the drawing to be dominated or guided by a single stimulus
- 2 Conceptual deficit: this error type reflects a loss or deficit in accessing knowledge of the attributes, features and meaning of a clock
- 3 Perseveration: the continuation or the recurrence of activity without an appropriate stimulus
- 4 Neglect of left hemispace: all attributes of the clock are written on the right half of the clock face
- 5 Plannning deficit: this error type is represented by gaps before 12, 3, 6 or 9
- 6 Nonspecific spatial error: a deficit in the spatial layout of numbers, without any specific pattern in spatial disorganization
- 7 Numbers written on the outside of the clock: numbers written either around the perimeter of the circle or the circle itself
- 8 Numbers written counterclockwise: arrangement of the numbers with '12' at the top of the clock face and then continuing around in a counterclockwise fashion

References

1 Schramm U, Berger G, Muller R: Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population. Int J Geriatr Psychiatry 2002;17:254–260.
2 Juby A, Tench S, Baker V: The value of clock drawing in identifying executive cognitive dysfunction in people with a normal Mini-Mental State Examination score. CMAJ 2002;167:859–864.
3 Flicker C, Ferris SH, Reisberg B: Mild cognitive impairment in the elderly: Predictors of dementia. Neurology 1991;41:1006–1009.
4 Petersen RC, Smith GE, Waring SC: Mild cognitive impairment: Clinical Characterization and outcome. Arch Neurol 1999;56:303–308.
5 Petersen RC, Doody R, Kurz A: Current concepts in mild cognitive impairment. Arch Neurol 2001;58:1985–1992.
6 Larrieu S, Letenneur L, Orgogozo JM: Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. Neurology 2002;59:1594–1599.
7 Petersen RC, Stevens JC, Ganguli M: Early detection of dementia: Mild cognitive impairment (an evidence-based review) report of the quality standards subcommittee of the American Academy of Neurology. Neurology 2001;56:1133–1142.
8 Shulman K, Shedletsky R, Silver IL: The challenge of time: Clock-drawing and cognitive function in the elderly. Int J Geriatr Psychiatry 1986;1:135–140.
9 Brodaty H, Moore CM: The clock drawing test for dementia of the Alzheimer's type: A comparison of three scoring methods in a memory disorders clinic. Int J Geriatr Psychiatry 1997;12:619–627.
10 Lam LC, Chiu HF, Ng KO: Clock-face drawing, reading and setting tests in the screening of dementia in Chinese elderly adults. J Gerontol B Psychol Sci Soc Sci 1998;53:353–357.
11 Royall DR, Mulroy AR, Chiodo LK: Clock drawing is sensitive to executive control: A comparison of six methods. J Gerontol B Psychol Sci Soc Sci 1999;54:328–333.
12 Richardson HE, Glass JN: A comparison of scoring protocols on the Clock Drawing Test in relation to ease of use, diagnostic group, and correlations with Mini-Mental State Examination. J Am Geriatr Soc 2002;50:169–173.
13 Shulman KI, Gold DP, Cohen CA: Clock drawing and dementia in the community: A longitudinal study. Int J Geriatr Psychiatry 1993;8:487–496.
14 Borson S, Brush M, Gil E: The Clock Drawing Test: Utility for dementia detection in multiethnic elders. J Gerontol A Biol Sci Med Sci 1999;54:534–540.
15 Manos P: Ten-point clock test sensitivity for Alzheimer's disease in patients with MMSE scores greater than 23. Int J Geriatr Psychiatry 1999;14:454–458.
16 Powlishta KK, Dras V, Stanford A: The Clock Drawing Test is a poor screening for very mild dementia. Neurology 2002;59:898–903.
17 Kitabayashi Y, Ueda H, Narumoto J: Qualitative analyses of clock drawings in Alzheimer's disease and vascular dementia. Psychiatry Clin Neurosci 2001;55:485–491.

18 Ueda H, Kitabayashi Y, Narumoto J: Relationship between Clock Drawing Test performance and regional cerebral blood flow in Alzheimer's disease: A single photon emission computed tomography study. Psychiatry Clin Neurosci 2002;56:25–29.
19 Yesavage JA: The use of self-rating depression scales in the elderly; in Poon LW, (eds): Clinical Memory Assessment of Older Adults. Washington, American Psychological Association, 1986, pp 213–217.
20 Mohs RC: The Alzheimer's disease assessment scale: An instrument for assessing treatment efficacy. Psychopharmacol Bull 1983;19:448–450.
21 Wechsler D: Wechsler Adult Intelligence Scale. New York, Psychological Corp, 1955.
22 Stroop JR: Studies of interference in serial verbal reaction. J Exp Psychol 1935;18:643–662.
23 Folstein MF, Folstein SE, McHugh PR: Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
24 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 3, revised. Washington, American Psychiatric Association, 1987.
25 Mckhann G, Drachman D, Folstein M: 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.
26 Roman GC, Tatemichi TK, Erkinjuntti T: Vascular dementia: Diagnostic criteria for research studies: Report of the NINDS-AIREN International workshop. Neurology 1993;43:250–260.
27 Smith GE, Petersen RC, Parisi JE: Definition, course, and outcome of mild cognitive impairment. Aging Neuropsychol Cogn 1996;3:141–147.
28 Sunderland T, Hill JL, Mellow AM: Clock drawing in Alzheimer's disease: A novel measure of dementia severity. J Am Geriatr Soc 1989;37:725–729.
29 Rouleau I, Salmon DP, Butters N: Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. Brain Cogn 1992;18:70–87.
30 Cahn DA, Salmon DP, Monsch AU: Screening for dementia of the Alzheimer type in the community: The utility of the Clock Drawing Test. Arch Clin Neuropsychol 1996;11:529–539.
31 Shulman KI: Clock-drawing: Is it the ideal cognitive screening test? Int J Geriatr Psychiatry 2000;15:548–561.
32 Moretti R, Torre P, Antonello RM: Ten-point clock test: A correlation analysis with other neuropsychological tests in dementia. Int J Geriatr Psychiatry 2002;17:347–353.

The Relationship Between Functional Disability and Depressive Mood in Japanese Older Adult Inpatients

Joji Onishi, MD, Hiroyuki Umegaki, MD, PhD, Yusuke Suzuki, MD, PhD, Katsuhiko Uemura, MA, Masafumi Kuzuya, MD, PhD, and Akihisa Iguchi, MD, PhD

ABSTRACT

Depression is commonly found in older adult patients and is often associated with handicaps. The authors administered the Comprehensive Geriatric Assessment (CGA), including basic activities of daily living (BADL), instrumental activities of daily living (IADL), Mini-Mental State Examination (MMSE), Geriatric Depression Scale (GDS)-15, and a socioenvironmental questionnaire to 198 patients who were admitted to Nagoya University Hospital, to examine the relationship between depressive mood and various physical and socioenvironmental outcomes. The overall GDS-15 score was correlated with the BADL and IADL. The factor analysis extracted 4 factors from the GDS-15 subscales. The factors labeled "loss of morale and hope" and "memory loss and reduction of social activity" were highly correlated with both ADLs, social variables, and the MMSE score. The results reveal that factor analysis of GDS-15 will help in understanding the etiology of depressive mood, thereby contributing to better therapeutic approaches. (*J Geriatr Psychiatry Neurol* 2004; 17:93-98)

Keywords: depressive mood; Geriatric Depression Scale; Comprehensive Geriatric Assessment; factor analysis

Depression is one of the most insidious problems faced by older adults, and its incidence is increasing with the growth of an aging population. Koenig and Blazer reported that the prevalence of major depression was about 1% among community-dwelling older adults and that less severe depressive disorder was present in over 25%.[1] Moreover, they reported that the rate of major depressive disorder in older adult hospitalized patients with illness was more than 10 times greater than that of the unhospitalized aging population. Depression is not only psychologically traumatic but also quite costly[2] because it is related to psychosomatic symptoms resulting in a higher frequency of examination and prescription of drugs. Furthermore, depression also decreases the morale of older people and increases the risk of being housebound. Although it is very important to adequately diagnose and treat depression in its early stage, it often remains unrecognized or untreated.[3] One of the main reasons for this is that depressive symptoms often resemble those of the aging process itself, such as progressive cognitive deterioration or physical disabilities.[4]

The Geriatric Depression Scale (GDS) is a self-administered questionnaire with 30 items[5] and is recommended by the Royal College of Physicians and British Geriatrics Society as a valid screening method for depression in older adults.[6] A short form of the GDS (GDS-15) was developed later[7] and was translated into Japanese.[8] The validity and reliability of the GDS-15 have been confirmed in both community and hospital settings.[9-11] Several studies have subjected the GDS-15 data to a factor analysis, which is a statistical technique to analyze interrelationships within a set of variables, resulting in the construction of a few hypothetical variables. To our knowledge, however, there has been only 1 study involving factor analysis of the Japanese version of the GDS-15, reported by Schreiner et al in poststroke patients.[12] In addition, there have been few studies demonstrating the relationship between GDS-15 factor loading and disabilities in the older population.

Received September 4, 2003. Received revised December 22, 2003. Accepted for publication February 25, 2004.

From Department of Geriatrics, Nagoya University Graduate School of Medicine, Japan (Drs Onishi, Umegaki, Suzuki, Kuzuya, and Iguchi), and Department of Interpersonal Communication, Aichi Shukutoku University Graduate School of Communication Studies (Mr Uemura).

Address correspondence to: Joji Onishi, MD, Department of Geriatrics, Nagoya University Graduate School of Medicine, 65 Tsuruma-Cho, Showa-Ku, Nagoya, Aichi, 466-8550, Japan.

DOI: 10.1177/0891988704264738

The GDS-15 is included as one of the components in the Comprehensive Geriatric Assessment (CGA), a tool developed in the late 1980s[13,14] to assess not only medical conditions but also overall functional status with respect to physical, psychological, and social problems of the older adults.

Although it is well known that depressive mood is often associated with functional disabilities, the mechanism by which the disabilities cause depressive mood in the older adults remains unclear. We hypothesized that some variables associated with functional disability may be associated with depressive mood. Therefore, we investigated the relationship between depressive mood and physical health and socioenvironmental variables in older adult inpatients. In addition, we attempted to clarify the structure of depression by performing a factor analysis of the GDS-15.

METHODS

Subjects

Among 355 consecutive patients aged 65 and older (mean age ± SD: 77.3 ± 6.8) who were admitted to Nagoya University Hospital between July 1998 and August 2001, patients who were admitted to nongeriatric wards were not included due to the absence of experienced CGA assessment team in the wards. Also, patients with communication impairments due to problems such as severe dementia or consciousness disturbance and patients under intensive care were not included in the study. If a patient was admitted more than once during the study period, only the data from the first admission was used for this analysis. As a result, 198 older adult patients in total were included in the study.

Measurements

The CGA was administered within a week after admission. The CGA included height; weight; Body Mass Index (BMI); blood pressure; basic activities of daily living (BADL), which were measured with the Barthel Index[15]; instrumental activities of daily living (IADL) using Lawton's scale[16]; Mini-Mental State Examination (MMSE)[17]; GDS-15; hearing ability and vision; communicative competence; and living environment including socioeconomic status. We scored IADL by 5 items (IADL-5), excluding food preparation, housekeeping, and laundry items from the Lawton's scale because the study samples included male patients, who did not normally perform these activities. The low scores of BADL and IADL-5 indicate greater functional disability. The GDS-15 is scored so that higher scores indicate a greater degree of depressive mood. The recent research clarified that the sensitivity of the GDS-15 was 97.3% and the specificity was 95.9% for screening major and minor depression when the cut-off score was set at 6/6+ in the Japanese geriatric population.[18] Socioenvironmental status was assessed by Ozawa's scale,[19] which includes items on economic, marital, family status, and the relationship between the patient and his or her family. The GDS-15 was self-administered by the patient. The attending nurse collected all other information by interview and/or assessment.

Statistical Analysis

Correlation coefficients were calculated by Pearson's method for parametric data and Spearman's for nonparametric data. We used the chi-square test with Yates correction and Fisher's exact test for categorical comparisons of the data. Differences in the means of continuous measurements between genders were tested using the Student's t test. In addition, after nonparametric data in the CGA were categorized into 2 groups (subjects with and those without a problem with respect to each parameter measured), the means of the continuous measurements between the groups were also compared by Student's t test. The internal consistency of the GDS-15 was calculated with Cronbach's alpha. Principal component analysis for the GDS-15 was performed with an eigenvalue of 1.0 or more as the extraction criterion, and factors were identified after Varimax rotation. The factor score, which shows the power of a factor's contribution, was calculated by regression method, which cumulated factor loadings of all items of GDS-15. In the present study, a higher score indicates a greater contribution of the factor to depressive mood. Differences in continuous variables among the disease groups were determined by 1-way analysis of variance (ANOVA). Tukey's test was used for multiple comparisons when homoskedasticity was assumed by Levene's method, and Dunnet's test was performed when homoskedasticity was not assumed. Multiple regression analysis, using the equation-building method with the variables of significant measures detected in the univariate analysis, was conducted to identify the variables contributing to GDS-15 scores. Values of $P < .05$ were considered to indicate statistical significance; all tests were 2-tailed. All statistical analyses were performed on a personal computer with the statistical package SPSS for Windows (Version 11.0 SPSS, Chicago).

RESULTS

Table 1 reports CGA variables for all patients, according to their diagnostic category. The mean GDS-15 score of all patients was 5.9 ± 3.8 SD, and 39.3% of the patients had scores above 6. The homoskedasticities were assumed in age, systolic blood pressure, BADL, IADL-5, and GDS-15, but not in BMI or MMSE. Significant intergroup differences were observed on the BADL and IADL-5, but not in BMI, MMSE, or GDS-15. The BADL score in patients with diabetes mellitus was higher than that in patients with collagen disease ($P = .005$), and the IADL-5 score in patients with diabetes mellitus was higher than that in patients

Table 1. Mean Values ± Standard Deviation of Comprehensive Geriatric Assessment (CGA) Variables by Admitting Diagnosis

Admitting Diagnosis	*n (%)*	*Age*	*BMI (kg/m²)*	*sBP (mm Hg)*	*BADL*	*IADL-5*	*MMSE*	*GDS-15*	*GDS > 6*
Neurological disease	40 (20%)	76.5 ± 6.6	20.9 ± 3.9	128.5 ± 23.7	16.9 ± 4.1	4.0 ± 1.3	24.9 ± 4.5	6.3 ± 37	42%
Cardiovascular disease	36 (18%)	77.7 ± 8.4	23.5 ± 3.8	132.8 ± 20.0	18.0 ± 3.7	4.0 ± 1.3	26.0 ± 4.3	5.7 ± 4.0	38%
Diabetes mellitus	34 (17%)	74.2 ± 5.3	23.5 ± 3.1	138.3 ± 19.4	19.0 ± 3.0*	4.5 ± 0.9*	26.6 ± 3.5	4.6 ± 3.5	27%
Psychological disease	20 (10%)	78.5 ± 6.5	20.0 ± 3.4	138.5 ± 22.3	17.9 ± 3.0*	3.1 ± 1.9*	22.4 ± 4.9	7.6 ± 3.8	15%
Gastroenterological disease	14 (7%)	78.9 ± 6.8	21.1 ± 4.8	132.3 ± 14.1	18.2 ± 3.2*	4.2 ± 0.9	25.9 ± 3.8	5.9 ± 4.7	64%
Collagen disease	12 (6%)	77.7 ± 5.1	21.6 ± 4.0	133.5 ± 20.6	14.2 ± 6.5*	3.3 ± 1.7	23.7 ± 4.7	5.4 ± 2.7	17%
Infectional disease	11 (6%)	83.1 ± 4.7	19.9 ± 3.0	122.0 ± 15.7	19.5 ± 0.8	4.8 ± 0.4	27.3 ± 2.8	2.8 ± 1.8	0%
Others	31 (16%)	78.0 ± 7.6	20.7 ± 3.5	142.4 ± 29.0	18.0 ± 3.9	4.3 ± 1.1	26.1 ± 4.1	6.3 ± 4.0	43%
Total	198 (100%)	77.3 ± 6.8	21.9 ± 3.8	133.9 ± 21.7	17.8 ± 3.8	4.1 ± 1.3	25.5 ± 4.3	5.9 ± 3.8	39%

Note: BMI = body mass index, sBP = systolic blood pressure, BADL = basic activities of daily living, IADL = instrumental activities of daily living, MMSE = Mini-Mental State Examination, GDS = Geriatric Depression Scale.

*$P < .05$.

Table 2. Principal Components (Varimax) Factor Analysis of the Geriatric Depression Scale-15

Items	*Factor 1 Unhappiness*	*Factor 2 Apathy and Anxiety*	*Factor 3 Loss of Hope and Morale*	*Factor 4 Memory Loss and Reduction of Social Activity*
1. Satisfied	***0.708***	0.270	0.061	−0.266
2. Dropped activities	0.058	***0.648***	0.350	−0.020
3. Emptiness	0.299	***0.621***	−0.134	0.179
4. Often bored	0.151	***0.675***	0.140	0.233
5. In good spirits	***0.627***	0.216	0.129	0.216
6. Afraid something bad will happen	0.336	***0.572***	0.163	−0.100
7. Feels happy	***0.769***	0.027	0.128	0.101
8. Often feels helpless	−0.186	***0.536***	0.493	0.013
9. Prefers to stay in	0.009	0.095	0.385	***0.445***
10. More problems with memory than most	0.082	0.074	0.043	***0.805***
11. Wonderful to be alive	***0.553***	0.077	0.458	0.033
12. Feels worthless	0.348	0.108	***0.605***	0.242
13. Full of energy	0.061	0.063	***0.753***	0.002
14. Feels situation is hopeless	0.270	0.235	***0.679***	0.090
15. Most people better off than self	***0.487***	0.396	0.013	0.368
Explained variance	2.4	2.2	2.2	1.2
Cumulative percentage of variance explained	16.6	31.5	46.3	54.8

Note: The factor score was calculated by regression method, which cumulated factor loadings of all items of GDS-15. Loadings in italic bold indicate those selected to define the factor.

with psychological disease ($P = .009$). The patients with psychological disease showed the highest mean score of GDS-15, (7.6 ± 3.8 SD). No significant intersex difference was observed in all parameters examined. Antidepressants had been administered to 7.2% of all patients, and to 9.0% of the patients with a GDS-15 score greater than 6.

The internal consistency of GDS-15 was found to be satisfactory, Cronbach's alpha being .83. Factor analysis of GDS-15 extracted 4 factors, whose loading values are shown in Table 2. The cumulative percentage of variance explained was 57.3%. Factor 1 represented "unhappiness," which included the items satisfied, in good spirits, feels happy, wonderful to be alive, and most people better off than self. Factor 2, "apathy and anxiety," was made up of the items, dropped activities, emptiness, often bored, afraid something bad will happen, and often feels helpless. Factor 3, "loss of hope and morale," included the items feels worthless, full of energy, and feels situation is hopeless. Finally, factor 4, "memory loss and reduction of social activity," included the items prefers to stay in and more problems with memory than most.

Table 3. Correlation Between Geriatric Depression Scale-15, Extracted Factors, and Parametric Data

Measure	*GDS-15*	*Factor 1 Unhappiness*	*Factor 2 Apathy and Anxiety*	*Factor 3 Loss of Hope and Morale*	*Factor 4 Memory Loss and Reduction of Social Activity*
Age	0.123	−0.001	−0.108	0.250**	0.166*
BMI	−0.141	0.006	−0.135	−0.121	−0.036
sBP	−0.038	−0.260	−0.040	−0.009	−0.101
BADL	−0.168*	−0.033	−0.044	−0.191*	−0.055
IADL-5	−0.201**	−0.076	0.023	−0.235**	−0.066
MMSE	−0.151*	−0.034	0.050	−0.167*	−0.214**

Note: Pearson's rho used for correlations. BMI = body mass index, sBP = systolic blood pressure, BADL = basic activities of daily living, IADL = instrumental activities of daily living, MMSE = Mini-Mental State Examination.

*$P < .05$. **$P < .01$.

Pearson's coefficients of continuous variables are shown in Table 3. The total GDS-15 score had a significant negative correlation with IADL-5 ($r = -.201$, $P = .005$), BADL ($r = -.168$, $P = .021$), and MMSE ($r = -.151$, $P = .034$). However, there was no significant relationship between the GDS-15 score and age, BMI, or systolic blood pressure.

The score of factor 3 (loss of hope and morale) correlated positively with age and negatively with IADL-5, BADL, and MMSE scores, whereas factor 4 (memory loss and reduction of social activity) showed a significant positive correlation with age and a significant negative correlation with MMSE score. However, there was no significant relationship between the scores of factor 1

Table 4. Relationship of Nonparametric Data in Comprehensive Geriatric Assessment With the Geriatric Depression Scale-15 and Extracted Factors

Measurement	*Percent With Problem*	*Spearman's ρ With GDS-15*	*t Test for Mean Score GDS-15*	*Factor 1 Unhappiness*	*Factor 2 Apathy and Anxiety*	*Factor 3 Loss of Hope and Morale*	*Factor 4 Memory Loss and Reduction of Social Activity*
Gender (male/female)	–	–	NS	NS	NS	–0.22/0.17**	NS
BADL (with/without problem)							
Grooming	7.1%	–	NS	NS	NS	0.75/–0.08**	NS
Feeding	8.1%	–0.087	NS	NS	NS	NS	NS
Bowel continence	12.2%	–0.062	NS	NS	NS	NS	NS
Using toilet	14.2%	–0.122	NS	NS	NS	NS	NS
Ambulation	16.8%	–0.102	NS	NS	NS	0.31/–0.09*	NS
Chair/bed transfer	16.8%	–0.142	7.1/5.6*	NS	NS	NS	NS
Dressing	17.8%	–0.122	NS	NS	NS	NS	NS
Bladder control	19.8%	–0.097	NS	NS	NS	NS	NS
Bathing	25.0%	–	6.9/5.5*	NS	NS	0.27/–0.12*	NS
Using staircase	29.9%	–0.271*	7.4/5.2**	NS	NS	0.33/–0.17**	NS
IADL(with/without problem)							
Going outside	10.4%	–	NS	NS	–0.41/0.10*	NS	NS
Using telephone	11.4%	–	NS	NS	NS	NS	NS
Managing money	20.3%	–	NS	NS	NS	NS	NS
Medication	37.1%	–	NS	NS	–0.14/0.15*	NS	NS
Shopping	39.4%	–	NS	NS	NS	0.21/–0.15*	NS
Physical (with/without problem)							
Seeing	23.1%	–0.141	NS	NS	NS	NS	NS
Hearing	23.0%	–0.091	NS	NS	NS	NS	NS
Communication	7.0%	–0.152*	8.2/5.7*	NS	NS	NS	0.51/–0.48*
Social							
Economic status (dependent/ independent)	–	–0.163*	NS	NS	NS	NS	NS
Marital status (with/without spouse)	–	–0.148*	NS	NS	NS	0.20/–0.21**	NS
Familial status (alone/not alone)	–	–0.136	7.2/5.6*	0.50/–0.08*	NS	NS	NS
Family relation (with/without interaction)	–	–0.220*	NS	NS	NS	0.71/–0.03*	NS

Note: NS = not significant. t-test for mean score compared between 2 groups with or without problem for each item.

*$P < .05$. **$P < .01$. Dashes indicate not calculated because the items have less than 3 alternatives

(unhappiness) or factor 2 (apathy and anxiety) and other CGA variables.

The patients were divided into 2 groups depending on their score for CGA variables. Then we compared the difference between the GDS-15 factor scores and these 2 groups using Student's *t* test. The correlations of nonparametric data with the score of GDS-15 and the extracted factors are shown in Table 4. The GDS-15 score had a significant negative correlation with BADL (using staircase), communicative ability, economic and marital status, and family relationship. Patients having problems in using the staircase, bathing, chair/bed transfer, and communication showed a significantly higher GDS-15 score than the patients without these problems ($P < .001$, $P = .041$, $P = .034$, $P = .028$, respectively). Also, patients living alone showed a significantly higher GDS-15 score than those not living alone ($P = .043$). The statistical analysis revealed that the score of factor 3 (loss of hope and morale) was significantly higher among women ($P = .007$). Factor 3 had a much stronger relationship with some variables of BADL and IADL-5, such as grooming, using staircase, ambulation, bathing, and shopping, than it did with other factors. On the other hand, factor 2 (apathy and anxiety) was inversely correlated with going outside and managing medication.

Multiple regression analysis was performed to predict the score of GDS-15 with significant variables, which were using stairs, bathing, communicative ability, economic status, marital status, familial status, and the total score of MMSE. This analysis elicited a model with an adjusted R^2 of .144 ($P < .001$) (Table 5).

DISCUSSION

The mean GDS-15 score in this study was 5.9, which was higher than those in previous studies. In a recent study of 1343 Japanese community-dwelling older adults, the mean GDS-15 score was 2.0 and 23.7% scored 6 or higher.[20] Meanwhile, Patrick et al reported that the mean score of hospitalized patients in their geriatric rehabilitation unit was 3.8 ± 2.8 SD.[21] The higher GDS-15 scores obtained in this study may imply that worsening medical conditions resulting in admission to the hospital relate to increased depressive symptoms. In particular, the neurological disease group showed the highest mean GDS-15 score, which is in line with findings in previous studies that depression

Table 5. Coefficients of Regression Model for Geriatric Depression Scale-15

Variable	ß	*Standardized* ß	*T*	*P Value*
Using stairs	−2.48	−0.48	−4.27	< .001
Bathing	2.59	0.29	2.44	< .001
Communicative ability	−0.57	−0.04	−0.558	.016
Economic status	−0.48	−0.07	−0.917	.577
Marital status	−0.34	−0.09	−1.25	.360
Familial status	−1.02	−0.17	−2.17	.211
MMSE	−0.04	−0.04	−0.55	.584

Note: MMSE = Mini-Mental State Examination. GDS-15 = −2.48 × (Using stairs) + 2.59 × (Bathing) − 0.57 × (Communication) − 0.48 × (Economic status) − 0.34 × (Marital status) − 1.02 × (Family status) − 0.04 × MMSE. Total adjusted $R^2 = 0.144$, $P < .001$.

frequently occurs after stroke.[10,22,23] In the present study, antidepressants were administered to only 9.0% of the patients who had a GDS-15 score of greater than 6, which supports claims that depression is overlooked by clinicians, or is not treated adequately.[4]

The results of this study are consistent with previous findings that physical disabilities relate to depressive symptoms.[24-27] In the present study, the GDS-15 score was negatively correlated with the BADL and IADL. Three BADL items in particular, using staircase, chair/bed transfer, and bathing, had strong negative correlations with the GDS-15 score. These results indicate that loss of lower body strength and impaired mobility may affect patient's mood. A possible explanation for the difference is that depressive mood may be associated with impaired abilities to maintain normality in life such as immobility, rather than the severity of disabilities.

We also found a weak but significantly negative correlation between the GDS-15 and MMSE scores. The findings of previous studies regarding the relationship between depression and the severity of dementia are varying, which may be attributable to differences in study design.[28] Although many investigators have reported a decrease in the frequency of depression in advanced dementia,[29,30] no such association was found in this study probably because the cognitive impairment of the patients in this study was rather mild with mean MMSE score of 25.5 ± 4.3 SD, and no patients with advanced dementia were included.

Liu et al reported that being female, older, and without spouse were related to depressive symptoms among Chinese older adults.[31] Our results did not demonstrate a significant relationship between the GDS-15 score and either gender or age, but a higher GDS-15 score was significantly related with economic dependence, absence of spouse, and poor family relationship particularly with "living alone."

Thus far, many researchers have reported on the factor analysis of GDS-15, but the relationship between the factors extracted and the physical, psychological, and socioenvironmental status of the older adults has not been extensively investigated. We found that factor 3, "loss of morale and hope," was highly related with BADL and IADL. Meanwhile, factor 4, "memory loss and reduction of social activity," was related with age and MMSE, although factor 1 (unhappiness) and factor 2 (apathy and anxiety) were not correlated with any of those parameters examined, which means they may be normal aspects of disabled state and hospitalization. Some investigators have reported that sense of loss or environmental change can induce depression in the aged.[32,33]

GDS-15 is often included in CGA, which is a useful tool to comprehensively assess older adult patients. The meta-analysis conducted by Stuck et al demonstrated that CGA was effective in improving mortality and in reducing hospitalization.[34] However, there have been few studies using CGA results to identify specific clinical strategies for patient care. The present study demonstrates that factor analysis of GDS-15 helps health care staffs establish better therapeutic strategies for depressive mood of older patients. For example, the present findings suggest that intervention to assist in coping with the functional impairment may decrease depressive symptoms in subjects suffering from them. However, pharmacological interventions may be more appropriate for nondisabled patients.

In conclusion, we carried out a structural analysis of the GDS-15 in older adult inpatients and extracted 4 factors related with functional disabilities. Factor 3, "loss of morale and hope," and factor 4, "memory loss and reduction of social activity," were highly related with ADL, social variables, and cognitive impairment. In addition, the results suggest that factor analysis will allow improved assessment and medical support of older adult inpatients. Thus, we believe that the results have indicated an extended utility of the GDS-15 not only as a simple screening method for depressive mood but also as a tool for better therapeutic approaches.

References

1. Koenig HG, Blazer DG. Epidemiology of geriatric affective disorders. *Clin Geriatr Med* 1992; 8:235-251.
2. Luber MP, Hollenberg JP, Williams-Russo P, et al. Diagnosis, treatment, comorbidity, and resource utilization of depressed patients in a general medical practice. *Int J Psychiatry Med* 2000; 30:1-13.
3. Jackson R, Baldwin B. Detecting depression in elderly medically ill patients: the use of the Geriatric Depression Scale compared with medical and nursing observations. *Age Ageing* 1993; 22:349-353.
4. Hirschfeld RM, Keller MB, Panico S, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. *JAMA* 1997; 277:333-340.
5. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982-1983; 17:37-49.
6. Royal College of Physicians of London and British Geriatrics Society. *Standardised assessment scales for elderly people.* Report of Joint Workshops of the Research Unit of the Royal College of Physicians and the British Geriatrics Society, 1992.
7. Yesavage JA. The use of self-rating depression scales in the elderly. In Ponn LW (Ed.). *Clinical memory assessment of older*

adults. Washington, DC: American Psychological Association, 1986:213-217.

8. Niino N, Imaizumi T, Kawakami N. A Japanese translation of Geriatric Depression Scale. *Clin Gerontol* 1991; 10:85-87.
9. Cwikel J, Ritchie K. Screening for depression among the elderly in Israel: an assessment of the Short Geriatric Depression Scale (S-GDS). *Isr J Med Sci* 1989; 25:131-137.
10. Herrmann N, Mittmann N, Silver I, et al. A validation study of the geriatric depression scale short form. *Int J Geriatr Psychiatry* 1996; 11:457-460.
11. Almeida OP, Almeida SA. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry* 1999; 14:858-865.
12. Schreiner AS, Morimoto T, Asano H. Depressive symptoms among poststroke patients in Japan: frequency distribution and factor structure of the GDS. *Int J Geriatr Psychiatry* 2001; 16:941-949.
13. AGS Public Policy Committee. Comprehensive geriatric assessment. *J Am Geriatr Soc* 1989; 37:473-474.
14. Health and Public Policy Committee, American College of Physicians. Comprehensive functional assessment for elderly patients. *Ann Intern Med* 1988; 109:70-72.
15. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J* 1965; 14:61-65.
16. Lawton MP, Moss M, Fulcomer M, et al. A research and service oriented multilevel assessment instrument. *J Gerontol* 1982; 37:91-99.
17. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12:189-198.
18. Schreiner AS, Hayakawa H, Morimoto T, et al. Screening for late life depression: cut-off scores for the Geriatric Depression Scale and the Cornell Scale for Depression in Dementia among Japanese subjects. *Int J Geriatr Psychiatry* 2003; 18:498-505.
19. Ozawa T. Comprehensive geriatric assessment. *Jpn J Geriatr* 1996; 35:1-9.
20. Muraoka Y, Oiji A, Ihara K. The physical and psychological and social background factor of elderly depression in the community. *Nippon Ronen Seishin Igakkai Zasshi* 1996; 7:397-407.
21. Patrick L, Knoefel F, Gaskowski P, et al. Medical comorbidity and rehabilitation efficiency in geriatric inpatients. *J Am Geriatr Soc* 2001; 49:1471-1477.
22. Kotila M, Numminen H, Waltimo O, et al. Depression after stroke: results of the FINNSTROKE Study. *Stroke* 1998; 29:368-372.
23. Carson AJ, MacHale S, Allen K, et al. Depression after stroke and lesion location: a systematic review. *Lancet* 2000; 8:122-126.
24. Aneshensel CS, Frerichs RR, Huba GJ. Depression and physical illness: a multiwave, nonrecursive causal model. *J Health Soc Behav* 1984; 5:350-371.
25. Berkman LF, Berkman CS, Kasl S, et al. Depressive symptoms in relation to physical health and functioning in the elderly. *Am J Epidemiol* 1986; 124:372-388.
26. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA* 1989; 262(7): 914-919.
27. Ormel J, Rijsdijk FV, Sullivan M, et al. Temporal and reciprocal relationship between IADL/ADL disability and depressive symptoms in late life. *J Gerontol B Psychol Sci Soc Sci* 2002; 57:338-347.
28. Katz IR. Diagnosis and treatment of depression in patients with Alzheimer's disease and other dementias. *J Clin Psychiatry* 1998; 59(Suppl 9):38-44.
29. Fischer P, Simanyi M, Danielczyk W. Depression in dementia of the Alzheimer type and in multi-infarct dementia. *Am J Psychiatry* 1990; 47:1484-1487.
30. Rovner BW, Broadhead J, Spencer M, et al. Depression and Alzheimer's disease. *Am J Psychiatry* 1989; 146:350-353.
31. Liu CY, Wang SJ, Teng EL, et al. Depressive disorders among older residents in a Chinese rural community. *Psychol Med* 1997; 27:943-949.
32. No authors listed. Grief versus depression in elderly patients. *JAMA* 1979; 241:1558.
33. Phifer JF, Murrell SA. Etiologic factors in the onset of depressive symptoms in older adults. *J Abnorm Psychol* 1986; 95: 282-291.
34. Stuck AE, Siu AL, Wieland GD, et al. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993; 23:1032-1036.

INTERNATIONAL JOURNAL OF GERIATRIC PSYCHIATRY
Int J Geriatr Psychiatry 2004; **19**: 727–733.
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/gps.1161

Smell identification test as an indicator for cognitive impairment in Alzheimer's disease

Y. Suzuki*, S. Yamamoto, H. Umegaki, J. Onishi, N. Mogi, H. Fujishiro and A. Iguchi

Department of Geriatrics, Medicine in Growth and Aging, Program in Health and Community Medicine, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan

SUMMARY

Objectives The aim of the present study was to assess olfactory dysfunction in patients with Alzheimer's disease (AD) and to compare utility of the olfactory tests as possible clinical markers.
Methods Two olfactory identification tests (The Cross-Cultural Smell Identification Test [CC-SIT] and the Picture-based Smell Identification Test [P-SIT]) and the Mini Mental State Examination (MMSE) were administered to patients with AD and age-matched controls. Apolipoprotein E (Apo E) genotypes of patients with AD were identified.
Results Patients with AD had significantly lower olfactory identification scores than age-matched non-demented elderly subjects in both olfactory assessments. In the AD group, the coefficient of correlation between the MMSE scores and the P-SIT scores was higher than that between the MMSE scores and the CC-SIT scores. Receiver operating curve (ROC) analyses for both tests indicated that the P-SIT discriminated AD patients from controls more reliably than did the CC-SIT. Within AD patients, those who were carrying one or two ApoE $\varepsilon 4$ alleles had a higher coefficient of correlation between the MMSE scores and the P-SIT scores than patients without the ApoE $\varepsilon 4$ allele.
Conclusions The results suggest that a short and simple non-lexical olfactory identification test can be useful as a clinical marker of AD appropriate for Japanese elderly population. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS — Alzheimer's disease; impaired olfactory identification; smell identification test; Apolipoprotein E genotypes

INTRODUCTION

Although elderly patients with Alzheimer's disease (AD) are steadily increasing and strategies to cope with various symptoms accompanied by the disease are drawing major concern from clinicians, the biological basis underlying the disease is not yet sufficiently understood. In terms of understanding the neurophysiological basis of the disease and efficacy of therapeutic intervention, it is imperative to establish biological markers with higher sensitivity and specificity by which patients would benefit from early therapeutic intervention (Cummings *et al.*, 2000). Whether preclinical AD can be screened assuredly by existing examinations remains controversial even though various possible biological markers or psychometric instruments have been proposed in recent years (Corder *et al.*, 1993; Minoshima *et al.*, 1997; Bondi *et al.*, 1999; Bookheimer *et al.*, 2000; Goldman *et al.*, 2001; Itoh *et al.*, 2001). A wealth of studies suggest olfactory dysfunction, in particular impairment of olfactory identification, exists in patients with AD (Morgan *et al.*, 1995; Mesholam *et al.*, 1998; Murphy, 1999). Furthermore, the dysfunction may occur at very early stage of the disease, prior to the advent of typical cognitive and behavioral disturbances, hence, assessment of olfactory function can be used as a predictor of the onset of the disease (Nordin and Murphy, 1996; Bacon *et al.*, 1998;

*Correspondence to: Dr Y. Suzuki, Department of Geriatrics, Medicine in Growth and Aging, Program in Health and Community Medicine, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya-shi, 466-8550, Japan. Tel: +81-52-744-2364. Fax: +81-52-744-2371.
E-mail: yus@med.nagoya-u.ac.jp

Contact/grant sponsor: Mitsui-Sumitomo Marine Welfare Foundation.

Received 5 May 2003
Accepted 19 April 2004

Graves *et al.*, 1999; Devanand *et al.*, 2000), as well as being a useful adjunctive screening measure in discriminating AD from other neuropsychiatric conditions of elderly patients (Solomon *et al.*, 1998).

In terms of onset of the disease, apolipoprotein E genotype has been widely acknowledged as a factor that can affect the onset of clinical manifestation (Noguchi *et al.*, 1993). The presence of one or more apolipoprotein E ε4 allele has been reported to increase the risk of cognitive decline in older adults with unexplained olfactory dysfunction (Murphy *et al.*, 1998; Borenstein Graves *et al.*, 1999). Volumetric MRI measurements of early AD patients have shown that apolipoprotein E ε4 allele is associated with the degree of atrophy in the entorhinal cortex, which receives direct sensory projection from the olfactory bulb (Juottonen *et al.*, 1998; Insausti *et al.*, 2002), and also robust relationships between mesial temporal lobe volumes and olfactory functional measures have been confirmed (Murphy *et al.*, 2003). These findings suggest a possible interaction of apolipoprotein E genotype with olfactory performance in AD patients.

In assessing olfaction in humans, various clinical tests of smell have been proposed in previous studies, most of which were aimed to assess ability to either detect smells or identify them (Cain and Gent, 1991; Doty *et al.*, 1996). Olfactory identification is believed to involve higher processing of olfactory stimuli than just detecting smells (Tanabe *et al.*, 1975). Previous studies on olfaction in patients with AD have suggested a possibility that a decline in olfactory identification may precede that of olfactory detection threshold (Serby *et al.*, 1991). Considering the influence of cultural background and personal history in identifying odors, clinical tests of smell applied in previous studies may not have been adequate in assessing olfaction of Japanese elderly population. Also, olfactory tests that rely on lexical function may not directly reflect olfactory dysfunction in patients with dementia (Morgan *et al.*, 1995). Thus, a necessity to generate a battery of standardized olfactory assessment, capable of bypassing cultural bias, emerges (Parola and Liberini, 1999).

In this study, we investigated olfactory identification in patients with AD and age-matched non-demented elderly subjects using two different olfactory identification tests (lexical and non-lexical), and examined the correlation between the degree of cognitive decline and the scores of the two tests with an aim to assess relevance of the tests for clinical application. We also examined a possible interaction of Apo E phenotype with patients' performance in olfactory identification.

SUBJECTS AND METHODS

We administered two olfactory tests to 85 subjects (mean age ± SD: 76.3 ± 7.2) who met Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) and National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria (MacKhann *et al.*, 1984) for AD, and 30 age-matched (mean age ± SD: 74.8 ± 8.5) non-demented elderly subjects. All subjects, who were screened to exclude conditions affecting olfactory function, (e.g. smoking, history of head trauma or nasal disease, presence of respiratory infection, metabolic disorders, medication or cognitive decline), were recruited from the Memory Impairment Clinic at Nagoya University Hospital. Written informed consent was obtained after complete description of the study.

All subjects were subjected to cognitive assessment using the MMSE (Folstein *et al.*, 1983). In order to examine the interaction of olfactory ability with Apo E status, blood samples of patients with Alzheimer's disease were taken for Apo E genotyping. The Apo E genotypes were analyzed and identified using the polymerase chain reaction (PCR) and Hha I digestion as described elsewhere (Emi *et al.*, 1988; Hixon and Vernier, 1990). Digested DNA fragments were separated with polyacrylamide gel electrophoresis and separated fragments of DNA were visualized using ethidium bromide staining.

Two tests for assessing olfactory identification were administered to all subjects. One was the Cross-Cultural Smell Identification test (CC-SIT), a widely used test of odor identification involving a scratch-and-sniff test of 12 microencapsulated odorants with a forced choice of four alternatives per item (Doty *et al.*, 1984a). The other test was picture-based smell identification test (P-SIT), in which subjects were asked to smell six distinctive odorants (ground coffee, incense, ground sesame, green tea, fermented soy-bean paste and soap) that were confirmed to be intense and familiar to cognitively intact Japanese elderly population by preliminary trials. The subjects were instructed to inhale the odorants with their eyes closed through both nostrils without sniffing. Twenty pictures of materials including six pictures corresponding to the smells and fourteen pictures of the materials unrelated to the odorants were presented, and the subjects were asked to identify the odorants by choosing a corresponding picture after smelling each odorant. Before both trials, subjects were confirmed to be able to either understand words presented

for choice or identify all the materials of the pictures, and those who could not answer correctly were eliminated from the study. Each trial was carried out with an interval of 30 seconds in order to avoid the interaction with odors previously presented. The number of correct choices was recorded as a score of the test.

Means age over the two study groups (elderly controls and AD group) were compared by Student's *t*-test. Gender difference of MMSE scores within each group, mean MMSE scores of patients with the $\varepsilon 4$ ($\varepsilon 4+$) allele and without the $\varepsilon 4$ ($\varepsilon 4-$) allele, mean P-SIT and CC-SIT scores of the two groups were compared by Mann–Whitney's U-test. Differences between means were considered statistically significant if $p < 0.05$. The relationship between the age of the subjects and the scores of the P-SIT and the CC-SIT in both groups, and the relationship between the MMSE scores and the scores of the two olfactory tests within the AD group and when divided by two ApoE phenotypes ($\varepsilon 4+$ and $\varepsilon 4-$) were examined using Spearman rank of order correlation coefficients. We also conducted simple regression analyses for both olfactory tests within the AD group and for the P-SIT when AD patients were divided by two ApoE phenotypes to calculate the predicted MMSE scores from the SIT scores as independent variables and compared coefficients of determination. In order to compare the validity of the smell identification tests in discriminating AD group from elderly controls, we carried out receiver operating curve (ROC) analysis at different cut-off scores in each test.

All statistical analyses were performed on a personal computer with the statistical package SPSS for Windows (Version 11.0; SPSS Inc., Chicago, IL, USA).

RESULTS

The two groups did not differ in terms of age ($p = 0.38$). Mean MMSE score of the AD group (mean ± SD: 19.6 ± 4.6) was significantly lower than that of the elderly controls (mean ± SD: 28.5 ± 2.2) ($p < 0.0001$) (Table 1). Within the AD group, there were no gender differences in age ($p = 0.18$), MMSE scores ($p = 0.93$) and in the scores of the two olfactory identification tests (P-SIT; $p = 0.68$, CC-SIT; $p = 0.23$). The ApoE genotypes for patients with AD were: $\varepsilon 23$, $n = 4$; $\varepsilon 33$, $n = 27$; $\varepsilon 24$, $n = 3$; $\varepsilon 34$, $n = 38$; $\varepsilon 44$, $n = 4$. Patients who were carrying one or two $\varepsilon 4$ alleles did differ in age ($p < 0.001$) but did not differ in the MMSE score ($p = 0.30$) and in the scores of the two olfactory identification tests (P-SIT: $p = 0.84$, CC-SIT: $p = 0.93$) from those without the $\varepsilon 4$ allele. All subjects (AD patients and elderly controls) were confirmed to correctly identify all the pictures of the materials in the P-SIT. However, many participants responded while administering the CC-SIT that it was difficult to imagine the corresponding odors of some of the alternatives such as turpentine or paint thinner.

Table 1. Demographic variables of participants

	Alzheimer's ($n = 85$)	Non-demented ($n = 30$)
Mean age ± SD	76.3 ± 7.2	74.8 ± 8.5
Age range (M/F)	58–89 (17/68)	60–94 (12/18)
Student's *t*-test	$p = 0.38$, $t = 0.89$, effect size = 0.20	
Mean MMSE score ± SD	19.6 ± 4.6	28.5 ± 2.2
Score range	12–28	24–30
Mann–Whitney's U Test	$p < 0.0001$, $z = 4.76$, effect size = 1.71	

A significant difference of the P-SIT scores between the AD group (mean ± SD: 1.5 ± 1.3) and the elderly controls (mean ± SD: 4.4 ± 1.2) was observed ($p < 0.0001$, $z = 6.848$). Difference of the CC-SIT scores between the two groups also reached a statistical significance (AD group; mean ± SD: 4.1 ± 2.5, elderly controls; mean ± SD: 6.8 ± 1.7), ($p < 0.0001$, $z = 4.339$) (Figure 1). Effect size of difference between the two groups for the P-SIT (1.59) was greater than that of the CC-SIT (1.04).

Regarding the effect of age on the performance of olfactory tests, negative correlation was observed between the age and the score of the P-SIT both in the elderly controls ($\rho = -0.77$, $p < 0.001$) and in the AD group ($\rho = -0.23$, $p = 0.045$). Meanwhile, correlations between the age of the subjects and the CC-SIT scores in both groups did not reach a level of statistical significance (elderly controls:

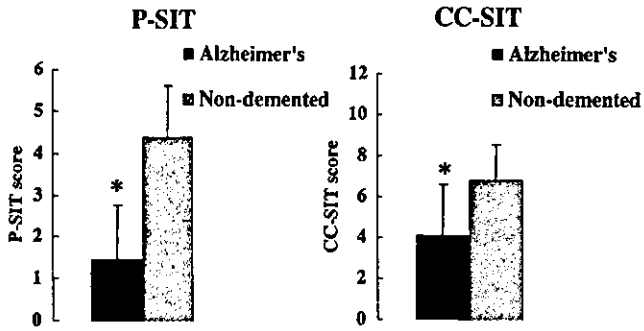


Figure 1. Comparison of the Smell Identification Test (P-SIT and CC-SIT) scores of patients with Alzheimer's disease and non-demented elderly subjects. *$p < 0.0001$