Quantitative and Qualitative Analyses of the Clock Drawing Test in Fall and Non-Fall Patients with Alzheimer’s Disease

Yukiko Suzuki a Hideki Mochizuki a Mayuka Oki b Miyuki Matsumoto c Mitsuko Fukushima c Yukiko Yoshikawa c Akira Nagasawa c Tomokazu Takakura c Nobuaki Shimoda d

Department of Occupational Therapy, Kyorin University, Tokyo, Japan; b Department of Rehabilitation Medicine, Juntendo University Hospital, Tokyo, Japan; c Department of Rehabilitation Medicine, Juntendo Tokyo Koto Geriatric Medical Center, Tokyo, Japan; d Department of Rehabilitation, Faculty of Health Sciences, Tokyo Kasei University, Saitama, Japan

Keywords
Alzheimer’s disease · Fall · Clock drawing test · Qualitative analysis

Abstract
Aim: The clock drawing test (CDT) is widely used as a visual spatial ability test and screening test for dementia patients. The appearance frequency of qualitative errors obtained through the qualitative analysis of CDT may be related to the participant’s falls. The aim of this study was to clarify the difference in the number of people who presented with qualitative errors in the CDT between a fall and non-fall group of patients with Alzheimer’s disease (AD). Methods: The CDT was implemented for 47 patients with AD. A quantitative analysis was conducted, and a qualitative analysis was performed for errors. The patients were divided into two groups based on their history of falls over the past year. The results of the CDT quantitative analysis were tested using the Mann-Whitney U test, and Fisher’s exact test was employed to determine the difference in the number of people who presented with error types between the two groups (fall group, non-fall group) in the CDT qualitative analysis. Results: In the quantitative analysis, a significant difference was found for the total scores, with the total CDT score of the fall group (n = 22) significantly lower than that of the non-fall group (n = 25) (p = 0.006, effect size: φ = 0.40). In the qualitative analysis, a significantly higher number of patients in the fall group than in the non-fall group presented with a conceptual deficit (p = 0.001, φ = 0.51). No differences were found in the number of patients in the two groups who presented with the other five error types. Conclusions: These results showed that a lower score in the CDT quantitative analysis might suggest an increased risk of falls. It was also clarified that a larger
number of patients in the fall group than in the non-fall group presented with a conceptual
deficit of the qualitative error types in the CDT. Therefore, these results suggest that the ap-
pearance of a conceptual deficit may be an index for the selection of patients with AD prone
to falling when implementing fall prevention measures.

Introduction

In patients with Alzheimer’s disease (AD), persons who fall easily account for slightly more
than 50% of all dementia patients [1]. This is because falls by dementia patients reduce motor
function, further promote dementia [2], and lead to difficulty in maintaining independence [3].

Researchers used the clock drawing test (CDT) as the evaluation method used to
determine fall-related factors in patients with AD [4, 5]. In the CDT, which is widely used as a
visual spatial ability test [6, 7] and screening test for dementia patients [8–15], the participant
is asked to draw a clock face. The advantages of the CDT include that it can be implemented
in a short period, participants have little resistance to the test, it tends to not be affected by
participants’ education level [16], and it can measure executive function, which is difficult to
evaluate with the Mini-Mental State Examination (MMSE) [17]. The methods used to analyze
the results of the CDT include quantitative and qualitative analyses. Maruta and Yoshikawa
[10] (2014), Nishiwaki et al. [11] (2004), Yoo and Lee [15] (2016), and Jeong et al. [18] (2016)
analyzed the CDT quantitatively and reported its utility as a screening evaluation of dementia.
Diegelman et al. [19] (2004) reported that the quantitative analysis of CDT was significantly
worse among non-drivers than individuals who were currently or recently driving. In this
way, many previous studies have used a quantitative analysis [10, 11, 15, 19–21]. However,
qualitative analysis reflects the severity of dementia better than quantitative analysis [22, 23]
and enables capturing the characteristics of the disease [7, 24]. Thus, qualitative analysis can
also be regarded as an important analysis method.

The qualitative analysis of the CDT can measure the appearance frequency of six quali-
tative errors (size of the clock, graphic difficulties, stimulus-bound response, conceptual
deficit, spatial/planning deficit, perseveration) [7]. The appearance frequency of these errors
may be related to participant falls. However, this has not been investigated in previous studies.

The aim of this study was to clarify the difference between the number of patients who
present with qualitative errors in the CDT between patients with AD in the fall and non-fall
groups.

Materials and Methods

Participants

The participants included 47 patients with AD (21 men and 26 women, mean age 80.7 ±
6.3 years) who were admitted to the dementia ward of the university hospital of Japan. The
selection criteria were patients with AD diagnosed by a doctor based on ICD-10, who were
able to respond to verbal commands, and whose condition had more or less not changed since
hospitalization. Exclusion criteria were the following six items (which may be the cause of falls): patients who (1) were unable to walk 50 m, (2) had motor paralysis, (3) had sensory
impairment, (4) had a history of disturbance or loss of consciousness, (5) had significant
visual impairment, and (6) were otherwise deemed unsuitable by the authors. Complications
included brain and nervous system diseases (n = 32), medical diseases (n = 40), orthopedic
diseases (n = 19), and other diseases (ophthalmological diseases after cataract surgery,
dermatological diseases, etc.) \( (n = 19) \) (multiple choices permitted). Information regarding the participants characteristics (sex, age, duration of AD, number of years of education after graduating elementary school, complications, number of types of medicines) was obtained from hospitalization medical records (Table 1). The aims of the study were explained verbally and in writing to participants and relatives living with the participants, and consent to participate in the study was obtained.

**Evaluation**

In the CDT, participants were presented with an A4-size blank sheet of paper, and instructed to draw a clock face showing 11:10, with all the numbers written on the clock dial.

### Table 1. Participant characteristics and evaluation results

|                              | Total (\( n = 47 \)) | Fall group (\( n = 22 \)) | Non-fall group (\( n = 25 \)) | \( p \) value |
|------------------------------|-----------------------|---------------------------|-------------------------------|--------------|
| Men/women\(^1\)             | 21/26                 | 7/15                      | 14/11                         | 0.143        |
| Age, years\(^2\)            | 80.7±6.3              | 82.5±6.1                  | 79.2±6.1                      | 0.075        |
| Duration of AD, months\(^3\) | 28 (2–144)            | 18 (2–132)                | 34 (3–144)                    | 0.564        |
| Education, years\(^3\)      | 6 (0–10)              | 6 (0–10)                  | 6 (3–10)                      | 0.352        |
| Complications\(^3\)         | 6 (0–35)              | 7 (0–35)                  | 5 (1–27)                      | 0.487        |
| Types of oral medicines\(^3\) | 5 (1–12)             | 5 (1–11)                  | 5 (1–12)                      | 0.889        |
| MMSE-J, points\(^2\)        | 16.3±5.6              | 15.9±5.6                  | 16.7±5.6                      | 0.623        |
| NPI, points\(^3\)           | 6 (0–40)              | 6 (0–39)                  | 6 (0–40)                      | 0.723        |
| FIM, points\(^2\)           | 96.2±14.5             | 92.5±17.7                 | 99.4±10.3                     | 0.113        |

Data are presented as number of persons, mean ± standard deviation, or median (range), as appropriate. AD, Alzheimer’s disease; MMSE-J, Japanese version of Mini-Mental State Examination; NPI, Neuropsychiatric Inventory; FIM, Functional Independence Measure. \(^1\) Fisher’s exact test. \(^2\) Welch’s \( t \) test. \(^3\) Mann-Whitney U test.

### Table 2. Revised scale used for scoring the clock drawings

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

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

- **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 (significantly out of course including 10 to 11)
  - 1: Only one hand or poor representation of two hands
  - 0: No hands or perseveration on hands

From Rouleau et al. [7] (1992).
The scoring of the quantitative analysis used the method proposed by Rouleau et al. [7] (Table 2), and the test was scored with 10 points as full marks. A higher score indicates better function. For the qualitative analysis, the method proposed by Rouleau et al. [7] was also adopted, and each participant was evaluated based on the appearance of the six error types (size of the clock, graphic difficulties, stimulus-bound response, conceptual deficit, spatial/planning deficit, perseveration). The specific analysis method used for the appearance of the six error types is shown in Table 3. Two research collaborators (occupational therapists with 13 and 16 years of experience) were asked to evaluate the error types. The research collaborators were not told the participant’s diagnosis or aim of the study, and they determined the appearance of each error type. When the judgment of the two research collaborators differed, it was determined to not be an error, and only matching judgments were set as errors [22] (Fig. 1).

Cognitive function (Japanese version of the MMSE, MMSE-J) [25], behavioral and psychological symptoms of dementia (BPSD) (Neuropsychiatric Inventory, NPI) [26], and activities of daily living (Functional Independence Measure version 3, FIM) [27] were also evaluated.

MMSE-J [25] is a simple test of cognitive functions. It consists of 11 items: orientation to time, orientation to place, attention and concentration, memory registration, memory recall, naming, repetition, comprehension, reading, writing, and drawing. It has a maximum score of 30 points, with higher scores indicating better cognitive function.

NPI asks a caregiver a question and evaluates the onset frequency of BPSD and the severity of dementia. The NPI is comprised of the following domains: delusions, hallucination, agitation, depression, anxiety, euphoria, apathy, disinhibition, irritability, aberrant motor behavior, night-time behavior disturbances, and appetite and eating abnormalities. It has a maximum score of 144 points, with higher scores indicating worse BPSD symptoms.

FIM consists of 18 items in total: 13 movement items and 5 cognitive items. Movement items include 6 self-care, 2 excretion control, 3 transferring, and 2 moving items. Cognitive items include 2 communication and 3 social cognition items. Each item is evaluated independently on a scale of 1 to 7 according to the degree of assistance required. It has a maximum score of 126 points, with higher scores indicating less assistance with day-to-day activities.

Procedure

All evaluations were implemented for all participants by physical therapists or occupational therapists within 1 week after hospitalization. During the evaluation period, there were no changes in the participants’ oral medicines or condition.

Falls were defined as “any unintentional action where a part of the body other than the soles of the feet comes into contact with the ground or lower location” [28]. Falls from steps, platforms, and bicycles were included in the definition. Participants who had one or more falls within 1 year preceding the survey date were classified as the fall group, and those without this fall history as the non-fall group. This information was obtained by asking the main caregiver, who was able to ascertain the participant’s lifestyle situation up to immediately before hospitalization.

Data Analysis

After confirming the normality of the variables related to characteristics and scoring based on the quantitative analysis of CDT using the Shapiro-Wilk test, age, MMSE-J, and FIM followed a normal distribution. Therefore, the Welch t test was used to compare the age, MMSE-J, and FIM of the two groups based on their history of falls (fall group vs. non-fall group). The duration of AD, number of years of education, number of complications, number of types of oral medicines, NPI, and scoring-based quantitative analysis of CDT did not have a normal distribution; thus, the Mann-Whitney U test was used for the analysis. Fisher’s exact test was performed to determine the difference in the number of people who presented with
Fig. 1. Actual examples of error types in qualitative analysis of the clock drawing test. There are no actual examples for size of the clock error types.

Table 3. Qualitative error analysis of clock drawing test

1. Size of the clock
   A clock was considered small if it measured less than 1.5 inches and large if it measured more than 5 inches
2. Graphic difficulties
   Mild: Some distortions in tracing the clockface and/or the hands and/or the numbers were present, but the overall performance remained adequate
   Moderate: Distortions were evident, but the overall performance remained interpretable
   Severe: Distortions were evident and precluded, in some cases, the unequivocal interpretation of the overall performance
3. Stimulus-bound response
   A: The minute hand pointed toward 11 instead of 2
   B: The time (min) is written in letters and numbers besides the 11 on the clock
4. Conceptual deficit
   A: Misrepresentation of the clock itself (only a clockface without numbers or inappropriate use of numbers), suggesting the unavailability of a correct graphic representation of a clock
   B: Misrepresentation of the time on the clock: the hands are either absent or inadequately represented; the time is written on the clock
5. Spatial and/or planning deficit
   A: Neglect of the left hemispace
   B: Deficit in planning, with gap before 12 or 3, 6, or 9 depending on the strategy used in drawing
   C: Deficit in spatial layout of numbers, without any specific pattern in spatial disorganization
   D: Numbers written outside the disorganization
   D: Numbers written outside the clockface
   E: Numbers written counterclockwise
6. Perseveration
   A: Perseveration of hands
   B: Perseveration of numbers

From Rouleau et al. [7] (1992). In points 2–6, when any one was applicable, it was judged as “with.”
error types between the two groups (fall group vs. non-fall group) for the qualitative analysis of CDT. The effect size ($\phi$) was calculated for the comparison of the quantitative and qualitative CDT analyses of the two groups (fall group vs. non-fall group). The index proposed by Cohen [29] was referenced to interpret the effect size, with 0.1–0.29 interpreted as low, 0.3–0.49 as moderate, and $\geq 0.5$ as high.

SPSS version 21.0 (IBM Corporation, Armonk, USA) was used for the statistical analysis, and the level of significance was set as less than 5%.

**Results**

The participant characteristics are shown in Table 1. Of the 47 participants (21 men and 26 women, mean age 80.7 ± 6.3, mean number of years of education 5.2 ± 2.7 years), 22 had fallen within the past year (46.8%). There was no significant difference in sex, age, AD duration, number of years of education, number of complications, number of types of oral medicines, MMSE-J, NPI, and FIM between the fall and non-fall groups.

There was a significant difference in the total scores for the quantitative analysis of CDT, with the fall group ($n = 22$) showing a significantly lower total CDT score than the non-fall group ($n = 25$) ($p = 0.006$, effect size: $\phi = 0.40$).

In the qualitative analysis of CDT, significantly more participants in the fall group than in the non-fall group presented with a conceptual deficit ($p = 0.001$, $\phi = 0.51$). Conversely, there was no significant difference between the two groups for the size of the clock ($p = 0.690$, $\phi = 0.09$), graphic difficulties ($p = 0.328$, $\phi = 0.18$), stimulus-bound response ($p = 0.352$, $\phi = 0.19$), spatial/planning deficit ($p = 1.000$, $\phi = 0.03$), and perseveration ($p = 0.270$, $\phi = 0.19$) (Table 4).

**Discussion**

In this study, patients with AD were divided into two groups, the fall group and non-fall group, and the differences revealed between the two groups in a quantitative and qualitative analysis of the CDT were investigated. The results of the quantitative analysis indicated that the CDT score of the fall group was significantly lower than that of the non-fall group. The qualitative analysis showed that significantly more participants presented with a conceptual...
deficit in the fall group than in the non-fall group. The effect size ($\phi$) was high at 0.51. There was no difference between the two groups in the number of participants who presented with the other five error types.

Yamada et al. [5] divided 31 patients with mild cognitive impairment into two groups, the fall group and non-fall group, and examined the intergroup differences in scoring through a quantitative analysis of the CDT. They found that the CDT score of the fall group was significantly lower than that of the non-fall group. The results of the quantitative analysis of the CDT in this study (examination of the intergroup difference in CDT scores) were similar to those reported by Yamada et al. [5]. This result indicates that a lower score for the quantitative analysis of the CDT may indicate an increased risk of falls.

In the qualitative analysis, significantly more participants presented with a conceptual deficit in the fall group than in the non-fall group. Specific examples of error judged as a conceptual deficit are numbers or clock hands that are missing or inappropriate, or writing the time using numbers instead of showing the time with the clock hands. These responses reflect a deficit in accessing knowledge about the attributes, characteristics, and meaning of a clock, that is, a lack of semantic association usually triggered by the word clock [7]. Patients with AD demonstrate this type of conceptual deficit in the CDT, which can be expected to increase the risk of falls in their daily activities. Using as an example the scenario of moving from a Western-style room to a Japanese-style room in their own home, a patient with AD may not be able to semantically associate the presence of a step that potentially and overtly manifests when moving between the two rooms. Similarly, when going to the toilet, a patient with AD may not be able to semantically associate this action with the step between the hallway and the toilet, and the action of lifting the lid of a Western-style toilet before sitting down. The risk of falling is expected to increase in these types of situations, which is estimated to be the reason for the large number of participants with a conceptual deficit in the group of AD patients with a history of falls. In the future, we would like to investigate the reasons that link a conceptual deficit with falls by analyzing in detail the fall scenarios of patients with AD that present errors in the CDT based on a conceptual deficit.

There was no significant difference between the fall and non-fall groups in the number of participants who presented with the five error types other than for conceptual deficit (size of the clock, graphic difficulties, stimulus-bound response, spatial/planning deficit, perseveration). These results suggest that the appearance of the five error types other than conceptual deficit in patients with AD has less effect on falls than conceptual deficit.

The limitations of this study include the small cohort. Further investigation of factors that cause falls is needed with more participants. The retrospective nature of the study, namely looking back on past fall history, can also be considered a limitation. A prospective study investigating falls during a set period after the CDT evaluation is preferable to accurately understand the relationship between the results of a qualitative analysis of the CDT and falls.

**Statement of Ethics**

This study was approved by the institutional review board of the Juntendo Tokyo Koto Geriatric Medical Center, Juntendo University Hospital (approval No. 96-4) and the Faculty of Health Sciences, Kyorin University (approval No. 29-64).

**Disclosure Statement**

The authors declare no conflict of interest.
References

1. Alzheimer's Disease International. World Alzheimer Report 2009. London: Alzheimer’s Disease International; 2009.
2. Yukinasa T, Nakano M, Matsuura H, Tamagawa M, Yamanaka M, Kubota M. [The investigation of risk factor of injury caused by falling]. J Biomed Fuzzy Systems Assoc. 2011;13:109–14. Japanese.
3. M. Morris JC, Rubin EH, Morris EJ, Mandel SA. Senile dementia of the Alzheimer’s type: an important risk factor for serious falls. J Gerontol. 1987 Jul; 42:4):412–7.
4. Suzuki Y, Mochizuki H, Oldi M, Matsumoto M, Fukushima M, Nagasawa A, et al. Fall prediction accuracy of visual spatial abilities tests in patients with Alzheimer’s disease: a retrospective study. J Ergon Technol. 2019; 19: 21–34.
5. Yamada M, Takechi H, Mori S, Aoyama T, Arai H. Global brain atrophy is associated with physical performance and the risk of falls in older adults with cognitive impairment. Geriatr Gerontol Int. 2013 Apr;13(2):437–42.
6. Olsson RH Jr, Wambold S, Brock B, Waugh D, Sprague H. Visual spatial abilities and fall risk: an assessment tool for individuals with dementia. J Gerontol Nurs. 2005 Sep;31(9):45–51.
7. Rouleau I, Salmon DP, Butters N, Kennedy C, McGuire K. Quantitative and qualitative analyses of clock drawings in Alzheimer’s and Huntington’s disease. Brain Cogn. 1992 Jan;18(1):70–87.
8. Konagaya Y, Watanabe T, Konagaya M. [Cognitive function screening of community-dwelling elderly people using the clock drawing test - quantitative and qualitative analyses]. Nippon Ronen Igakkai Zasshi. 2012; 49(4):483–90. Japanese.
9. Manos PJ, Wu R. The ten point clock test: a quick screen and grading method for cognitive impairment in medical and surgical patients. Int J Psychiatry Med. 1994;24(3):229–44.
10. Maruta T, Yoshikawa H. [Usefulness of clock drawing test in detection of dementia in brain dock]. Off J Japan Soc Ningen Dock. 2014;29:571–6. Japanese.
11. Nishiwaki Y, Breeze E, Matsuura H, Tamagawa M, Yamanaka M, Kubota M. [Cognitive function screening of community-dwelling elderly people using the clock drawing test - quantitative and qualitative analyses]. Nippon Ronen Igakkai Zasshi. 2012; 49(4):483–90. Japanese.
12. Ricci M, Pigliautile M, D’Ambrosio V, Errolandi S, Bianchini C, Ruggiero C, et al. The clock drawing test as a screening tool in mild cognitive impairment and very mild dementia: a new brief method of scoring and normative data in the elderly. Neurol Sci. 2016 Jun;37(6):867–73.
13. Sato H, Tabata E, Takahashi R, Kato J, Honda T. [About the utilities of the Clock Drawing Test in the rehabilitation setting]. Cogin Rehab. 2010;15:36–44. Japanese.
14. Shulman KI. Clock-drawing: is it the ideal cognitive screening test? Int J Geriatr Psychiatry. 2000 Jun;15(6): 548–61.
15. Yoo DH, Lee JS. Clinical usefulness of the clock drawing test applying rasch analysis in predicting of cognitive impairment. J Phys Ther Sci. 2016 Jul;28(7):2140–3.
16. Hoshino H, Takaqhi Y, Miyaoka H, Takaqhi M. [Clock drawing task as screening for cognitive impairment in the elderly population]. Nippon Ronen Igakkai Zasshi. 1993 Sep;30(9):826–31. Japanese.
17. Royall DR, Cordes JA, Polk M. CLOX: an executive clock drawing task. J Neurol Neurosurg Psychiatry. 1998 May; 64(5):588–94.
18. Jeong HY, Lee JY, Park HK, Oh S, Lee JY. Clock drawing test to screen for dementia in parkinsonian patients with low educational backgrounds. Neurol Asia. 2016;21:357–65.
19. Diegelmam NM, Gilbertson AD, Moore JL, Banou E, Meager MR. Validity of the Clock Drawing Test in predicting reports of driving problems in the elderly. BMC Geriatr. 2004 Oct;4:10.
20. Hubbard EJ, Santini V, Blankevoort CG, Volkers KM, Barrup MS, Byerly L, et al. Clock drawing performance in cognitively normal elderly. Arch Clin Neuropsychol. 2008 May;23(3):295–327.
21. Mittal C, Gorthi CP, Rohatgi MG. Early cognitive impairment: role of Clock Drawing Test. Med J Armed Forces India. 2010;66(1): 25–8.
22. Kitabayashi Y, Ueda H, Narumoto J, Nakamura K, Kita H, Fulari K. Qualitative analyses of clock drawings in Alzheimer’s disease and vascular dementia. Psychiatry Clin Neurosci. 2001 Oct;55(5):485–91.
23. Surh J, Grace J, Allen J, Nadler J, McKenna M. Quantitative and qualitative performance of stroke versus normal elderly on six clock drawing systems. Arch Clin Neuropsychol. 1998 Aug;13(6):495–502.
24. Konagaya Y, Konagaya M, Watanabe T, Washimi Y. [Quantitative and qualitative analyses for characteristics of the clock drawing in Alzheimer’s disease]. Rinsho Shinkeigaku. 2004;54(2):109–15. Japanese.
25. Sugishita M, Koshizuka Y, Sudou S, Sugishita K, Hemmi I, Karasawa H, et al. The validity and reliability of the Japanese version of the Mini-Mental State Examination (MMSE-J) with the original procedure of the attention and calculation task (2001). Ninchi Shinkei Kagaku. 2018;20:91–110.
26. Shigenobu K, Hiroko N, Tabushi K, Ikeda M. [Validity and reliability of the Japanese Version of the Neuropsychiatric Inventory-Nursing Home Version (NPI-NH)]. Brain Nerve. 2008 Dec 60(12):1463–9. Japanese.
27. Data management service of the Uniform Data System for Medical Rehabilitation and the Center for Functional Assessment Research. Guide for use of the uniform data set for medical rehabilitation. version 3.0. Buffalo: State University of New York at Buffalo; 1990.
28. Kondo S, Miyamae T, Tsutsuzumi F. [Fear of falling among elderly persons]. Gen Rehab. 1999;27:775–80. Japanese.
29. Cohen J. A power primer. Psychol Bull. 1992 Jul;112(1):155–9.