Original Research Article

CIBIC Plus-J Assessment Using a Videotaped Method in Alzheimer's Disease Patients

Yu Nakamura^a Mikiko Usui^b Takashi Nishikawa^e
Masashi Takita^f Masahiro Shigeta^c Yukimichi Imai^g
Katsuya Urakami^h Hiroshi Kitaⁱ Akira Homma^d

^aKagawa University, Miki, ^bJiyukai Health Academy, ^cTokyo Metropolitan University, and
^dTokyo Dementia Care Research and Training Center, Tokyo, ^eOsaka Prefecture University,
Habikino, ^fTakita Memory Mental Clinic, Fukuoka, ^gJapan College of Social Work, Kiyose,
^hTottori University, Yonago, and ⁱDainippon Sumitomo Pharma Co., Ltd., Suita, Japan

Key Words

CIBIC plus-J • Dementia • Alzheimer's disease • Reliability

Abstract

Background/Aims: CIBIC plus-J is the Japanese language version equivalent to CIBIC plus. Variability of CIBIC plus-J arises among raters in accordance with their experience and their memories of patients' conditions at baseline. Therefore, in a multicenter trial of Alzheimer's disease, CIBIC plus-J interviews with Alzheimer's disease patients were videotaped, and the tapes were assessed by central raters as a means to improve the reliability of CIBIC plus-J assessment. **Methods:** Two of eight central raters were randomly selected and independently assessed the CIBIC plus-J of each patient. **Results:** CIBIC plus-J of 41 patients was assessed. The agreement rate between the two raters was 46.3% (19/41), when two raters assessed the CIBIC plus-J of the same patient. However, when considering disagreement between adjacent points as 'agree', the agreement rate was 97.6% (40/41). Although the kappa coefficients contained coincidence, simple and quadratic weighted kappa coefficients [95% confidential interval (CI)] were 0.226 (0.066–0.386) and 0.633 (0.507–0.759), respectively, and when considering disagreement between adjacent points as 'agree', the agreement kappa was 0.896 (0.752–1.041). The interclass coefficient from the two-way layout model was 0.639. **Conclusion:** The reliability of the CIBIC plus-J assessment with the videotaped method was acceptable.

Copyright © 2012 S. Karger AG, Basel

Introduction

ADAS-J cog [1, 2] and CIBIC plus-J have been widely used in Japan as primary outcome assessment tools in clinical trials in Alzheimer's disease (AD). CIBIC plus-J is a clinical global assessment method for senile dementia patients, and is the Japanese language version equivalent to CIBIC plus [3]. The Clinician's Global Impression of Change (CGIC) of CIBIC plus-J is comprehensively assessed on a 7-point scale based on rater's impression in consideration of the results of the domains that comprise activities of daily living, psychological symptoms and cognition of the patients, which are assessed by subscales of Disability Assessment of Dementia (DAD) [4], Behavioural Pathology in Alzheimer's Disease (BEHAVE-AD) [5, 6], and Mental Function Impairment Scale (MENFIS) for cognitive and emotional impairment [7], respectively [8, 9].

Therefore, CIBIC plus-J raters must be well trained in clinical evaluation of anti-dementia drugs. A report on the reliability of CIBIC plus-J described that eleven physicians who were familiar with dementia evaluated CGIC, videotaping the interviews of 13 AD patients and 7 virtual patients at baseline and after follow-up periods of 3–14 months. As a result, the kappa coefficient showed a moderate agreement of 0.453 [10]. However, CIBIC plus-J is generally conducted at clinical sites, and we consider that greater variability would arise among raters at such sites in accordance with their experience. In addition, CIBIC plus-J depends on raters' memories, because it is assessed by comparing conditions between baseline and follow-up period. Thus, for longer study duration, the assessment shows less reliability and objectivity. Furthermore, in clinical trials with AD patients, the rate of deterioration of symptoms in patients receiving placebo slows, thus making it difficult to measure the difference in efficacy between the placebo and the active drug [11–13].

Therefore, we presumed that a more objective assessment could be made by videotaping the interview of CIBIC plus-J to avoid that raters would have to rely on their memories of patients' conditions, and that assessment by central raters who are well experienced in the clinical evaluation of dementia would show decreased inter-rater variability and increased reliability.

In a multicenter clinical trial in AD patients sponsored by Dainippon Sumitomo Pharma Co., Ltd., we videotaped CIBIC plus-J interviews with AD patients, and central raters assessed the patients according to CIBIC plus-J by watching the videotapes. Here, we present the method and results. The institutional review board at each site approved the conduct of this study prior to commencement. In addition, an ethics committee approved the conduct upon request at some sites and their institutional review boards.

Materials and Methods

Informed Consent

Informed written consent was obtained from patients and caregivers before their enrollment into this clinical trial. Informed written consent forms contained not only GCP requirements but also the videotaping of patients' and their caregivers' facial expressions and voices, the watching of videotapes for assessment by central raters, and ensuring the protection of personal information handling the videotapes.

Subjects

The patients were diagnosed according to DSM-IV-TR diagnosis criteria [14] as dementia of the Alzheimer's type and by NINCDS-ADRDA criteria [15] as possible AD. In addition, they had mild-to-moderate AD assessed by MMSE [16] (with a severity score of 12–22). Interviews at baseline were performed at 4 weeks after the confirmation of eligibility.

Interviewers

The CIBIC plus-J interviews were performed by clinicians, nurses, clinical psychologists, and psychiatric social workers who were familiar with dementia, as the central raters assessed CIBIC plus-J based on videotaping. This study allowed CIBIC plus-J interview by primary physicians, ADAS raters of the patients, or under unavoidable circumstances different interviewers between baseline and follow-up period. Interviewers agreed to their voice being recorded before each interview.

Instruments and CIBIC Plus-J Work Sheet

A work sheet and assessment manual for CIBIC plus-J were prepared based on a report by Homma et al. [17], and provided to clinical sites to standardize the inquiries of patients and their caregivers.

Video Recording

Voices of patients, caregivers, and CIBIC plus-J interviewers were recorded, and facial expressions of patients and caregivers were videotaped. Voices and visual images remained unretouched for the purpose of quality assurance.

The interviewers interviewed patients and caregivers twice, once at baseline and once after a follow-up period of 1–24 weeks. The interviews were performed in accordance with the work sheet, and videotaped by professional camera operators. The videotapes were transported between clinical sites and the locations of central raters under closely guarded conditions in respect of patients' and caregivers' personal information protection.

Raters

The central raters consisted of eight AD experts. Two raters were randomly selected and independently assessed the CIBIC plus-J score of each patient.

Assessment Procedure

First, the central raters watched the videotaped CIBIC interviews of patients and caregivers to assess the patients by subscales at baseline and after the follow-up period. Then, taking into consideration the changes in the subscale values and impressions from patients' facial expressions, the raters assessed CIBIC plus-J on a 7-point scale of 'markedly improved' (score 1), 'moderately improved' (score 2), 'minimally improved' (score 3), 'no change' (score 4), 'minimally worsened' (score 5), 'moderately worsened' (score 6), and 'markedly worsened' (score 7).

Analysis

Agreement rate, kappa coefficient [17–19], and interclass coefficient (ICC) [20] from two raters' scores were used to investigate the reliability of the CIBIC plus-J assessment. Complete agreement rate and agreement rate when considering disagreement between adjacent points as 'agree' were calculated. Although the kappa coefficients contained coincidence, quadratic weighted kappa coefficient [95% confidential interval (CI)] was calculated taking into consideration that CGIC is ordinal data. Simple kappa coefficient and kappa coefficient when considering disagreement between adjacent points as 'agree' were also calculated for reference. ICC was determined using a two-way layout model with CGIC as continuous data and with patient and rater as random effects. SAS® software was used for statistical analysis.

Table 1. Summary of baseline demographic characteristics of patients

Patients	41
Gender, male/female	9/32
Age, years	72.3 (7.9)
Duration of AD, months	4.2 (1.9)
MMSE	16.9 (3.2)
ADAS-J cog	27.6 (9.1)
DSM-IV, early/late onset	12/29

Figures in parentheses indicate SDs.

Results

Subjects

Forty-one patients were selected at 16 hospitals and their CIBIC plus-J interviews were assessed by central raters. Table 1 shows baseline demographic characteristics of the patients. Of the 41 patients, 32 (78.0%) were female; mean age (SD) was 72.3 (7.4) years; mean disease duration was 4.1 (2.0) years; the mean MMSE score was 16.6 (3.0) and the ADAS-J cog score was 27.5 (8.0).

CIBIC Plus-J Assessment

The agreement rate between the first and second raters for CGIC was 46.3% (19/41) when the two raters assessed the CIBIC plus-J scores of the same patients. However, when considering disagreement between adjacent points as 'agree', the agreement rate was 97.6% (40/41) (table 2).

Although kappa coefficients contained coincidence, simple and quadratic weighted kappa coefficients (95% CI) were 0.226 (0.066–0.386) and 0.633 (0.507–0.759), respectively, and when considering disagreement between adjacent points as 'agree', the agreement kappa was 0.896 (0.752–1.041). ICC determined by the two-way layout model was 0.639.

Discussion

The central raters assessed CIBIC plus-J scores of patients based on videotaped interviews conducted at baseline followed by follow-up. This avoided that raters had to rely on their memories of the patients' conditions at baseline when evaluating CIBIC plus-J scores. Although this was a multicenter study, the precision of CIBIC plus-J assessment was also likely to be improved, because the total number of raters was decreased to seven central raters, with two raters assessing each patient.

Simple kappa was 0.226 for complete agreement between the two raters, which is low in comparison to 0.453 of Homma's report [10]. This report is based on data from a clinical trial for efficacious assessment, and not from a trial for assessment of inter-rater reliability as primary objective. As a result, CGIC was concentrated in two domains, 'no change' (score 4) and 'minimally worsened' (score 5), and expected agreement rate was higher while simple kappa coefficient was lower. CGIC should be evaluated using quadratic weighted kappa rather than simple kappa, taking into consideration that CIBIC plus-J is a 7-point assessment [22]. Quadratic weighted kappa was 0.633 in this study. It is known that the larger the sample size, the more asymptotically quadratic weighted kappa accords with ICC [20–22]. ICC in the present study was 0.639, and was nearly identical to quadratic weighted kappa coefficient of 0.633. Kappa coefficient when considering disagreement between adjacent points as 'agree' was 0.896, showing a high agreement rate.

Table 2. Gap between two raters for CIBIC plus-J assessment in the same patients

First rater's assessment	Second rater's assessment				
	moderately improved	minimally improved	no change	minimally worsened	moderately worsened
Moderately improved		1			
Minimally improved	1	1	4		
No change		3	7	6	
Minimally worsened		1	5	8	1
Moderately worsened					3
<hr/>					
L	Gap for CIBIC	Correspondence		Gap in one level	Gap in two levels
	Number of patients	19 (46.3%)		21	1
	Total, n (%)	40 (97.6%)			Total 41

These values were judged as 'substantial' and 'almost perfect', respectively, according to the criteria of Landis et al. [18]. The latter was similar to the kappa coefficient of 0.894, which Homma et al. [10] obtained when they assessed the reliability of CIBIC plus-J. Therefore, the reliability of CIBIC plus-J assessment with this videotaped method was acceptable.

Some shortcomings with the videotaped method of CIBIC plus-J were as follows:

- (1) Some CIBIC plus-J interviewers did not ask the patients several questions that would have satisfied the central raters' needs, because the interviewers had different backgrounds than the raters. We have come to realize that it is important to train CIBIC plus-J interviewers in standardized procedures for asking patients and caregivers the CIBIC plus-J questions in the same manner as usual clinical trials.
- (2) Patients attempted to answer with more effort in comparison to usual CIBIC interviews, because the videotaped method was atypical. This would give rise to a high placebo effect in the evaluation of drug efficacy. Even if the precision of CIBIC plus-J can be improved, the control of placebo effect still remains to be solved.
- (3) This method is not appropriate for multi-national clinical trials, which are increasing in number, because raters may not understand the languages that the patients and their caregivers speak in the interviews. In fact, tremendous effort was made to understand the dialects of patients and their caregivers in this study.
- (4) Professional camera operators were employed to videotape the CIBIC plus-J interviews in this study. The sites where this method can be conducted are limited, because a sufficiently large room must be available for the videotaping, as well as sufficient time for the videotaping process. The option of clinical research coordinators using home video cameras should be considered.

A problem with the CIBIC plus-J interview method itself was that reliability of caregiver responses was doubtful. Caregivers obviously did not comprehend and/or underestimated the disease condition of their patients, as observed by the central raters who watched video interviews with caregivers followed by their patients. Since the public nursing-care insurance system started in 2000 in Japan, there has been an increasing tendency for caregivers such as family members to spend less time with their patients. This could be among the reasons why caregivers do not accurately comprehend patients' disease conditions [23]. CIBIC plus-J (and CIBIC plus) may no longer be appropriate for assessment in clinical studies under this condition. The guideline on medicinal products for the treatment of AD was revised in July 2008

in Europe, and global assessment including CIBIC plus has changed from a primary to secondary variable [24].

We assessed CIBIC plus-J using a videotaped method, and it was clear that there are various issues regarding clinical studies of AD in addition to CIBIC plus-J. However, this method is expected to be more appropriate than when local raters assess CIBIC plus-J at each site in mono-national clinical studies.

Acknowledgements

We would like to thank the following principal investigators and their staffs for patient recruiting, interviews and video-recording: T. Ohashi, Seirei Hamamatsu General Hospital; S. Hozumi, Kyowa Hospital; S. Uchiyama, Tokyo Women's Medical University Hospital; M. Yamazaki, Nippon Medical School Hospital; T. Ieda, Yokkaichi Municipal Hospital; K. Shigematsu, National Hospital Organization Minami-Kyoto Hospital; Y. Inoue, Nara Medical University Hospital; Mitsuhiro Tsujihata, Nagasaki-kita Hospital; M. Murata, National Center of Neurology and Psychiatry; T. Nakajima, National Hospital Organization Niigata Hospital; T. Kobayashi, Nakano General Hospital; M. Hamamoto, Nippon Medical School Chiba Hokusoh Hospital; S. Takahashi, Iwate Medical University Hospital; S. Katayama, National Hospital Organization Hiroshima-Nishi Medical Center; K. Nishiyama, Yuge Hospital, and S. Murata, Azumi General Hospital.

We also wish to thank the following clinical research associates who were involved in collecting data for the study: Y. Sato, R. Sakaguchi, T. Tada, and K. Okamoto, Department of Clinical Development, and the following statisticians who were involved in the statistical analysis: K. Kochi and H. Horio, Department of Data Science, Dainippon Sumitomo Pharma Co., Ltd.

Disclosure Statement

This report is based on the data of a clinical trial that was sponsored by Dainippon Sumitomo Pharma Co., Ltd. Approval was obtained from the sponsor before submitting this report. The first author was involved in this trial as the chairperson of the central raters for CIBIC plus-J.

References

- 1 Rosen WG, Mohs RC, Davis KL: A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984; 141:1356–1364.
- 2 Homma A, Hukuzawa K, Tsukada Y, Ishii T, Hasegawa K, Mohs RC: Development of Japanese version of Alzheimer's Disease Assessment Scale (ADAS) (in Japanese). *Jpn J Geriatr Psychiatry* 1992;3: 647–655.
- 3 Reisberg B, Ferris SH: CIBIC-plus Interview Guide, 1994.
- 4 Gelinas I, Gauthier L, McIntyre M, Gauthier S: Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. *Am J Occup Ther* 1999;53:471–481.
- 5 Reisberg B, Borenstein J, Salob SP, Ferris SH, Franssen E, Georgotas A: Behavioral symptoms in Alzheimer's disease: phenomenology and treatment. *J Clin Psychiatry* 1987;48(suppl):9–15.
- 6 Asada T, Homma A, Kimura M, Uno M: Study on the reliability of the Japanese version of the BEHAVE-AD (in Japanese). *Jpn J Geriatr Psychiatry* 1999;10:825–834.

- 7 Homma A, Niina R, Ishii T, Hasegawa K: Development of a new rating scale for dementia in the elderly: Mental Function Impairment Scale (MENFIS) (in Japanese). *Jpn J Geriatr Psychiatry* 1991;2:1217–1222.
- 8 Homma A, Asada T, Arai H, Isse K, Imai Y, Nishikawa T, Kobune S: Clinician's Interview-Based Impression of Change plus-Japan (CIBIC plus-J) concept and assessment manual (in Japanese). *Jpn J Geriatr Psychiatry* 1997;8:855–869.
- 9 Homma A, Asada T, Arai H, Isse K, Imai Y, Nishikawa T, Kobune S, Utsuki T, Kimura F: Clinical assessment for patients with age-associated dementia – Global and Psychometric Assessment (in Japanese). *Jpn J Geriatr Psychiatry* 1999;10:193–229.
- 10 Homma A, Nakamura Y, Kobune S, Haraguchi H, Kodani N, Takami I, Matsuoka J, Matsuda H, Kusunoki T: Reliability study on the Japanese version of the clinician's interview-based impression of change. *Dement Geriatr Cogn Disord* 2006;21:97–103.
- 11 Schneider LS, Dagerman KS, Shaikh Z, Insel P: No secular trend and high variability for ADAS-cog change among placebo groups from clinical trials: Alzheimer's Association International Conference on Alzheimer's Disease, Chicago, IL, July 26–31, 2008. *Alzheimers Dementia* 2008;4(suppl 2):T167.
- 12 Irizarry MC, Webb DJ, Bains C, Barrett SJ, Lai RY, Laroche JP, Hosford D, Maher-Edwards G, Weil JG: Predictors of placebo group decline in the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-Cog) in 24 week clinical trials of Alzheimer's disease. *J Alzheimers Dis* 2008;14:301–311.
- 13 Homma A, Nakamura Y, Saito T, Shikinami K, Ishida R: A placebo-controlled, double-blind, comparative study of Galantamine hydrobromide in patients with Alzheimer-type dementia (in Japanese). *Jpn J Geriatr Psychiatry* 2011;22:333–345.
- 14 Delirium, dementia, and amnesic and other cognitive disorders; in The Task Force on DSM-IV and other committees and work groups of the American Psychiatric Association (eds): *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. 4th ed. Washington, DC: American Psychiatric Association, 1994, pp 123–163.
- 15 McKhann G, Drachman D, Folstein M, Kazman R, Price DL, Stadlan EM: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944.
- 16 Folstein MF, Folstein SE, McHugh PR: 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
- 17 Homma A, Amari K, Ueki A, Usui M, Shigeta M, Takita M, Otani T, Kawaguchi H, Haraguchi H, Fujimoto M, Kusunoki T, Kobune S: Clinical global assessment for patient with age-associated dementia: additional comments for subscale in CIBIC plus-J and made out a supplemented worksheet (in Japanese). *Jpn J Geriatr Psychiatry* 2002;13:939–959.
- 18 Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
- 19 Fleiss JL, Cohen J: The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973;33:613–619.
- 20 Fisher RA: *Statistical Methods for Research Workers*, ed 13. Edinburgh, Oliver & Boyd, 1958.
- 21 Slim J, Wright CC: The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* 2005;85:257–268.
- 22 Kundel HL, Palansky M: Measurement of observer agreement. *Radiology* 2003;228:303–308.
- 23 Nakamura Y: Investigation of the effects of use of nursing care services on changes of status of nursing care for patients with dementia and consideration of the effect on CIBIC-plus by these changes (in Japanese). *Jpn J Geriatr Psychiatry* 2010;21:685–694.
- 24 Guideline on medicinal products for the treatment of Alzheimer's disease and other dementia. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP) 2008.