Journal of Brain Research

Kikuchi et al., J Brain Res 2019, 2:1

Research article Open Access

Concurrent Validity of EQ-5D-5L by Caregiver Proxy Rating with the ABC Dementia Scale for Alzheimer Patients

Takashi Kikuchi[1*], Kenji Wada-Isoe[2], Takahiro Mori[3], Yu Nakamura[3], Yumi Umeda-Kameyama[4] and Masahiro Akishita[4]

[1]Translational Research Informatics Center for Medical Innovation, Foundation for Biomedical Research and Innovation at Kobe, 1-5-4 Minatojima-Minamimachi, Chuo-ku Kobe, Hyogo 650-0047, Japan

[2]Division of Neurology, Department of Brain and Neurosciences, Tottori University, 36-1, Nishi-cho, Yonago, Tottori 683-8504, Japan

[3]Department of Neuropsychiatry, Kagawa University, 1750-1, Ikenobe, Miki-cho, Kita-gun, Kagawa 761-0793, Japan

[4]Department of Geriatric Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8665, Japan

[*]Corresponding author: Takashi Kikuchi, Translational Research Informatics Center for Medical Innovation, Foundation for Biomedical Research and Innovation at Kobe, 1-5-4 Minatojima-Minamimachi, Chuo-ku Kobe, Hyogo 650-0047, Japan, Phone: +81-78-303-9107; Fax: +81-78-303-9094; Email: tkikuchi@tri-kobe.org

Received date: January 24, 2019; **Accepted date:** February 08, 2019; **Published date:** February 13, 2019

Abstract

Background: It is difficult for individuals with Alzheimer's disease (AD) to measure their quality of life via EQ-5D-5L due to their symptoms, and the caregiver proxy rating is often used. Hence, it is important to understand the validity of EQ-5D-5L by proxy rating for the patients.

Objectives: To validate the EQ-5D-5L by caregiver proxy rating and confirm its concurrent validity using the ABC dementia scale (ABC-DS) for Alzheimer's patients.

Methods: We used item response category characteristics curves (IRCCC) to investigate the responsiveness of five dimensions in five items of the EQ-5D-5L over the possible range of the QOL. We also calculated the correlation coefficient between EQ-5D-5L and ABC-DS, and obtained a regression model to estimate the expected utility score using the ABC-DS score.

Results: IRCCC indicated that three questions of the EQ-5D-5L for mobility, self-care, and usual activities accurately rated the QOL; however, the other two questions for pain/discomfort and anxiety/depression did not sufficiently contribute to the rating. EQ-5D-5L utility scores correlated well with the three-dimensional distance (TDD) scores of the ABC-DS. By using a regression model to estimate the expected utility score of EQ-5D-5L by TDD, our simulation estimated that the therapeutic effect obtained by revising a drug regimen resulted in a monetary benefit of 110 US dollars on average.

Conclusions: We confirmed the concurrent validity of EQ-5D-5L by proxy rating with the ABC-DS. Although two items of EQ-5D-5L did not contribute to estimating the QOL, the other three items did, thus reflecting the severity of AD.

Keywords: Alzheimer's disease; EQ-5D-5L; Anxiety/depression; Correlation coefficient

Introduction

Current drugs for Alzheimer's disease (AD) can delay disease progress but cannot cure the disease. As general practitioners or family physicians may not be able to confidently measure the exact symptom changes, they often continue current prescriptions, which can lead to a waste of medical expenditures.

The number of patients with AD is rapidly growing in Japan, and the costs for the treatment and patient care will be a heavy social burden. The cost-benefit analysis for treatment and care is therefore important for AD. Patients' QOL must be properly measured for this purpose, and many previous studies have measured QOL using EQ-5D-3L or EQ-5D-5L for patients with dementia [1-4].

EQ-5D was developed as a self-rating scale, but many studies have questioned the validity of self-rating for Alzheimer's disease patients (ADP) because their capability of judgment and cognitive function were not stable [1]. Many studies have thus compared self-rating and proxy-rating and have reported that the agreement between self- and proxy-ratings was poor [3]. Kawano et al. reported that EQ-5D in Japanese by self-rating had relatively good reliability, but did not have sufficient validity and feasibility due to weak correlation with standard external scales because of missing items and a ceiling effect [5]. They also demonstrated that EQ-5D by proxy rating had strong reliability, validity, and feasibility and concluded that the proxy rating was the best measurement method to rate QOL in ADP.

There are problems in the diagnosis and assessment of AD by general practitioners. The accurate assessment of the disease is usually based on three parameters–activities of daily living (ADL) and behavioral and psychological symptoms (BPSD) [6]. By following the item response theory and other statistical approaches, the authors established a new assessment scale called the ABC Dementia Scale (ABC-DS); the ABC-DS has 13 questions about aspects of AD, each with a nine-point scale from least severe to most severe [7,8].

Evaluators use the ABC-DS to interview the caregiver about a recent patient's episode. The mean measurement time was ten minutes, and the interviewers did not need any training.

The primary purpose of this study is to use the item response theory to discuss the validity of EQ-5D-5L for ADP by proxy rating. Second, by using a correlation and a regression model, we investigated the concurrent validity of EQ-5D-5L with the ABC-DS; we evaluated whether EQ-5D-5L utility scores can well reflect on the severity of AD. Finally, we presented a useful application of the regression model as for EQ-5D-5L and ABC-DS; we carried out statistical simulations to estimate EQ-5D-5L utility scores when physicians revised prescriptions for the treatment drug. We then compared it with the values for a hypothetical control group assuming the physicians had not changed the prescriptions. Our research question in this simulation is whether the physicians can increase the QOL for ADP if they proactively revise the prescriptions.

Methods

The study (TRIAD1710)

We registered patients who were diagnosed with dementia according to the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition Text Revision [8], National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association [9] or National Institute on Aging-Alzheimer's Association workgroup [10], along with their caregivers who had regular contact or lived together with the subject at least three days. We did not restrict the severity of AD for enrolment in the study. The present data were a part of the TRIAD1710 study that evaluated the effect of revising prescription drugs over 12 weeks [6]. In this study, revising the prescription meant switching or adding treatment drugs. If switching was performed, the number of drugs was identical; if adding was performed, the number of drugs was changed from one to two. We used the ABC dementia scale to evaluate the effect of revised prescriptions over 12 weeks. Research assistants or physicians administered the assessment scales, and the caregivers were interviewed by identical evaluators at baseline and at 12 weeks [10,11].

The research ethics committee approved this study, and all ADP and their caregivers provided written consent to participate in the study. The clinical trial registration number of TRIAD1710 is UMIN000029610.

Dataset

We registered 104 patients in TRIAD1710 but removed two patients due to a protocol violation or a withdrawal of the research consent. Accordingly, we used data of 102 patients (Dataset A) for the evaluation of the concurrent validity of the ABC dementia scale with EQ-5D-5L. Eighty-two (Dataset B) out of 104 patients provided evaluable paired data that were measured both at baseline and 12 weeks. Therefore, we used 82 patients' data for our simulation study.

ABC dementia scale (ABC-DS)

The study group developed this scale for clinical trials requiring the accurate evaluation of sequential changes of symptoms and the severity for comparing them between treatment groups. We proposed a new algorithm known as three-dimensional distance (TDD) to enable integration of three types of scores that assess ADL, BPSD, and cognitive function of a patient [8].

The most favourable feature of TDD in clinical trials is to detect the changes in symptoms with high sensitivity. If the change in TDD was greater than 1.6, we were confident with 75% probability of some recognizable improvement of the patient's symptoms (data not shown). In this paper, we used this value as an arbitrary threshold for a significant improvement.

We verified inter-rater and intra-rater reliabilities in the previous studies [7,8]. We also confirmed the concurrent validity of the ABC-DS with Disability Assessment for Dementia [12], Neuropsychiatric Inventory Caregiver Distress Scale [13,14], Mini-Mental State Examination [15], and Clinical Dementia Rating (CDR) [16]. We could also estimate global CDR using the values of TDD [8].

ABC-DS in English, French, Chinese, and Korean language can be downloaded from the following site under the terms and conditions applied: https://eprovide.mapi-trust.org/instruments/abc-dementia-scale. TDD is a patent protected technique and needs a contract with the owner (abc_scale@tri-kobe.org).

EQ-5D-5L

The 5-level EQ-5D consists of five dimensions or questions abbreviated as Q in this paper: Q1 mobility, Q2 self-care, Q3 usual activities, Q4 pain/discomfort, and Q5 anxiety/depression. Each dimension has five ordinal levels: 1. No problems, 2. Slight problems, 3. Moderate problems, 4. Severe problems and 5. Extreme problems [17]. The scale was originally designed such that a rating could be selected by the patient that reflected his/her health state associated with the most appropriate statement in each of the five dimensions. The analysts of the scale can convert a combination of the selected 5-digit numbers to a utility score (EQ-5D-5L values) that reflects the patient's health state by using a tariff table [18].

In the TRIAD1710 study, we employed a caregiver rating and measured EQ-5D-5L at baseline only because the purpose of this study was to investigate the validity of proxy rating of EQ-5D-5L for ADP, which is the concurrent validity of the ABC-DS with EQ-5D-5L. Performing a cost-benefit analysis was not the original intention.

Statistical analysis

We used R version 3.5.1 and SAS version 9.4 for the statistical analyses.

First, we evaluated the validity of EQ-5D-5L for AD by proxy rating. If EQ-5D-5L works well, the five ordinal levels must reasonably correspond to the QOL levels of AD that is often called "ability" in the item response theory. We thus investigated the validity by item response category characteristic curves (IRCCCs) following the item response theory [19,20]. These curves show the probabilities that each level can be chosen over a possible range of QOL for ADP. The curves indicated how accurately each question could measure the QOL. We calculated the difficulty parameter or DIF (location) and discrimination parameter or DIS (steepness) of the IRCCCs. The values of DIF must exist between -4 and 4, which is a possible range of QOL. We also used item information curves (IIC) to estimate the relative amount of information each question could give about the QOL. The area under the curve indicates the relative amount of total information that the questions provide. We used the package software ltm of R for these analyses.

We calculated the Pearson product-moment correlation coefficient to measure of the linear correlation between two variables TDD and EQ-5D-5L utility score and obtained a formula for a univariate regression line L that the explanatory and the outcome variables are TDD and the utility score, respectively.

By using the results of our previous study TRIAD1402 [7,8], we simulated the values of TDD for a hypothetical group as if the patients in TRIAD1710 had not changed the prescriptions.

In TRIAD1402, we measured changes in TDD values when physicians did not alter their drug regimen for 12 weeks. We obtained four normal distributions N (mean, standard deviation) for changes in TDD: N(-0.93, 3.00), N(-1.11, 3.89), N(-1.37, 4.15), and N(-0.85, 3.37) if baseline CDR were 0/0.5, 1, 2, and 3 respectively. We can calculate the changes in TDD of the hypothetical group by constructing a mixed normal distribution F weighting the patient rate at baseline CDR in TRIAD1710, where

A mixed normal distribution is a common technique in statistics [19,20], and we briefly discussed the validity of this simulation in the Supplemental material.

Global CDR	N	Mean of Δ	SD	P-value
CDR0/0.5	11	-0.67	2.96	0.470
CDR1	36	-0.19	3.77	0.759
CDR2	21	0.24	3.88	0.782
CDR3	14	4.24	6.07	0.022

Table 1: The changes in TDD classified by the estimated global CDR at baseline TRIAD1710 study. Global CDR: the estimated class by TDD; N: the number of patients; Δ: the differences in TDD between baseline and 12th-week measurements; SD: standard deviation; P-value: the null hypothesis is H0: Δ=0

Table 1 shows the distributions for the observed changes in TDD in TRIAD1710, which were classified by the estimated global CDR at baseline. These distributions indicated the time at which the physicians revised the prescriptions. All patients' results showed a normal distribution G; N(0.61, 4.44).

We repeatedly selected 82 samples from F and G and calculated Δ or the difference in the sample means between two distributions. We carried out this process up to 10,000 times and calculated the probability for the event that Δ >1.6. We converted the mean of Δ into EQ-5D-5L utility scores by using L, and calculated the monetary benefit for revising prescriptions assuming that a Quality Adjusted Life years (QALY) deserved 49,000 US dollars.

Results

Patient characteristics

The number of males and females in Dataset A were 30 (29.4%) and 72 (70.6%), respectively; the median (range) of education in years was 12 (6-21). Of 102 patients, 41 (40.2%) patients switched their treatment drugs, while 61 (59.8%) patients included drugs at baseline and the details are shown in Table 2.

IRCCC and IIC

Figure 1 shows the IRCCCs for EQ-5D-5L by proxy rating. The gradients of IRCCCs were sufficiently steep in Q2 and Q3, and the gradient was moderate in Q1. The locations of IRCCCs for Q1, Q2, and Q3 reasonably reflected the fact that the five levels of EQ-5D-5L were an ordered category. However, the gradients for Q4 and Q5 were almost linear and flat. The item information curves show that Q1, Q2, and Q3 were informative as for the QOL of the patients, but Q4 and Q5 gave a little information about the QOL.

The numbers (rate) of patients per estimated global CDR in Dataset B were 16 (15.7%), 41 (40.2%), 27 (26.2%), and 18 (17.6%) for CDR0/0.5, CDR1, CDR2, and CDR3, respectively.

	Drug regimen		Number	%
	Original	Revised		
Switch 41 (40.2%)	ChE inhibitor A	ChE inhibitor B	24	58.5
	ChE inhibitor A and NMDA receptor antagonist	ChE inhibitor B and NMDA receptor antagonist	12	29.3
	ChE inhibitor A	NMDA receptor antagonist	4	9.8
	NMDA receptor antagonist	ChE inhibitor A	1	2.4
Addition 61 (59.8%)	ChE inhibitor A	ChE inhibitor A and NMDA receptor antagonist	47	77
	NMDA receptor antagonist	NMDA receptor antagonist and ChE inhibitor	14	23

Table 2: Revised drug regimen by two strategies, switch or addition. ChE: cholinesterase, NMDA: N-methyl-D-aspartate.

These findings in IRCCCs are described in further detail in Table 3. We summarized the difficulty and discrimination parameters and item information for each question in the table. The values of DIS for Q2 and Q3 were relatively high, 3.60 and 2.67, respectively. However, the values of the difficulty parameter for level five (DIF-5) in Q4 and Q 5 were 0.69 and 0.61, respectively and less than those of the fourth level (DIF-4). These results contradicted the fact that the fifth level must express the worst condition. Accordingly, the software caused a computational error so that it failed to show the curves for the fifth level in Q4 and Q 5 in Figure 1. Furthermore, the total information for Q4 and Q5 was 1.50 and 1.20, respectively, which was far less than that for others.

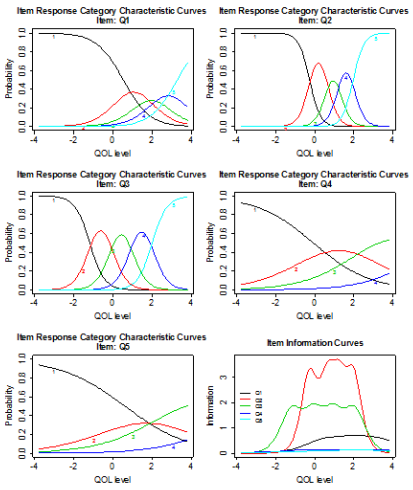


Figure 1: IRCC and IIC for EQ-5D-5L by proxy rating. The numbers in IRCCCs indicate the five level of a Likert scale on the EQ-5D-5L questionnaire. The software package ltm was not able to show the curves for the fifth level of Q4 and Q5 because of contradicting results that the fifth level curves were computationally located between first and second levels.

	DIF-1	DIF-2	DIF-3	DIF-4	DIF-5	DIS	Total-Info
Q1	0.52	1.05	1.97	2.82	3.28	1.47	3.35
Q2	-0.28	0.18	0.93	1.59	1.95	3.60	11.30
Q3	-1.19	-0.64	0.42	1.47	2.00	2.67	8.66
Q4	-0.08	1.20	4.22	3.33	0.69	-0.08	1.50
Q5	0.61	1.70	4.73	3.65	0.61	0.61	1.20

Table 3: Difficulty and location parameters of IRCCCs for EQ-5D-5L by proxy rating. DIF-1, 2, 3, 4, and 5: difficulty parameters for first, second, third, fourth, and fifth level of the Likert scale for EQ-5D-5L, DIS: discriminate parameter, and Total-Info: the area under the curve for Item Information Curves.

Concurrent validity

The correlation coefficient and 95% confidence interval (lower band, upper band) of TDD with EQ-5D-5L utility score was 0.692 (0.575, 0.781). We plotted patient data for the values of TDD and the score in Figure 2 and imposed a regression line in the graph. We write the formula of the regression line L as

EQ-5D-5L utility score=0.145+0.0112TDD. (1)

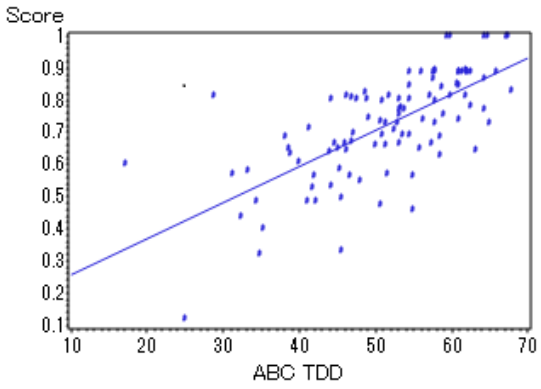


Figure 2: Correlation between TDD and EQ-5D-5L utility scores.

Simulation

To show a useful application of the equation (1), we carried out the following simulation.

By using the patient rates per the estimated global CDR at baseline in TRIAD1710, we constructed a mixed normal distribution of a hypothetical control group for Dataset B as follows:

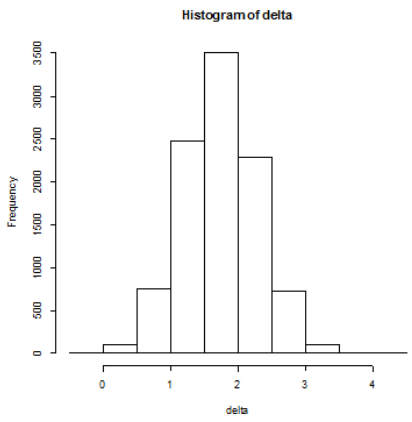
This mixed normal distribution simulated a distribution as if patients in TRIAD1710 had not changed their drug regimen.

We repeatedly selected a random sample from F and G, put them in the formula L and calculated the difference of the utility score. We showed the distribution for the difference Δ of the utility scores between F and G in Figure 3. By this simulation, we estimated the probability that the mean of the difference is larger than 1.6 as 0.60.

The mean of Δ was 1.74, and we estimated that the increments in monetary benefit by revising prescriptions was on average 0.002 QALY (=1.74×0.0112×12/520.5) or 110 US dollars per patient.

Discussion

We investigated the validity of EQ-5D-5L by proxy rating using the item response theory. We found that Q1, Q2, Q3 can be applied for the assessment of the QOL for ADP, but Q4 and Q5 were not sufficiently informative; Q4 and Q5 investigate pain/discomfort and anxiety/depression, respectively. These items did not contribute to the assessment because the patients at a progressed disease stage could experience difficulties in communicating their feelings to others and their caregivers cannot perceive the degrees of pain/discomfort and anxiety/depression [21].



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0.28	1.36	1.73	1.74	2.10	4.01

Min: minimum, 1st Qu.: first quarter, 3rd Qu.: third quarter, Max: maximum

Figure 3: Histogram for the changes in TDD score (delta) by revision of prescriptions: a simulation result.

Two of five questions of EQ-5D-5L failed to estimate the QOL, but this did not indicate that the questionnaire could not provide any information about the QOL. The utility score had a good correlation with TDD of the ABC-DS, one of the validated assessment scales for Alzheimer's disease. Accordingly, it is likely that EQ-5D-5L can provide information about the QOL reflecting the severity of AD, and this is important when health-economic evaluations are performed. We showed an example of the evaluation by using a simulation that evaluated the monetary values of the effect of revising the drug regimen. We obtained a regression model to estimate the expected values of the utility score by using the value of TDD. This result is however applicable for Japanese patients only, and a similar validation study will be needed to obtain a formula that can be used for individuals in other countries.

Conclusions

IRCCC indicated that EQ-5D-5L questions for self-care, usual activities, and mobility could contribute to the estimation of the QOL, but other questions for pain/discomfort, and anxiety/depression could not. However, because the EQ-5D-5L values correlate well with the TDD scores of the ABC-dementia scale, we concluded that the EQ-5D-5L utility scores by proxy rating could reflect the severity of AD. Accordingly, we could estimate the QOL by taking into account of the severity of AD if we calculated the expected values of the EQ-5D-5L by using the TDD score. We could thus convert estimate the monetary value based on the treatment effect by using a simple regression formula.

References

1. Orgeta V, Edwards R, Hounsome B, Orrell M, Woods B (2015) The use of the EQ-5D as a measure of health-related quality of life in people with dementia and their carers. Qual Life Res 24: 315-324.
2. Diaz-Redondo A, Rodriguez-Blazquez C, Ayala A, Martinez-Martin P, Forjaz MJ (2013) EQ-5D rated by proxy in institutionalized older adults with dementia: Psychometric pros and cons. Geriatr Gerontol Int 14: 346-353.
3. Hounsome N, Orrell M, Edwards RT (2011) EQ-5D as a quality of life measure in people with dementia and their carers: evidence and key issues. Value Health 14: 390-399.
4. Naglie G, Tomlinson G, Tansey C, Irvine J, Ritvo P, et al. (2006) Utility-based Quality of Life measures in Alzheimer's disease. Qual Life Res 15: 631-643.
5. Yasuyuki K, Toru Y, Kinoshita T, Atsuko Y, Yuriko I, et al. (2009) A study of the reliability and validity of the Japanese version of EuroQol EQ-5D and VAS in patients with Alzheimer's disease. Japanese Journal of Geriatric Psychiatry 20: 1149-1159.
6. Wada-Isoe K, Kikuchi T, Umeda-Kameyama Y (2019) ABC dementia scale can evaluate the changes of symptoms caused by alternation or addition of anti-Alzheimer's disease drugs (in Japanese). Journal of Geriatric Psychiatry 30
7. Mori T, Kikuchi T, Umeda-Kameyama Y, Wada-Isoe K, Kojima S, et al. (2018) ABC Dementia Scale: A quick assessment tool for determining Alzheimer's disease severity. Dement Geriatr Cogn Dis Extra 8: 85-97.
8. Kikuchi T, Mori T, Wada-Isoe K, Umeda-Kameyama Y, Kagimura T, et al. (2018) A novel dementia scale for Alzheimer's disease. J Alzheimers Dis Parkinsonism 8: 429.
9. American Psychiatric Association. (2000) Diagnostic and Statistical Manual of Mental Disorders, 4th ed., Text Revision. Washington DC, American Psychiatric Association.
10. McKhann G, Drachman D, Folstein M, Katzman R, Price D, et al. (1984) 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 34: 939-944.
11. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, et al. (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7: 263-269.
12. Gelinas I, Gauthier L, McIntyre M, Gauthier S (1999) Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. Am J Occup Ther 53: 471-481.
13. Kaufer, DI, Cummings, JL, Christine D, Bray T, Castellon S, et al. (1998) Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease: the neuropsychiatric inventory caregiver distress scale. J Am Geriatr Soc 46: 210-215.
14. Matsumoto N, Ikeda M, Fukuhara R, Hyodo T, Ishikawa T, et al. (2006) Validity and reliability of the Japanese version of the Neuropsychiatric Inventory Caregiver Distress Scale (NPI-D) and the Neuropsychiatric Inventory Brief Questionnaire Form (NPI-Q). Brain and Nerve 58: 785-790.
15. Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12: 189-198.
16. Morris JC (1993) The Clinical Dementia Rating (CDR): Current version and scoring rules. Neurology 43: 2412-2414.
17. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, et al. (2011) Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 20: 1727-1736.
18. Ikeda S, Shiroiwa T, Igarashi A, Noto S, Fukuda T, et al. (2015) Developing a Japanese version of the EQ-5D-5L value set. J Natl Inst Public Health 64: 47-55.
19. O'hagan A, Foster J (2004) Kendall's advanced theory of statistics-Bayesian Inference (2nd ed), Oxford University Press, New York.
20. Congdon P (2001) Bayesian statistical modeling, John Wiley & Sons, Ltd, West Sussex, England.
21. Delva F, Auriacombe S, Letenneur L, Foubert-Samier A, Bredin A (2014) Natural history of functional decline in Alzheimer's disease: a systematic review. J Alzheimers Dis 40: 57-67.