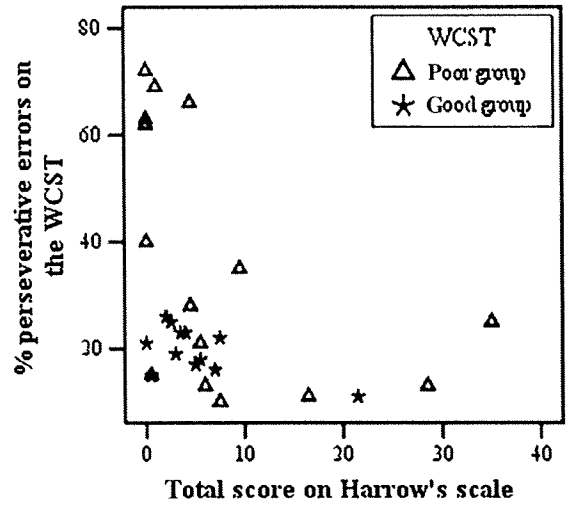

Figure 1. Correlation between total score on Harrow's scale and % perseverative errors on the WCST. The good WCST group exhibited a significant negative partial correlation, while controlling for age, between total score on Harrow's scale and % perseverative errors on the WCST ($r = -0.87$, $p < .001$; raw data are shown in the Figure). WCST: Wisconsin Card Sorting Test; poor WCST group: no category completed on WCST; good WCST group: number of categories completed on WCST ≥ 1.

Relationships Between Thought Disorder and Demographic and Clinical Characteristics

Although, some studies reported that thought disorder did not exhibit significant correlations with clinical variables (Mazumdar, Chaturvedi, & Gopinath, 1994) or demographic factors (Harrow & Marengo, 1997), the present study found significant correlations between WAIS comprehension test score on Harrow's scale and both age at onset and duration of illness after controlling for age. Correlations were assessed in this study while controlling for age since clinical variables such as duration of illness are affected by current age, andbecause significant positive correlations were found in this study between age and estimated IQ and BPRS total score, and age-related differences in thought disorder were reported by a previous study (Harvey et al., 1997). Parameters of thought disorder exhibited no significant correlations with the duration of education or SES, unlike the results of the WCST. While Saito et al. (1997) reported a significant negative correlation between severity of thought disorder and duration of education, the results of the present study suggests that no direct relationship exists between positive thought disorder and IQ in schizophrenia.

Previous studies reported that negative and positive thought disorders and negative symptoms were already present in mid-childhood in subjects with adult schizophrenia years before the onset of psychosis (Ott, Roberts, Rock,

Allen, & Erlenmeyer-Kimling, 2002). The findings of the present study showing a lack of significant correlation between severity of positive thought disorder and daily dose of antipsychotics and that some inpatients with schizophrenia had no positive thought disorder suggest that there may be subtypes with and without positive thought disorder in schizophrenia, and that thought disorder, if it exists, appears in the childhood and persists for a long time, independent of pharmacological treatment.

Relationships Between Executive Function and Demographic and Clinical Characteristics

Number of categories completed on the WCST exhibited significant positive correlations with both duration of education and estimated IQ after controlling for age, consistent with previous studies (Laws, 1999). No significant associations were found between WCST performance and daily doses of medications.

Among the variables measured on the WCST, perseveration appears to be the most diagnostically useful and characteristic feature of schizophrenia (Koren et al., 1998). Previous studies reported that earlier age at onset was associated with poor performance on the WCST in schizophrenia (Bellino et al., 2004; Stratta et al., 1997), which is apparently contrary to our results; however, their previous findings were based on number of categories completed (Stratta et al., 1997) and number of perseverative errors (Bellino et al., 2004), and not the perseverative error rates as in our study. The results of higher perseverative error rates on the WCST with significant associations with later age at onset and shorter duration of illness were beyond our expectation. These may be caused by the difference in the subjects' population, as the subjects in the present study showed apparently poorer performance level in the WCST than those in the previous studies. This should be reconfirmed using a larger and wider-ranging sample in the future.

Association Between Thought Disorder and Executive Function

In the present study, no significant correlations were found between Harrow's scale scores and performance on the WCST in the group of all subjects, whereas the TD group had significantly lower % perseverative errors than that in the NTD group. Moreover, as the mean number of categories completed in this study, 1.2, was much lower than those reported in previous studies (2.3 in Nestor et al., 1998) (3.4 in Stratta et al., 2000), and more than half of the subjects completed no category in this study; we divided the subjects according to the number of categories completed in the WCST. The results of analyses using the subjects who completed more than one category showed that more

severe thought disorder was associated with lower perseverative dysfunction in schizophrenia. These findings clearly contradict our expectation and suggest that thought disorder and executive function play independent or relatively independent roles in schizophrenia.

Despite the performance of WCST, including perseverative error has been considered to reflect the flexibility of thought (Heaton, Chelune, Talley, Kay, & Curtis, 1993a). Why did the patients with higher perseverative error rates on the WCST show less thought disorder evaluated from the content of speech in the present study? One of the possible reasons for our findings may be due to the methodological design of batteries; that is, Harrow's scale is essentially designed for the purpose of assessing positive thought disorder, but not negative thought disorder, such as poverty of speech or alogia. The patients with negative thought disorder might even have scored less than expected as no response (I don't know) is estimated as 0 point. It can be speculated that the patients with stronger negative thought disorder, scoring lower in the Harrow's scale, showed higher perseverative error rates in the WCST because they lacked flexibility of thought as compared with the patients scoring higher in the Harrow's scale.

These findings suggest the possibility that (1) positive thought disorder and perseverative function play different roles in schizophrenia, and (2) there are subtypes of schizophrenia with severe perseverative dysfunction but less positive thought disorder.

Limitations

Some limitations in this study must be mentioned. First, this study was limited in particular by its small sample size, which may have given rise to type II errors. Moreover, the extremely poor WCST performance may have affected the results in this study. One of the reasons for poor performance on the WCST may be that all subjects were recruited from hospitalized patients. Finally, the multiplicity of statistical analyses was not corrected, implying that our findings should be taken as exploratory.

In conclusion, our findings suggest that positive thought disorder and executive function may play different roles in schizophrenia. The finding that lower perseverative error was associated with more severe thought disorder in the good WCST group suggests the possibility of existence of a subgroup of patients with these features among those with schizophrenia. Further investigations with large sample sizes and longitudinal studies are needed to elucidate the associations among outcome, thought disorder, and executive function.

REFERENCES

American Psychiatric Association (APA) (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th ed.* Washington, DC: American Psychiatric Association.

Bellino, S., Rocca, P., Patria, L., Marchiaro, L., Rasetti, R., Di Lorenzo, R., et al. (2004). Relationships of age at onset with clinical features and cognitive functions in a sample of schizophrenia patients. *Journal of Clinical Psychiatry, 65*(7), 908–914.

Breier, A., Schreiber, J. L., Dyer, J., & Pickar, D. (1991). National Institute of Mental Health longitudinal study of chronic schizophrenia: Prognosis and predictors of outcome. *Archives of General Psychiatry, 48*(3), 239–246.

Curtiss, G., & Tuttle, K. (1993). *Wisconsin card sorting test: Computer version 2 [Program Manual]*. Odessa, FL: Psychological Assessment Resources, Inc.

Dixon, T., Kravariti, E., Frith, C., Murray, R. M., & McGuire, P. K. (2004). Effect of symptoms on executive function in bipolar illness. *Psychological Medicine, 34*(5), 811–821.

Eckblad, M., & Chapman, L. (1983). Magical Ideation as an Indicator of schizotypy. *Journal of Consulting and Clinical Psychology, 51*(2), 215–225.

Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry, 54*(2), 159–165.

Gooding, D. C., Kwapil, T. R., & Tallent, K. A. (1999). Wisconsin Card Sorting Test deficits in schizotypic individuals. *Schizophrenia Research, 40*(3), 201–209.

Gorham, D. R. (1956). A proverbs test for clinical and experimental use. *Psychological Reports Monograph, Supplement, 1*, 1–12.

Guelfi, G., Faustman, W., & Csernansky, J-Y. (1989). Independence of positive and negative symptoms in a population of schizophrenic patients. *Journal of Nervous and Mental Disease, 177*(5), 285–290.

Harrow, M., & Marengo, J. T. (1986). Schizophrenic thought disorder at follow up: Its persistence and prognostic significance. *Schizophrenia Bulletin, 12*(3), 373–393.

Harrow, M., & Marengo, J. T. (1997). Longitudinal courses of thought disorder in schizophrenia and schizoaffective disorder. *Schizophrenia Bulletin, 23*(2), 273–285.

Harrow, M., Marengo, J. T., & McDonald, C. (1986). The early course of schizophrenic thought disorder. *Schizophrenia Bulletin, 12*(2), 208–224.

Harrow, M., & Quinlan, D. M. (1985). *Disordered Thinking and Schizophrenic Psychopathology*. New York: Gardner Press.

Harvey, P. D., Lombardi, J., Leibman, M., Parrella, M., White, L., Powchick, P., et al. (1997). Age-related differences in formal thought disorder in chronically hospitalized schizophrenic patients: a cross-sectional study across nine decades. *American Journal of Psychiatry, 154*(2), 205–210.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993a). *Wisconsin Card Sorting Test (WCST) manual revised and expanded.* Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Curtiss, G., & Tuttle, K. (1993b). *Wisconsin Card Sorting Test: Computer (Version 2) [Computer Software]*. Odessa, FL: Psychological Assessment Resources.

Higashima, M., Urata, K., Kawasaki, Y., Maeda, Y., Sakai, N., Mizukoshi, C., et al. (1998). P300 and the thought disorder factor extracted by factor-analytic procedures in schizophrenia. *Biological Psychiatry, 44*(2), 115–120.

Hollingshead, A. B., & Redlich (1958). *Redlich FC: Social Class and Mental Illness*. New York: John Wiley & Sons.

Iwanami, A., Okajima, Y., Kuwakado, D., Isono, H., Kasai, K., Hata, A., et al. (2000). Event-related potentials and thought disorder in schizophrenia. *Schizophrenia Research, 42*(3), 187–191.

Kazumdar, P. K., Chaturvedi, S. K., & Gopinath, P. S. (1994). Thought disorder and its correlation with clinicodemographic variables in schizophrenia. *Psychopathology, 27*(1–2), 37–42.

Koren, D., Seidman, L. J., Harrison, R. H., Lyons, M. J., Kremen, W. S., Caplan, B., et al. (1998). Factor structure of the Wisconsin Card Sorting Test: Dimensions of deficit in schizophrenia. *Neuropsychology, 12*(2), 289–302.

Laurent, J. P., & Baribeau, J. (1992). AERPs and clinical evolution of thought disorders over five years. *International Journal of Psychophysiology, 13*(3), 271–282.

Laws, K. R. (1999). A meta-analytic review of Wisconsin Card Sort studies in schizophrenia: General intellectual deficit in disguise? *Cognitive Neuropsychiatry, 4*(1), 1–30.

Marengo, J. T., & Harrow, M. (1987). Schizophrenic thought disorder at follow-up. A persistent or episodic course? *Archives of General Psychiatry, 44*(7), 651–659.

Marengo, J. T., Harrow, M., Lanin-Kettering, I., & Wilson, A. (1986). Evaluating bizarre-idiosyncratic thinking: a comprehensive index of positive thought disorder. *Schizophrenia Bulletin, 12*(3), 497–511.

Mazumdar, P. K., Chaturvedi, S. K., & Gopinath, P. S. (1994). Thought disorder and its correlation with clinicodemographic variables in schizophrenia. *Psychopathology, 27*(1–2), 37–42.

McConaghy, N., Catts, S. V., Michie, P. T., Fox, A., Ward, P. B., & Shelley, A. M. (1993). P300 indexes thought disorder in schizophrenics, but allusive thinking in normal subjects. *Journal of Nervous and Mental Disease, 181*(3), 176–182.

Miyata, R., Fujii, Y., Inagaki, A., Inada, T., & Yagi, G. (1995). Reliability of the Japanese version of Brief Psychiatric Scale (BPRS) (in Japanese). *Clinical Evaluation, 23*(2), 357–367.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12*(4), 313–324.

Nestor, P. G., Shenton, M. E., Wible, C., Hokama, H., O'Donnell, B. F., Law, S., & McCarley, R. W. (1998). A neuropsychological analysis of schizophrenic thought disorder. *Schizophrenia Research, 29*(3), 217–225.

Ott, S. L., Roberts, S., Rock, D., Allen, J., & Erlenmeyer-Kimling, L. (2002). Positive and negative thought disorder and psychopathology in childhood among subjects with adulthood schizophrenia. *Schizophrenia Research, 58*(2–3), 231–239.

Overall, J. E., & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports, 10*, 799–812.

Parellada, E., Catarineu, S., Catafau, A., Bernardo, M., & Lomena, F. (2000). Psychopathology and Wisconsin Card Sorting Test performance in young unmedicated schizophrenic patients. *Psychopathology, 33*(1), 14–18.

Racenstein, J. M., Penn, D., Harrow, M., & Schleser, R. (1999). Thought disorder and psychosocial functioning in schizophrenia: The concurrent and predictive relationships. *Journal of Nervous and Mental Disease, 187*(5), 281–289.

Reed, R. A., Harrow, M., Herbener, E. S., & Martin, E. M. (2002). Executive function in schizophrenia: Is it linked to psychosis and poor life functioning? *Journal of Nervous and Mental Disease, 190*(11), 725–732.

Saito, K., Sugiura, M., Hata, T., Nakagome, K., Iwanami, A., & Niwa, S. (1997). The usefulness of Harrow Thought Disorder Scale (Japan Edition) for thought disorder in schizophrenic patients (in Japanese). *Japanese Journal of Clinical Psychiatry, 26*(11), 1443–1451.

Sponheim, S. R., Surerus-Johnson, C., Leskela, J., & Dieperink, M. E. (2003). Proverb interpretation in schizophrenia: The significance of symptomatology and cognitive processes. *Schizophrenia Research, 65*(2–3), 117–123.

Stratta, P., Daneluzzo, E., Bustini, M., Prosperini, P., & Rossi, A. (2000). Processing of context information in schizophrenia: Relation to clinical symptoms and WCST performance. *Schizophrenia Research, 44*(1), 57–67.

Stratta, P., Mancini, F., Mattei, P., Daneluzzo, E., Bustini, M., Casacchia, M., et al. (1997). Remediation of Wisconsin Card Sorting Test performance in schizophrenia. *Psychopathology, 30*(2), 59–66.

Sugiura, M., Saito, K., Hata, A., Nakagome, K., Iwanami, A., Ichikawa, I., et al. (1995). Reliability of the Japanese version of Harrow's Thought Disorder Scale (in Japanese). *Archives of Psychiatric diagnostics and Clinical evaluation, 6*(3) 333–344.

Uetsuki, M., Kasai, K., Araki, T., Yamasue, H., Maeda, K., Seino, K., et al. (2004). Varlidity of the short forms of the WAIS-R in schizophrenia (in Japanese). *Clinical Psychiatry, 46*(8), 845–848.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale.* New York: The Psychological Corporation.

Wechsler, D. (1981). *WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised.* New York: The Psychological Corporation.

doi:10.1111/j.1479-8301.2008.00256.x PSYCHOGERIATRICS 2009; 9: 4–10

ORIGINAL ARTICLE

Effectiveness of the Takeda Three Colors Combination Test as a screening test for dementia

Shinya TAKEDA,[1] Kazuyuki NAKAGOME,[2] Kayo TAJIME[3] and Motoi SAITO[3]

Departments of [1]Clinical Psychology and [2]Psychiatry, Tottori University Faculty of Medicine and [3]Tottori Seikyo Hospital, Tottori, Japan

Correspondence: Lecturer Shinya Takeda MA, Department of Clinical Psychology, Tottori University Faculty of Medicine, 86 Nishicho, Yonago, Tottori 683-8503, Japan. Email: takedas@med.tottori-u.ac.jp

Received 21 January 2008; accepted 11 August 2008.

Abstract

Background: The aged population is increasing worldwide and it is expected that dementia, for which aging is a risk factor, will increase as well. A critical issue then becomes a community's capacity for the early detection of dementia among its senior citizens. In the present paper, we report on the development and potential use of a screening test for dementia that can be administered and assessed easily in a short period of time by non-specialist clinicians and represents no burden for those undergoing the screening.

Methods: Three hundred and sixty senior citizens, over 60 years of age, participated in the study. Of these, 126 had Alzheimer's disease (AD), 60 had vascular dementia (VaD), 41 had mild cognitive impairment (MCI), and 133 were healthy volunteers (control group). A screening test for dementia, which consisted of a colored cards configuration memory task (the Takeda Three Colors Combination Test; TTCC) was examined for sensitivity, specificity, reliability and criterion-related validity.

Results: Sufficient sensitivity and specificity were demonstrated for each clinical group (AD, VaD) and the control group. The sensitivity of the TTCC was 0.94, 0.82, and 0.71 for the AD, VaD, and MCI groups, respectively; specificity was 0.83. In addition, sufficient reliability and validity were established. Administration of the TTCCand assessment procedures required only 1 or 2 min.

Conclusion: Satisfactory sensitivity and specificity were indicated for both the AD and VaD groups, with sufficient reliability and validity also indicated. Thus, the TTCC is an effective dementia screening test.

Key words: *Alzheimer's disease, mild cognitive impairment, screening test for dementia, vascular dementia.*

INTRODUCTION

An increase in the elderly population has been seen not only in Japan, but also worldwide. According to a United Nations report,[1] by 2050 senior citizens will account for one-fifth of the world's population. As the elderly population increases, the need to manage dementia will also increase; this has become an international issue.

With an increase in the number of dementia cases, research efforts aimed at developing treatments for dementia have also increase, which has resulted in the development of acetylcholinesterase inhibitors, which slow the progress of Alzheimer's disease (AD) in its early stage.[2] Studies of an AD vaccine to decompose β-amyloid protien are also under way.[3] Some studies have reported on the efficacy of non-medical interventions, such as prevention programs for dementia, and cognitive rehabilitation, which is effective for mild cognitive impairment (MCI) and people with dementia in its early stages.[4,5] Delayed morbidity and the progression of dementia leads to a reduced need for care and a reduction in medical expenses. Thus, the early detection of dementia is becoming an increasingly important issue in mental gerontology.

However, because diagnosis of dementia in its early stages generally requires time, human resources, and

diagnostic imaging equipment, and is conducted within a special clinical setting, it is not unusual to find that the patient's condition has already progressed considerably by the time a diagnosis is made. It is estimated that 20–30% of elderly people in any one community experience dementia.[6,7] Aging is a major risk factor for dementia[8] and so early detection is important. In order to facilitate early detection, it is suggested that the screening test: (i) be easy to administer for non-specialists; (ii) imposes no strain on those undergoing the screening; (iii) be accomplished in a short period of time; and (iv) be easy to evaluate.

With these conditions in mind, Takeda *et al.* developed the Takeda Three Colors Combination Test (TTCC), an easy and time-saving screening test.[9] However, in their previous study the following problems remained: (i) vascular dementia (VaD) was not considered in their investigation, which is as typical a dementia as AD; (ii) the validity of the TTCC was not examined; and (iii) MCI was not investigated.

Thus, the aim of the present study was to determine whether the TTCC could be a useful screening test for dementia by comparing the results of the TTCC between the AD and control groups, as well as the VaD and control groups, and to calculate the screening's reliability and validity. In addition, we discuss the plausibility of the TTCC as a screening test for MCI by comparing results obtained using the test between the MCI and control groups.

METHODS

Subjects

There were 360 people included in the present study (both men and women) that were divided into four groups: AD, VaD, MCI and the control group (Table 1). The age of subjects in all groups was restricted over 60 years. The people in the AD and MCI groups were patients receiving treatment at the Department of Psychiatry, Tottori Seikyo Hospital, who were complaining of memory lapses. The people in the VaD group were patients with the same complaint who were either receiving treatment at the Department of Psychiatry or the Cranial Nerve Department in the hospital. The *Diagnostic Statistical Manual of Mental Disorders*, 4th edition (DSM-IV),[10] was used for the diagnosis of AD and VaD. People with mixed dementia were excluded from the present study. The people in the MCI group were patients who did not meet the criteria of any dementia, but who scored 0.5 on the Clinical Dementia Rating (CDR) scale[11] and over 21 on the Revised Hasegawa Dementia Scale (HDS-R),[12] based on the definition by Petersen *et al.*[13]

Table 1 Subject characteristics

	AD group (n = 126)	VaD group (n = 60)	MCI group (n = 41)	Control group (n = 133)
Age (years)	72.6 (8.0)	71.5 (7.3)	70.6 (7.1)	68.1 (6.5)
No. men/women	42/84	35/25	14/27	46/87
HDS-R socre	14.7 (4.7)	15.7 (5.5)	23.5 (1.6)	28.4 (1.5)

Data shos the mean value, with the SD given in parentheses.
AD, Alzheimer's disease; VaD, vascular dementia; MCI, mild cognitive impairment; HDS-R, Revised Hasegawa Dementia Scale.

The control group consisted of patients who had undergone a brain examination at Tottori Seikyo Hospital, had no complaint of memory lapses, were not diagnosed as having dementia according to DSM-IV standards, showed no precipitant conditions of cognitive dysfunction on magnetic resonance imaging scans, and scored 25 or over on the HDS-R. The objectives and methods of the study were explained to the subjects (or their families) and consent was obtained prior to participation in the study.

Procedures

About the test

The dementia screening test used in the present study was the TTCC developed by Takeda *et al.* in 2004.[9] The examiner was not told which of the four groups a patient belonged to. The test is described in more detail below.

Implements The following items were used: three wooden colored cards (red, blue, and yellow), each 5 cm^2 and 5 mm thick, and a card with a diagram of the three colored squares in a certain configuration. Hereafter, the card with the diagram is referred to as the model (Fig. 1).

Outline The model is presented to the subject for 5 s before it is hidden. Following a simple interference task (a mathematical calculation), the subject is asked to arrange the three wooden colored cards to match the configuration shown in the model.

Procedure The test procedure was as follows:

1. Show the subject the three wooden colored cards and confirm that he or she can distinguish the colors.

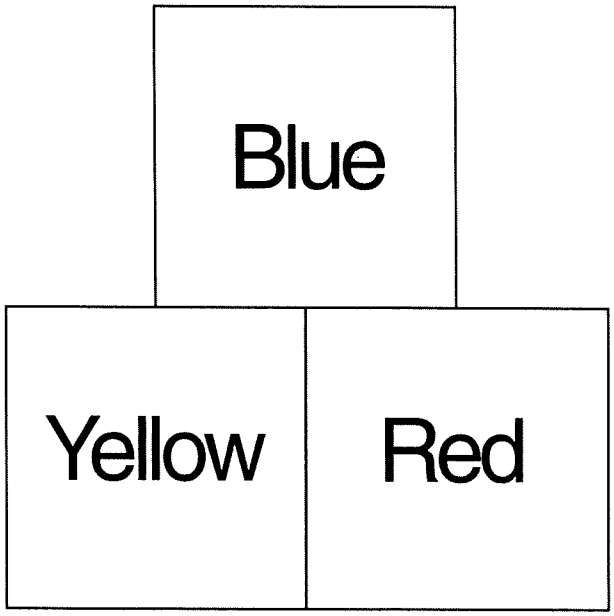

Figure 1 Arrangement of the three colored squares that should be remembered and reproduced later, after an interference task, using three wooden colored cards.

2. Explain what to do; for example, 'I'll show you a figure for 5 s, please remember it. I'll ask you to make the same figure from memory later using these three cards. Do you understand?' If the subject understands the instructions, hide the cards and proceed to the next step. If the subject does not understand, explain again.
3. Present the model for 5 s and then hide it quickly. Ask the subject, 'What is 100 minus 7?'
4. After the subject has answered the question, put the three colored cards in a pile and hand him or her the pile of cards. Ask the subject, 'Please make the same figure I showed you earlier'. Allow the subject 1 min to accomplish the task.

Conditions for stopping the test

The test is stopped if: (i) the subject is not able to complete the arrangement within 1 min; or (ii) the subject completes the arrangement within 1 min.

Assessment

If the subject succeeded in making exactly the same figure as shown in the model, he or she is assessed as being normal. If the arrangement of the colored cards differs from that shown in the model, dementia is suspected. Assessment does not include determing the correctness of the calculation, which is performed as an interference task.

Any of the following is regarded as an incorrect answer: (i) the cards overlap; (ii) the top square is placed off-center, such that the ratio of the lengths of the bottom side contacting the two lower squares exceeds 1.5; or (iii) the cards are arranged more than 5 mm apart.

Statistical analysis

The TTCC was administered to all groups. The difference between the control group and each of the other groups was examined using the Chi-squared test. The sensitivity of the TCC was calculated using data for each disease group who responded incorrectly in the first trial.

We conducted the TTCC again 1 month later for all groups, except the control group, to calculate the reliability of the test. The agreement between the first and second trials was calculated, as was the Phi coefficient.

Criterion-related validity was determined using HDS-R as an external criterion. With the TTCC, we allotted [1] to a correct response and [0] to an incorrect response as dummy variables and calculated Spearman's rank correlation coefficient between TTCC and HDS-R scores.

Concerning the age difference among groups, when a significant difference was detected in one-way analysis of variance, we once again allotted [1] to a correct response and [0] to an incorrect response as dummy variables to calculate Spearman's correlation coefficient between age and TTCC results. Concerning gender differences among groups, when a significant difference in gender ratio was detected among groups, we conducted a Chi-squared test to identify any relationship between gender and TTCC results.

RESULTS

Control group

In the control group, 110 people responded correctly and 23 responded incorrectly. Based on these results, the specificity of the TTCC is 0.83.

Comparison between the AD and control groups

In the AD group, seven patients responded correctly and 119 responded incorrectly (Table 2). Thus, the sensitivity of the TTCC was 0.94 for the AD group. In the Chi-squared test (AD group [2] × Control group [2]),

Table 2 The number of correct and incorrect responses to the Takeda Three Colors Combination test in each group

	Correct response	Incorrect response	Total
Control group (*n*)	110	23	133
AD group* (*n*)	7	119	126
VaD group** (*n*)	11	49	60
MCI group*** (*n*)	12	29	41

*Alzheimer's disease (AD) versus control: $\chi^2 = 155.5$, d.f. = 1, $P < 0.001$; **vascular dementia (VaD) versus control: $\chi^2 = 73.3$, d.f. = 1, $P < 0.001$; ***mild cognitive impairment (MCI) versus control: $\chi^2 = 42.7$, d.f. = 1, $P < 0.001$.

the difference in the performance of the TTCC between groups was significant ($\chi^2 = 155.5$; $P < 0.001$). The analysis revealed a significantly increased number of incorrect responses in the AD group and correct responses in the control group.

Including the results of the second trial in the AD group, five patients responded correctly in both trials, two responded correctly in the first trial but incorrectly in the second trial, 116 responded incorrectly in both trials, and three responded incorrectly in the first trial but correctly in the second trial.

Comparison between the VaD and control groups

In the VaD group, 11 patients responded correctly to the TTCC and 49 responded incorrectly (Table 2). Thus, the sensitivity of the TTCC for the VaD group was 0.82. In the Chi-squared test (VaD group [2] × Control group [2]), the difference between groups was significant ($\chi^2 = 73.3$; $P < 0.001$). The analysis revealed a significantly increased number of incorrect responses in the VaD group and correct responses in the control group.

Including the results of the second trial for the VaD group, seven patients responded correctly in both trials, four responded correctly in the first trial but incorrectly in the second trial, 49 responded incorrectly in both the first and second trails, and none responded incorrectly in the first trial but correctly in the second trial.

Comparison between the MCI and control groups

In the MCI group, 12 patients responded correctly to the TTCC and 29 responded incorrectly (Table 2). Thus, the sensitivity of the TTCC for the MCI group was 0.71. In the Chi-squared test (MCI group [2] × Control group [2]), the difference between groups was significant ($\chi^2 = 42.7$; $P < 0.001$). Secondary analysis showed a significantly increased number of incorrect responses in the MCI group and correct responses in the control group.

Including the results of the second trial for the MCI group, 11 patients responded correctly in both trials, one patient responded correctly in the first trial but incorrectly in the second trial, 24 responded incorrectly in both trials, and five responded incorrectly in the first trial but correctly in the second trial.

Reliability of the TTCC

Including the results of the second trials in the AD, VaD, and MCI groups, a total of 23 patients responded correctly in both trials, seven responded correctly in the first trial but incorrectly in the second trial, 189 responded incorrectly in both trials, and eight responded incorrectly in the first trial but correctly in the second trial. The Phi coefficient was 0.72 and the consistency percentage between the first and second trials was 93.4%.

Validity of the TTCC

Spearman's rank correlation analysis, using HDS-R as an external standard, revealed a significant correlation between the HDS-R score and results of the TTCC (Rho = 0.66; $P < 0.01$).

Age and gender differences between groups

One-way analysis of variance using 'age' as a dependent variable and 'group' as an independent variable yielded a significant effect of 'group' ($F_{(3)} = 8.98$; $P < 0.001$). In Tukey's multiple comparison analyses, a significant difference was obtained between the control and AD groups, as well as between the control and VaD groups. However, Spearman's Rho between age and the TTCC outcomes was not significant (Rho = –0.10).

The difference in gender ratio among groups was significant ($\chi^2 = 12.6$; $P < 0.01$) using the Chi-squared test (group [4] × gender [2]). Secondary analysis revealed that the VaD group was comprised of significantly more men and fewer women compared with the control group. However, the Chi-squared test (gender [2] × the TTCC outcomes [2]) for all subjects revealed no significant relationship between gender and TTCC outcomes ($\chi^2 = 0.15$).

Time requirements for the TTCC

The TTCC (including instruction and assessment) was accomplished within 1 min for 78% of subjects and

within 2 min for 22% of subjects. No refusal or resistance to this test was observed by any of the subjects.

DISCUSSION

Effectiveness of TTCC in screening for dementia

We have suggested the importance of screening for dementia in the community. However, it is not plausible to accomplish dementia screening in the community using only a limited number of specialists. If we are to take effective measures against dementia, naturally we need to develop a screening test that can be administered and assessed simply and quickly by non-specialists.

The leading cause of dementia is AD, followed by VaD. These two types of dementia together account for 70–80% of all dementia cases.[14,15] Thus, any screening test developed should enable detection of both AD and VaD. The TTCC screening test meets these criteria, as demonstrated in the comparisons of the AD, VaD, and control groups.

First, a significant difference was found between the AD and control groups. The sensitivity of the test was 0.94 and specificity was 0.83. Second, were able to differentiate the VaD group, which was investigated for first time in the present study, from the control group with a sensitivity of 0.82.

The relatively high sensitivity of the TTCC test for AD shown in the screening is explained by the fidelity of the TTCC in detecting impairment in recent memory and space perception, both of which are seen in the early stages of AD.[16] Recent memory, among all memory classifications, is the one damaged at the earliest stage of AD.[17] A recent memory task requires subjects to recall items after a short interval, during which an interfering task has been inserted.[18] With the TTCC, although there is no appreciable time lapse following memorization, an interfering task is imposed before the subject is required to retrieve the model, implicating recent memory rather than immediate icon memory.

The TTCC is also relevant to visuospatial cognition, because the subject is required to reconstruct the spatial arrangement of the cards from visuospatial memory of the model figure. Impairment in visuospatial cognition is prominently seen at an early stage of AD.[19]

The sensitivity of the test in diagnosing VaD was somewhat less than that for AD. We consider this is because the main symptoms of VaD are not necessarily impairments in recent memory or visuospatial cognition. Cognitive impairments of VaD are more varied than in AD depending on the area of the brain that is damaged.[20]

Despite the difference in sensitivity of the TTCC for AD and VaD, the detection rate compares favorably with other screening tests for AD and VaD.[21–23] Therefore, it may be assumed that the TTCC is as effective a screening test for both AD and VaD.

As for the time required to administer the TTCC, in all cases the task was completed within 2 min, with most people able to complete the task in less than 1 min. Considering that subjects tend to get tired easily and experience a decline in attention and concentration, screening tests should be conducted as promptly as possible. Thus, the short time required to administer the TTCC makes it an appropriate screening test for seniors. None of the subjects refused to take the TTCC and none showed any resistance during the task. The task of remembering a figure and reproducing it was accepted as enjoyable, motivating subjects to participate in the screening. In fact, with other widely used screening tests, such as the Mini-Mental State Examination, some subjects may get embarrassed or sometimes refrain from taking the task because of the offensive content of some of the test items.[22] In this light, the TTCC is taken not as unpleasant as some other tests.

The effectiveness of the TTCC as a screening test for MCI was also examined in the present study. Comparing MCI and control groups indicated a significant difference in the number of correct responses, but the sensitivity of the TTCC for MCI was 0.71, which is not high. We cannot determine whether the TTCC is an effective screening test for MCI.

Validity and reliability of the TTCC

The TTCC was administered to subjects again after 1 month, except for the control group, to determine its reliability. Considerable correlation in the Phi coefficient was indicated between the first and second trials. The consistency score was also very high. These results are almost equal to those of our previous study.[9] As for inter-rater reliability, taking into account that the TTCC is quite simple to administer and requires no special training, the effect of the examiner on the results was thought to be negligible. Given the results noted above, we may say that the TTCC is a highly reliable screening test for dementia.

Criterion-related validity, measured with the HDS-R as an external standard, indicated a rather high correlation between the TTCC results and HDS-R scores. Because the HDS-R is widely used in Japan, we think that the high correlation between the TTCC results and HDS-R scores suggests the criterion-related validity of the TTCC as a screening test for dementia.

Significant differences were obtained among groups in age and gender ratio; however, no significant relationship was shown between these factors and the TTCC results. Therefore, we did not take into account these factors in statistical tests.

Effectiveness of the TTCC as a medical checkup tool

The importance of detecting dementia in aged individuals is growing as the aged population increases. It is desirable that a screening test that can be administered by non-specialists be developed.

Takayama[23] suggests that if screening is to be conducted quickly in locations other than hospitals by non-specialist clinicians in the field of neuropsychology, then the first structural condition to be met is that the test be easy to take, as well as motivating, for both the aged who are otherwise healthy and those who are cognitively impaired. The equipment should be low cost and the procedure simple.

The results of the present study have shown that the TTCC is satisfactory in terms of sensitivity and specificity and that it is an effective screening test for dementia, with adequate reliability and criterion-related validity. It is administered quickly and was accepted favorably by all subjects. It requires no special training or techniques to administer. Therefore, we can say that the TTCC is an easy-to-use screening test, especially effective for use in the community.

The content of the TTCC test is quite simple and so the user's education or cultural background, or even language, is likely to have little effect on the outcome. Given also that it requires little in terms of equipment expenditure, the TTCC appears to be well suited for worldwide use as a screening test for dementia.

Limitations and problems to be overcome in the present research

The present study examined the usability of a screening test developed for dementia. Because the principal aim of screening tests is early detection, it would have been preferable to include more subjects at early stages with mild dementia. However, most subjects in the present study were patients with relatively severe advanced AD and VaD. An MCI patient group was included in the present study and the sensitivity of the TTCC for MCI was approximately 70%. However, it has been reported that approximately 50% of subjects diagnosed with MCI show improvement in subsequent screening tests.[24,25] Therefore, the validity of the TTCC for detecting those MCI patients who develop dementia in the future is not yet clear. It would be of considerable interest to follow those subjects with MCI who responded incorrectly in the TTCC in the present study to see whether they actually convert to dementia at some time in the future.

The subjects in the present study were all patients of a single medical institute, making it difficult to generalize the results across samples. Further testing with patients from a number of medical institutions from different areas will enable the results to be standardized, enhancing the usability of the TTCC.

In order to strengthen the benefit of the TTCC as a screening test used worldwide, it is also worth investigating its relevance to biological indices, such as functional neuroimaging findings, in future studies.

ACKNOWLEDGMENTS

The authors thank the staff of Tottori Seikyo Hospital for their help with this study.

REFERENCES

1 United Nations. *World Population on Aging 1950–2050*. New York: United Nations Publications, 2002.
2 Doody RS, Geldmacher DS, Gordon B, Perdomo CA, Pratt RD. Open-label multicenter, phase 3 extension study of the safety and efficacy of donepezil in patients with Alzheimer disease. *Arch Neurol* 2001; **58**: 427–433.
3 Birmingham K, Frantz S. Set back to Alzheimer vaccine studies. *Nat Med* 2002; **8**: 199–200.
4 Urakami K. Preventive medicine for Alzheimer's disease. *Anti Ageing Igaku* 2007; **3**: 304–308 (in Japanese).
5 Yatomi N. Intervention program for cognitive impairment in its early or prodromal stage. *Rhonen Seishin Igaku Zasshi* 2003; **14**: 20–25 (in Japanese).
6 Bozoki A, Giordani B, Heidebrink JL, Berent S, Foster NL. Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss. *Arch Neurol* 2001; **58**: 411–416.
7 Richards M, Touchon J, Ledesert B *et al.* Cognitive decline in ageing: Are AAMI and AACD distinct entities? *Int J Geriatr Psychiatry* 1999; **14**: 534–540.
8 Fratiglioni L, Palliard-Borg S, Winblad B. An active and socially integrated lifestyle in late life might protect against dementia. *Lancet Neurol* 2004; **3**: 343–353.

9 Takeda S, Tajime K, Murakami Y. Development of Takeda Three Colors Combination Test. *Rhonen Seishin Igaku Zasshi* 2004; **15**: 957–963 (in Japanese).
10 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. Washington, DC: American Psychiatric Association, 1994.
11 Hughes CP, Berg L, Danziger WL *et al.* A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982; **140**: 566–572.
12 Kato S, Shimogaki H, Onodera A *et al.* Development of the revised version of Hasegawa's Dementia Scale (HDS-R). *Rhonen Seishin Igaku Zasshi* 1991; **2**: 1339–1347 (in Japanese).
13 Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: Clinical characterization and outcome. *Arch Neurol* 1999; **56**: 303–308.
14 Kashima H. Distinguishing and diagnosing cognitive impairment in Alzheimer's types. In: The Committee for Manual of Diagnosis and Treatment for Cognitive Impairment in Alzheimer's Types (eds). *Manual of Diagnosis and Treatment of Cognitive Impairment in Alzheimer's Types*. Tokyo: World Planning, 2001; 101–128 (in Japanese).
15 Nakamura S, Shigeta M, Iwamoto M *et al.* Prevalence and predominance of Alzheimer type dementia in rural Japan. *Psychogeriatrics* 2003; **3**: 97–103.
16 Furuta N, Mimura M. The cognitive dysfunction of the Alzheimer's disease in early stages. *Rhonen Seishin Igaku Zasshi* 2006; **17**: 385–392 (in Japanese).
17 Nakamura Y, Takeda M. The symptom of the Alzheimer's disease. In: The Committee for Manual of Diagnosis and Treatment for Cognitive Impairment in Alzheimer's Types (eds). *Manual of Diagnosis and Treatment of Cognitive Impairment in Alzheimer's Types*. Tokyo: World Planning, 2001; 13–41 (in Japanese).
18 Tanaka Y, Hashimoto R. Episodic memory. In: Asai M, Kashima H (eds). *The Clinical Psychiatry Lecture S2: Clinical Memory*. Tokyo: Nakayama Syoten, 1999; 75–87 (in Japanese).
19 Kashima H. Small test for dementia of Alzheimer's type. *Psychogeriatrics* 2007; **7**: 1–3.
20 Graham NL, Emery T, Hodges JR. Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia. *J Neurol Neurosurg Psychiatry* 2004; **75**: 61–71.
21 Meulen EF, Schmand B, van Campen JP *et al.* The seven minute screen: A neurocognitive screening test highly sensitive to various types of dementia. *J Neurol Neurosurg Psychiatry* 2004; **75**: 700–705.
22 Sakai Y, Kotaka A, Murayama N *et al.* Japanese version of the Rapid Dementia Screening Test: Effectiveness in detecting possible patients with dementia. *Rhonen Seishin Igaku Zasshi* 2006; **17**: 539–551 (in Japanese).
23 Takayama Y. Necessary conditions of screening tests for early detection of dementia. *Rhonen Seishin Igaku Zasshi* 2003; **14**: 13–19 (in Japanese).
24 Collie A, Maruff P, Currie J. Behavioral characterization of mild cognitive impairment. *J Clin Exp Neuropsycol* 2002; **24**: 720–733.
25 Comijs HC, Dik MG, Deeg DJH, Jonker C. The course of cognitive decline in older persons: Results from the longitudinal aging study Amsterdam. *Dement Geriatr Cogn Disord* 2004; **17**: 136–142.

Psychiatry and Clinical Neurosciences 2010; 64: 44–51 doi:10.1111/j.1440-1819.2009.02049.x

Regular Article

Does daily Naikan therapy maintain the efficacy of intensive Naikan therapy against depression?

Mari Sengoku, PhD,[1]* Hiroaki Murata, PhD,[2] Takanobu Kawahara, PhD,[2] Kaori Imamura, PhD[2] and Kazuyuki Nakagome, MD, PhD[1]

[1]*Division of Neuropsychiatry, Department of Multidisciplinary Internal Medicine, Tottori University Faculty of Medicine, Tottori, and* [2]*Kunimi Research Institute of Internal Medicine/Midorigaoka Mental Clinic, Kagoshima, Japan*

Aim: Naikan Therapy, which has been applied to treating patients with various mental difficulties, can be classified into two major categories: intensive Naikan therapy, which lasts for seven days in a Naikan center or a clinical institute secluded from the outside world for the purpose of deep introspection, and daily Naikan therapy, which can be integrated into regular daily activities. The aim of this research is to evaluate daily Naikan therapy as a maintenance treatment for depression.

Methods: Forty-seven patients, who were diagnosed as having major depressive disorder using DSM-IV criteria and who practiced intensive Naikan therapy participated in the present study. Two groups of patients were compared: 24 patients who conducted daily Naikan therapy and 23 patients who did not, after practicing intensive Naikan therapy. To evaluate efficacy, the Beck Depression Inventory was used as a primary outcome measure for the assessment of depression. The State–Trait Anxiety Inventory and the Cornell Medical Index were also used as secondary outcome measures to evaluate anxiety and psychosomatic conditions before, immediately after and three months after intensive Naikan therapy.

Results: Significant between-group differences were obtained in the time course change of depression, anxiety and psychosomatic scores within three months following the completion of intensive Naikan therapy.

Conclusion: The current study indicates that conducting daily Naikan therapy is effective for maintaining the psychological and psychosomatic state at 3 months following the intensive Naikan therapy, while a lack of therapy may allow the patients to exacerbate their conditions to the level they held before practicing intensive Naikan therapy.

Key words: anxiety, daily Naikan therapy, depression, intensive Naikan therapy, psychosomatic condition.

MAJOR DEPRESSION MAY require long-term treatment because it is a debilitating and recurrent disorder.[1] Efficacious alternatives to medication, including psychological intervention, are necessary, especially for patients who may not tolerate or respond to medication. Approximately 40% of patients with depression discontinue their medication within the initial 30 days, and the dropout rate even reaches 72% within 90 days after starting medication.[2–4] A meta-analysis study comparing the efficacy between pharmacotherapy and psychotherapy in patients with depression reported that both were about equally effective, whereas the dropout rates seemed to be smaller in psychotherapy compared with pharmacotherapy.[5] Moreover, a systematic review comparing both the efficacy and adherence to antidepressant medication use between antidepressant treatment in combination with a psychological intervention and antidepressant treatment

Correspondence: Mari Sengoku, PhD, Division of Neuropsychiatry, Department of Multidisciplinary Internal Medicine, Tottori University Faculty of Medicine, 36-1 Nishi-cho, Yonago, Tottori 683-8504, Japan. Email: sengokumari@live.jp
Received 4 February 2009; revised 4 October 2009; accepted 12 November 2009.

alone indicated that the former was associated with a higher improvement rate that the latter.[6] In longer therapies, the combination with psychotherapy significantly reduced the dropout rates. However, most studies cited in the above reviews are those from Western countries, and the psychological interventions adopted are cognitive–behavioral therapy, interpersonal psychotherapy, problem-solving therapy and others.

Naikan was established by Ishin Yoshimoto (1916–1988), who was a Jodo Shin Buddhist minister. Naikan literally means 'looking within'. Naikan was originally rooted in a spiritual practice called *Mishirabe*, which was practiced among Jodo Shin Buddhist followers in a rural district in Nara, Japan, where Yoshimoto was born. In the practice of *Mishirabe*, in addition to introspection, the practitioners fasted completely even without liquids, and went without sleep for long periods to reflect on their thoughts and behaviors. What Yoshimoto discovered through this practice was the degree to which his life had been sustained by others, by all living creatures and ultimately by nature instead of only through self-centeredness and, by extension, ignorance. Based on his experience, Yoshimoto realized the possible use of this therapy for all people, not only Buddhist practitioners. So, he established a relatively easier method by eliminating the strict physical restrictions and formed three themes for 'healthy' practitioners to be discussed later. Nowadays, the main goal of Naikan is not to achieve religious enlightenment, but to experience self-understanding and self-improvement, goals that can be applied to any person regardless of religion, ethnic group or social background. Naikan therapy (NT) has been adapted to prisons, detention homes, schools and business training. In the field of psychiatry, it has been reported that Naikan as psychotherapy is an efficient means of treating patients with various mental difficulties.[7]

In intensive Naikan therapy (INT), the patient sits in the corner of a room, walled off by a folding screen to cut off visual stimulation from the outside so that it is easier for them to observe their own thinking. Sitting in a quiet place and staying in a relaxed position, the patient begins to seriously look into his/her thoughts, continuing his/her introspection daily from 06.00 hours to 21.00 hours. The patients examine how they have lived according to three themes: (i) What have I received from a particular person? (ii) What have I returned to that person? and (iii) What troubles and difficulties have I caused that person? To begin with, the patients are asked to examine the relationship with their mothers or their main caretakers through every period of their life, starting from childhood and gradually moving to the present. Then, they are asked to examine themselves regarding other people who are close to them, such as their fathers, spouses, friends, colleagues, and so forth.

In their articles, both Tashiro,[8] who studied the efficacy of INT for prolonged depression, and Nukina,[9] who presented the efficacy of INT for general anxiety disorders and panic disorders, have referred to the importance of daily Naikan therapy (DNT) as an important factor for preventing the recurrence of various mental disorders, although detailed data have not yet been provided. Yoshimoto, the founder of Naikan, emphasized to his clients that doing DNT for at least one hour a day after completing INT would be highly beneficial.[10] He insisted that INT is the preparation for DNT, which should be integrated into daily life. His view was that, although the patients successfully completed INT, their neglect of DNT may cause a quick recurrence of their original perspectives and habits. Yoshimoto's encouragement to do DNT could be explained by his understanding of Dr Ikemi's argument that the neocortex of humans completes its development by the age of 15–16, so the character of each person has generally been formed around that time and is thus very difficult to change. The paleocortex is more fragile and sensitive than the neocortex to various stressful conditions humans are exposed to. These experiences, especially emotional ones experienced by infants through interaction with parents and others, greatly influence the formation of their personalities. The pathetic characteristics of some patients cannot form by experiences and memories from their childhood unless they are recognized and, thus become deeply rooted in their consciousness.[11] Therefore, even though patients experience dramatic transformation by practicing INT, by ignoring DNT, they would be drawn to their previous habits and perspectives and may have additional episodes.

The method of DNT for patients who do it for one hour a day is as follows. Patients examine their past as they did in INT. They divide their lives into chronological segments and reflect on themselves in relation to persons important in their lives for half an hour each day as they did in INT. For example, at the end of the day one patient may examine his/her

relationship with his/her mother from three to six years for half an hour. Then, the patient may use the remainder of the hour examining his/her day according to the three Naikan themes as follows: (i) What did I receive from others today? (ii) What did I give to others? and (iii) What troubles and difficulties did I cause others today? He/she may then go on to examine a later period of life, for example, four to seven years, on the following day for half an hour, then use the remaining half hour to ask himself/herself the three Naikan themes for that day.

The purpose of this research is to evaluate the efficacy of DNT as a maintenance treatment for depression by comparing patients: (i) who experienced INT and continued DNT, with (ii) those patients who did not conduct DNT after practicing INT. To evaluate efficacy, the Beck Depression Inventory (BDI) was used as a primary outcome measure and the State–Trait Anxiety Inventory (STAI) and the Cornell Medical Index (CMI) were used as secondary outcome measures before, immediately after and three months after INT.

METHODS

Subjects

The subjects were 47 inpatients who were diagnosed as having major depressive disorder through interviews and close examination according to DSM-IV criteria[12] by an experienced psychiatrist. All subjects underwent INT at the Midorigaoka Mental Health Clinic between July 2006 and February 2008. The patients with strong suicidal tendencies, paranoia and severe depression requiring psychophysical restraint were excluded. All of the subjects gave written informed consent before participating.

Study design

All subjects had completed INT shortly before they were discharged from the Midorigaoka Mental Health Clinic. Immediately after the completion of INT and at discharge, all patients were advised to practice DNT. However, 24 patients consented to practice daily Naikan (DNT group) while 23 did not (non-DNT group) owing to participant-determined attendance. Fifteen out of 24 patients in the DNT group were admitted to hospital with the aim of undergoing INT for 7 days, which was designated as 8 days for the length of hospitalization. As for the nine patients, they were recommended to practice INT during their stay in the hospital and the average length of hospitalization of these nine patients was 58.3 days. Thirteen out of 23 patients in the non-DNT group were admitted to hospital with the aim of undergoing INT. The rest of the patients were recommended to practice INT during their stay in hospital and the average length of hospitalization of these subjects was 50.2 days. There was no significant between-group difference either in the rate of the patients admitted to hospital with the aim of undergoing INT or the length of hospitalization. Other clinical and demographic backgrounds at baseline for both groups immediately after INT completion are shown in Table 1. All subjects were assessed immediately before and after INT, and after the three months' trial starting shortly after INT completion for depression, anxiety and psychosomatic symptoms severity. The DNT group complied fairly well with the one-hour/day session of DNT within the three months' period and the average rate of practicing the therapy was 3.5 days/week. Twenty-two out of 24 patients in the DNT group were under medication during the trial, which was similar to 22 out of 23 in the non-DNT group.

Assessments

1 Psychological scales

The BDI,[13] the STAI,[14] and the CMI[15] tests were used for evaluation. The CMI was adopted to observe whether physical functions of patients would be normalized as patients became mentally stable from the perspective of psychosomatic medicine. In cases where NT helped to make patients mentally stable, stress would be alleviated and physical functions would also be normal.[16] Therefore, various subjective somatic symptoms (such as hearing, visual, cardiac, respiratory, gastro, dermatological, neural, urological and fatigue symptoms) and subjective psychological symptoms of depression (such as anxiety, tension, hypersensitivity, indignation and inadaptability) may differ between those patients practicing DNT and those not practicing DNT.

2 Quality of Naikan: Evaluation after the completion of INT in terms of psychological transformation

It is desirable in NT to recall events in the past by taking another person's perspective along with seri-

Table 1. Clinical and demographic background of the subjects

	DNT	non-DNT	*P*-value
Number of subjects	24	23	
Male subjects	9	16	$P < 0.05$†
Mean age (SD)	38.5 (11.0)	32.7 (12.5)	n.s.‡
Subjects with single episode	16	17	n.s.†
Mean years of illness duration (SD)	1.7 (1.7)	1.7 (1.5)	n.s.‡
Frequency of practicing DNT			
Everyday	3	0	
5–6 times per week	2	0	
3–4 times per week	13	0	
1–2 times per week	6	0	
Mean BDI scores (SD) before INT	25.1 (7.8)	25.5 (10.4)	n.s.‡
Mean STAI trait scores (SD) before INT	51.0 (12.5)	58.5 (13.7)	n.s.‡
Mean STAI state scores (SD) before INT	49.5 (12.2)	52.7 (17.3)	n.s.‡
Mean CMI-soma scores (SD) before INT	27.0 (18.5)	40.9 (21.7)	$P < 0.05$‡
Mean CMI-psyche scores (SD) before INT	16.7 (10.8)	22.3 (13.5)	n.s.‡

†Chi-square test.
‡Wilcoxon's rank sum test.
BDI, Beck Depression Inventory; CMI-psyche, Cornell Medical Index, psychological; CMI-soma, Cornell Medical Index, somatic; DNT, daily Naikan therapy; INT, intensive Naikan therapy; n.s., not significant; SD, standard deviation; STAI, State–Trait Anxiety Inventory.

ously considering the three already mentioned main themes. As a result, patients may realize that they have received so much love and care from others (feelings of being loved), and will recognize a self-centeredness that they never realized before. When they seriously feel guilt and remorse, they may decide that they would like to discard their ego-centeredness. As discussed above, in terms of psychological transformation, the factors for evaluation are: (i) seeing things from another person's perspective; (ii) the feeling of being loved; (iii) awareness of ego-centeredness; and (iv) the decision to discard ego-centeredness. These four categories are evaluated as shown in Table 2.[17]

In addition to the above four categories, questions as to whether the patients had a sense of accomplishment and self-acceptance after INT were also added for evaluation.

Table 2. Criteria for evaluation of the four categories of psychological transformation

Psychological transformation		Score
Seeing things from another person's perspective	seeing things only from the patient's perspective	0
	seeing things objectively without personal emotions	1
	seeing things from the other person's point of view	2
Feelings of being loved	no feelings of being loved	0
	feelings of being loved	1
	feelings of being loved unconditionally despite the problems to the person	2
Awareness of ego-centeredness	no ability to be aware of ego-centeredness	0
	only aware of ego-centeredness	1
	awareness of ego-centeredness and remorseful and/or guilty feelings	2
Decision to discard ego-centeredness	no ability to be aware of ego-centeredness	0
	only aware of ego-centeredness	1
	a decision to discard ego-centeredness	2

Data analysis

First between-group differences were tested using χ^2-tests for dichotomous variables and Wilcoxon rank sum tests for continuous variables in terms of clinical and demographic backgrounds of the DNT and non-DNT groups. Moreover, assessments regarding psychological transformation after INT were compared between the two groups. In addition, to assure the efficacy of INT for depression, we calculated the remission rate according to the definition of the BDI $\leqq 8$.[18] Next, a repeated ANCOVA measure was performed using the BDI, STAI-trait (T), STAI-state (S), CMI-somatic (soma) and CMI-psychological (psyche) scores as dependent variables; group and sex as interindividual factors; time of assessment (before INT, immediately after INT, 3 months after INT) as an intraindividual factor; and the clinical, demographic and psychological transformation data as covariates if they showed significant between-group differences either before or immediately after INT. For the variables that showed significant interaction between group and time, the difference scores between the first and second assessments (difference-INT) and those between the second and third assessments (difference-DNT) were submitted for secondary ANCOVA.

As depression, anxiety and psychosomatic symptoms are assumed to hold a correlative relationship with each other, the effect of DNT on depression may well be obtained indirectly via its efficacy on anxiety and/or psychosomatic symptoms. With the aim of investigating whether the effect of DNT on depression is independent from its efficacy on anxiety and/or psychosomatic symptoms, we calculated Spearman's rank correlation coefficient between each pair of the difference-DNT variables of BDI, STAI-T, STAI-S, CMI-soma and CMI-psyche scores. Next, in case significant correlations were observed between at least some pairs, we reanalyzed whether significant between-group differences could be obtained in difference-DNT BDI scores, including those variables as covariates that showed a significant relationship with difference-DNT BDI scores in additional ANCOVA.

RESULTS

First, the number of patients that attained the remission criteria immediately after INT was 18 (38.3 %) in total. A significant between-group difference was scored for the number of remitted patients, that is, there were 13 out of 24 patients (54.1%) in the DNT group whereas there were five out of 23 (21.7%) in the non-DNT group ($P < 0.05$, χ^2-test). Next, as there were significant between-group differences in sex, CMI-soma scores assessed before INT, and the scores for 'feelings of being loved' and 'the decision to discard ego-centeredness' after INT (see Table 3), we adopted the latter three variables as covariates and added sex as an interindividual variable in the following repeated measures ANCOVA.

BDI

There was a significant interaction between group and time ($F [2, 80] = 23.31$, $P < 0.0001$) along with a significant main effect of time. The secondary analyses revealed significant between-group differences in both difference-INT ($F [1, 40] = 4.79$, $P < 0.05$) and difference-DNT scores ($F [1, 40] = 34.72$, $P < 0.0001$) (Fig. 1).

Table 3. Comparison of the psychological transformation after INT between the DNT and non-DNT subjects

	DNT	non-DNT	*P*-value
Psychological transformation			
Seeing things from another person's perspective; mean (SD)	2.7 (0.5)	2.4 (0.8)	n.s.‡
Feelings of being loved; mean (SD)	2.7 (0.5)	2.3 (0.8)	$P < 0.05$‡
Awareness of ego-centeredness; mean (SD)	2.7 (0.5)	2.5 (0.6)	n.s.‡
Decision to discard ego-centeredness; mean (SD)	2.5 (0.7)	2.0 (0.7)	$P < 0.05$‡
Number of subjects who gained a sense of accomplishment	24	21	n.s.†
Number of subjects who gained a sense of self-acceptance	22	18	n.s.†

†Chi-square test.
‡Wilcoxon's rank sum test.
DNT, daily Naikan therapy; INT, intensive Naikan therapy; n.s., not significant; SD, standard deviation.

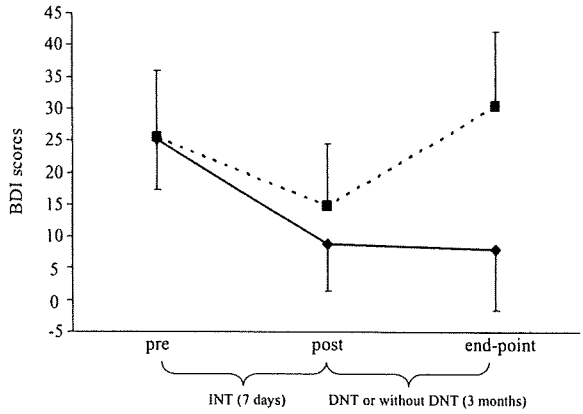

Figure 1. Comparison of Beck Depression Inventory (BDI) scores between (-◆-) the daily Naikan therapy (DNT) and (-■-) non-DNT subjects at pre, post and 3 months following intensive Naikan therapy (INT). INT was conducted for 7 days between 'pre' and 'post', and DNT for 3 months between 'post' and 'end-point' for DNT subjects but not for non-DNT subjects. The vertical bars indicate standard deviations. Note the improvement in DNT subjects exceeded that in non-DNT subjects in either difference-INT (pre-post) or difference-DNT (post-end-point) scores.

STAI

For either STAI-T or STAI-S scores, there was a significant interaction between group and time (Table 4). The secondary analyses revealed a significant between-group difference in difference-DNT for either STAI-T ($F [1, 40] = 17.71$, $P < 0.0001$) or STAI-S scores ($F [1, 40] = 19.49$, $P < 0.0001$), but not in difference-INT scores.

CMI

For either CMI-soma scores or CMI-psyche scores, there was a significant interaction between group and time (Table 4). The secondary analyses revealed a significant between-group difference in difference-DNT for either CMI-soma ($F [1, 40] = 22.81$, $P < 0.0001$) or CMI-psyche scores ($F [1, 40] = 14.05$, $P < 0.001$), but not in difference-INT scores.

Relationship between difference-DNT BDI, STAI and CMI scores

Spearman's rank correlation coefficient between each pair of the difference-DNT variables of BDI, STAI-T, STAI-S, CMI-soma and CMI-psyche scores revealed that all the pairs showed significant correlations (rho = 0.45–0.70). Additional ANCOVA using difference-DNT BDI scores as dependent variables, including those variables as covariates that showed a significant relationship with difference-DNT BDI scores revealed the non-significant effect of the group ($F [1, 36] = 3.27$, $P < 0.1$).

DISCUSSION

Despite the shortcomings in the present study design as noted below, whereas BDI showed significant between-group differences in the improvement during both DNT and INT, significant between-group

Table 4. Comparison of secondary outcomes between the DNT and non-DNT subjects at pre, post and 3 months following intensive Naikan therapy

Measures	Group	*time*	Pre	Post	End-point	Analysis
STAI-T	DNT		51.0 (12.5)	43.5 (13.5)	37.1 (10.3)	F (2, 80) = 11.71
	non-DNT		58.5 (13.7)	51.7 (15.1)	58.9 (12.6)	$P < 0.0001$
STAI-S	DNT		49.5 (12.2)	32.0 (9.9)	35.2 (11.7)	F (2, 80) = 9.24
	non-DNT		52.7 (17.3)	38.1 (14.8)	56.2 (14.6)	$P < 0.0005$
CMI-soma	DNT		27.0 (18.5)	20.9 (18.8)	19.4 (16.0)	F (2, 80) = 12.92
	non-DNT		40.9 (21.7)	30.8 (18.6)	46.6 (24.0)	$P < 0.0001$
CMI-psyche	DNT		16.7 (10.8)	9.8 (10.9)	8.1 (10.1)	F (2, 80) = 7.24
	non-DNT		22.3 (13.5)	16.9 (10.5)	24.3 (13.7)	$P < 0.005$

Analysis: interaction between time and group in repeated measures ANCOVA.
CMI-psyche, Cornell Medical Index, psychological; CMI-soma, Cornell Medical Index, somatic; DNT, daily Naikan therapy; STAI-S, State-Trait Anxiety Inventory-State; STAI-T, State-Trait Anxiety Inventory-Trait.