Author’s response to reviews

Title: Validation of Addenbrooke’s Cognitive Examination III for detecting mild cognitive impairment and dementia in Japan

Authors:

Shintaro Takenoshita (shintaro.take@gmail.com)
Seishi Terada (terada@okayama-u.ac.jp)
Hidenori Yoshida (hyoshida416@yahoo.co.jp)
Megumi Yamaguchi (m.shimamoto@cc.okayama-u.ac.jp)
Mayumi Yabe (m-yabe@cc.okayama-u.ac.jp)
Nao Imai (tana-o@cc.okayama-u.ac.jp)
Makiko Horiuchi (horiuc-m@okayama-u.ac.jp)
Tomoko Miki (tomokomoreaux@gmail.com)
Osamu Yokota (oyokota1@yahoo.co.jp)
Norihito Yamada (nyamada@okayama-u.ac.jp)

Version: 2  Date: 22 Mar 2019

Author’s response to reviews:

Answer to Editor

Comment 1. Thank you for including regarding ethics approval and consent to participate. As your study includes participants with mild cognitive impairment and dementia, we ask that you please provide additional details about the consent process for those participants. Please specify whether consent was obtained from the participants and/or a family member or guardian.

Response 1. Thank you for your pointing this out. We modified our manuscript as follows.

Before (Ethics approval and consent to participate)
The Internal Ethical Committee of the Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences approved this study. We obtained written informed consent after a complete explanation of the study to the subjects and their relatives.

After (Ethics approval and consent to participate)

The Internal Ethical Committee of the Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences approved this study. We gave all participants written explanations of this research, taking into consideration the cognitive impairment of the participants. After a description of the study, written informed consent was obtained from the subjects who were able to express consent. In addition, written informed consent was obtained from their relatives in all cases.

Comment 2. We notice that two authors have the same initials, please distinguish between them by designating them with numbers 1 and 2. For example: YZ1 and YZ2. YZ1 would correspond to the author furthest up on the author list.

Response 2. Thank you for your pointing this out. We modified our manuscript as you suggested.

Comment 3. We note that the current submission contains some textual overlap with other previously published works, in particular:

This overlap mainly exists in the Results section. While we understand that you may wish to express some of the same ideas contained in these publications, please be aware that we cannot condone the use of text from previously published work. We would therefore be grateful if you could reformulate this section to resolve the overlap between your manuscript and other sources.

Response 3. Thank you for your pointing this out. We modified our manuscript.

Before (Methods)

Reliability

Inter-rater reliability was measured by determining the intraclass correlation coefficient (ICC) on 25 consecutive patients. Two investigators assessed patients at the same time, and they were blind to each other’s scores. Ten participants were actively assessed by one of them while the other passively observed, and their roles were reversed for the other 15. Test-retest reliability was evaluated using 26 consecutive patients. The second session for test-retest reliability was done four to eight weeks after the first session. The test-retest reliability was determined by the ICC. The internal consistency reliability within ACE-III-J was evaluated using Cronbach's coefficient alpha [27].

After (Methods)

Reliability

Inter-rater reliability was measured by determining the intraclass correlation coefficient (ICC) of 25 consecutive patients. Two clinical psychologists assessed subjects at the same time, and they were blind to each other’s scores. One of them actively assessed ten patients while the other passively observed, and their roles were reversed for the other 15. We evaluated test-retest reliability using the ICC of 26 consecutive patients. The second session for test-retest reliability was done four to eight weeks after the first session. We evaluated the internal consistency reliability within ACE-III-J using Cronbach's coefficient alpha [27].
Statistical analyses were performed using the IBM SPSS Statistics 23.0 software program. A value of P<0.05 was accepted as significant. Two groups were compared by independent sample t-tests. Three groups were compared using one-way analysis of variance, followed by Bonferroni correction at the time of post hoc analysis. χ² tests were used for comparison of categorical data (gender). We used a multiple regression analysis to examine possible associations between the demographic variables (gender, age, and education) and the total ACE-III score.

The sensitivity and specificity of ACE-III, HDS-R, and MMSE were determined using a receiver operating characteristic (ROC) curve that plotted sensitivity and specificity across the range of possible cut-off scores [7]. We used the area under the curve (AUC) as a measure of the ability of each test to distinguish between groups of participants (dementia vs. MCI and normal; MCI vs. normal).

In this study, we used StAR software to assess statistical differences between AUCs of the three tests [28]. The optimal cut-off scores for identifying dementia and MCI were determined as the scores that led to the maximal accuracy of classification. Subsequently, positive predictive values (PPV) and negative predictive values (NPV) were estimated at different prevalence rates (5%, 10%, 20%, and 40%) for each optimal cut-off score.

Correlation between the CDR sum of box (CDR SoB) score and ACE-III scores was evaluated using Spearman’s correlation coefficient. A value of P<0.05 was accepted as significant.

After (Methods)

Statistical analysis

Statistical analyses were performed using the IBM SPSS Statistics 23.0 software program. A value of P<0.05 was accepted as significant. Two groups were compared by independent sample t-tests. Three groups were compared using one-way analysis of variance, followed by Bonferroni correction at the time of post hoc analysis. χ² tests were used for comparison of categorical data (gender). We used a multiple regression analysis to examine possible associations of the clinical characteristics (gender, age, and years of education) with the total ACE-III score.

We determined the sensitivity and specificity of ACE-III, MMSE, and HDS-R using a receiver operating characteristic (ROC) curve [7]. We used the area under the curve (AUC) as a scale of each test’s ability to differentiate between groups of participants (dementia vs. MCI and normal; MCI vs. normal).

In this study, we used StAR software to assess statistical differences between AUCs of the three tests [28]. The most suitable cut-off scores for identifying dementia and MCI were determined as the scores that led to the maximal accuracy of classification. Subsequently, positive predictive
values (PPV) and negative predictive values (NPV) were estimated at different prevalence rates (5%, 10%, 20%, and 40%) for each optimal cutoff score.

Correlation between the CDR sum of box (CDR SoB) score and ACE-III scores was evaluated using Spearman’s correlation coefficient. A value of $P<0.05$ was accepted as significant.

Before (Results)

Demographics of control, MCI, and dementia groups

Table 1 shows the clinical characteristics, MMSE scores, HDS-R scores, and ACE-III-J total and subscores of dementia, MCI, and control groups.

Age ($F(2, 386)=20.93$, $P<0.001$) and education ($F(2, 386)=7.47$, $P=0.001$) were significantly different between the three groups. Post hoc analysis showed that the dementia group was significantly older and less educated than the control and MCI groups, and that the MCI group was older than the control group. The multiple regression analysis using the total score of ACE-III-J as a dependent variable and clinical data (gender, age, and education) as independent variables revealed that age ($\beta$; standard partial regression coefficient=-0.282, $P<0.001$) and education ($\beta=0.129$, $P<0.05$) had a significant impact on the ACE-III-J score. When the same analysis was done on the control subjects (n=74), it revealed that only age ($\beta=-0.266$, $P<0.05$) affected ACE-III-J performance significantly.

ACE-III-J total ($F(2, 386)=288.562$, $P<0.001$), MMSE ($F(2, 386)=184.793$, $P<0.001$), and HDS-R ($F(2, 386)=189.996$, $P<0.001$) scores were significantly different between the three groups. On ACE-III-J, subscores of all five subdomains differed significantly among the three groups. According to the post hoc analysis, the dementia group had lower scores in all five domains than the control and MCI groups ($P<0.001$). The MCI group had lower scores than the control group in attention/orientation, memory, and fluency domains, but the differences between the two groups in language and visuospatial scores were not significant.

After (Results)

Clinical characteristics of dementia, MCI, and control groups

Table 1 shows the clinical characteristics, MMSE scores, HDS-R scores, and ACE-III-J total and subdomain scores of dementia, MCI, and control groups.

Age ($F(2, 386)=20.93$, $P<0.001$) and years of education ($F(2, 386)=7.47$, $P=0.001$) were significantly different between the three groups. The dementia group was significantly older and less educated than the control and MCI groups, and the MCI group was older than the control
group. The multiple regression analysis showed that age (β; standard partial regression coefficient=-0.282, P<0.001) and education (β=0.129, P<0.05) had a significant impact on the ACE-III-J score. When the same analysis was done on the normal controls (n=74), it revealed that only age (β=-0.266, P<0.05) affected ACE-III-J performance significantly.

ACE-III-J total (F(2, 386)=288.562, P<0.001), MMSE (F(2, 386)=184.793, P<0.001), and HDS-R (F(2, 386)=189.996, P<0.001) scores were significantly different between the three groups. On ACE-III-J, scores of all five subdomains differed significantly among the three groups. According to the post hoc analysis with Bonferroni correction, the control and MCI groups had higher scores in all five domains than the dementia group (P<0.001). The control group had higher scores than the MCI group in attention/orientation, memory, and fluency domains, but the differences between the two groups in language and visuospatial scores were not significant.

Before (Results)

Demographics of dementia group (very mild and mild)

The dementia group (n=178) was subdivided into two groups, very mild (CDR=0.5) and mild (CDR=1), according to the CDR score. The clinical characteristics are shown in Table 2.

Significant differences in education or gender distribution between the groups were not found. The very mild dementia group was significantly younger than the mild dementia group (P<0.05) and had significantly higher scores than the mild dementia group on ACE-III-J, MMSE, and HDS-R (P<0.001). On four of the subscores of the ACE-III-J, excluding the memory score, the very mild dementia group had significantly higher scores than the mild dementia group.

After (Results)

Demographics of dementia group (very mild and mild)

The dementia group (n=178) was subdivided into two groups, very mild (CDR=0.5) and mild (CDR=1), according to the CDR score. The clinical characteristics are shown in Table 2.

There were no significant differences in education or gender distribution between the groups. The mild dementia group was significantly older than the very mild dementia group (P<0.05) and had significantly lower scores than the very mild dementia group on ACE-III-J, MMSE, and HDS-R (P<0.001). On four of the subscores of the ACE-III-J, excluding the memory score, the mild dementia group had significantly lower scores than the very mild dementia group.
Normative data

Normative scores were generated for the ACE-III-J total and subdomain scores using data of the control group, based on the mean minus two standard deviations (lower limits of normal) for three age bands (≤69, 70–79, and ≥80 years old) as well as all age groups, as shown in Table 3.

Among the three age groups, the number of years of education differed (F(2, 71)=5.228, P<0.01). Post hoc analysis revealed that the ≤69 age group had more years of education than the 70–79 and ≥80 age groups (respectively, P=0.036 and 0.016).

Among these three age groups, ACE-III-J total scores (F(2, 71)=3.857, P<0.05) and visuospatial subscores (F(2, 71)=4.031, P<0.05) differed. There were no significant differences between the three groups in the attention/orientation (F(2, 71)=0.279, P=0.757), memory (F(2, 71)=1.173, P=0.315), fluency (F(2, 71)=1.900 P=0.157), and language (F(2, 71)=0.174, P=0.841) subscores. In the total ACE-III-J score, post hoc analysis disclosed no significant difference among the three groups. The scores of the ≤69 age group were significantly higher than those of the ≥80 age group (P<0.05) in the visuospatial domain.
Diagnostic interpretation

The ROC curves of ACE-III-J, HDS-R, and MMSE for diagnosing MCI or dementia are shown in Figure 1 (for MCI in Figure 1A, and for dementia in Figure 1B). The AUCs of ACE-III-J, HDS-R, and MMSE for diagnosing MCI were 0.914 (0.876–0.953), 0.859 (0.807–0.912), and 0.838 (0.780–0.896), respectively. The AUC of ACE-III-J was significantly larger than those of HDS-R and MMSE (ACE-III-J vs. HDS-R, P<0.05; ACE-III-J vs. MMSE, P<0.01). The optimal cut-off score of ACE-III-J for discriminating MCI patients from controls was 88/89 (sensitivity 0.77, specificity 0.92), and those of the HDS-R and MMSE were 24/25 (sensitivity 0.57, specificity 0.89) and 26/27 (sensitivity 0.61, specificity 0.86), respectively. The PPV and NPV of ACE-III for identifying MCI at different prevalence rates were 5% (PPV 0.35, NPV 0.99), 10% (PPV 0.52, NPV 0.97), 20% (PPV 0.72, NPV 0.94), and 40% (PPV 0.87, NPV 0.86) (Table 4).

The AUCs of ACE-III-J, HDS-R, and MMSE for diagnosing dementia were 0.938 (0.915–0.960), 0.881 (0.847–0.915), and 0.881 (0.847–0.914), respectively. The AUC of ACE-III-J was significantly larger than those of HDS-R and MMSE (ACE-III-J vs. HDS-R, P<0.001; ACE-III-J vs. MMSE, P<0.001). For discriminating dementia patients from MCI patients and controls, the optimal cut-off score of ACE-III-J was 75/76 (sensitivity 0.82 and specificity 0.90), and those of HDS-R and MMSE were 20/21 (sensitivity 0.67, specificity 0.89) and 23/24 (sensitivity 0.64, specificity 0.87), respectively. The PPV and NPV of ACE-III for identifying dementia at different prevalence rates were 5% (PPV 0.30, NPV 0.99), 10% (PPV 0.48, NPV 0.98), 20% (PPV 0.67, NPV 0.95), and 40% (PPV 0.85, NPV 0.88) (Table 4).

After (Results)

Diagnostic interpretation

The ROC curves of ACE-III-J, HDS-R, and MMSE for diagnosing MCI or dementia are shown in Figure 1 (for MCI in Figure 1A, and for dementia in Figure 1B). The AUCs of ACE-III-J, HDS-R, and MMSE for diagnosing MCI were 0.914 (0.876–0.953), 0.859 (0.807–0.912), and 0.838 (0.780–0.896), respectively. The AUC of ACE-III-J was significantly larger than those of HDS-R and MMSE (ACE-III-J vs. HDS-R, P<0.05; ACE-III-J vs. MMSE, P<0.01). The most suitable cut-off score of ACE-III-J for discriminating MCI patients from controls was 88/89 (sensitivity 0.77, specificity 0.92), and those of the HDS-R and MMSE were 24/25 (sensitivity 0.57, specificity 0.89) and 26/27 (sensitivity 0.61, specificity 0.86), respectively. The PPV and NPV of ACE-III for identifying MCI at different prevalence rates were 5% (PPV 0.35, NPV 0.99), 10% (PPV 0.52, NPV 0.97), 20% (PPV 0.72, NPV 0.94), and 40% (PPV 0.87, NPV 0.86) (Table 4).

The AUCs of ACE-III-J, HDS-R, and MMSE for diagnosing dementia were 0.938 (0.915–0.960), 0.881 (0.847–0.915), and 0.881 (0.847–0.914), respectively. The AUC of ACE-III-J was
significantly larger than those of HDS-R and MMSE (ACE-III-J vs. HDS-R, P<0.001; ACE-III-J vs. MMSE, P<0.001). For discriminating dementia patients from MCI patients and controls, the most suitable cut-off score of ACE-III-J was 75/76 (sensitivity 0.82 and specificity 0.90), and those of HDS-R and MMSE were 20/21 (sensitivity 0.67, specificity 0.89) and 23/24 (sensitivity 0.64, specificity 0.87), respectively. The PPV and NPV of ACE-III for identifying dementia at different prevalence rates were 5% (PPV 0.30, NPV 0.99), 10% (PPV 0.48, NPV 0.98), 20% (PPV 0.67, NPV 0.95), and 40% (PPV 0.85, NPV 0.88) (Table 4).

Before (Results)

Reliability

The inter-rater reliability of ACE-III-J was very good, with an ICC of 0.996. The test-retest reliability of ACE-III-J was also very good (ICC= 0.918). The internal consistency reliability for ACE-III-J was high (Cronbach’s coefficient =0.870).

After (Results)

Reliability

The inter-rater reliability of ACE-III-J was very good, with an ICC of 0.996. The test-retest reliability of ACE-III-J was also very good (ICC= 0.918). The internal consistency of ACE-III-J was high (Cronbach’s coefficient =0.870).

Before (Discussion)

Discussion

The internal consistency, test-retest reliability, and inter-rater reliability of ACE-III-J were excellent. ACE-III-J was found to be a sensitive and specific screening test to diagnose MCI and dementia in a Japanese sample, and it was better than the MMSE and HDS-R in accuracy for identifying MCI and dementia. These results suggest that ACE-III-J is a reliable and valid screening instrument.

After (Discussion)

Discussion
The reliability of ACE-III-J was excellent. ACE-III-J was found to be a sensitive and specific screening test to diagnose MCI and dementia in a Japanese sample, and it was better than the MMSE and HDS-R in accuracy for identifying MCI and dementia. These results suggest that ACE-III-J is a reliable and valid screening instrument.

Comment 4. Please include a Conclusions heading in the text of your manuscript, as both a Conclusion and a Discussion section should be provided. You may consider placing it before the last paragraph of the Discussion section.

Response 4. Thank you for your pointing this out. We modified our manuscript as follows.

Comment 5. Please add a “Declarations” heading after the list of Abbreviations.

Response 5. Thank you for your pointing this out. We added “Availability of data and materials” to our manuscript.

After (Availability of data and materials)

Availability of data and materials

The datasets of this study are available from the corresponding author on request.

Comment 6. At this stage, please upload your manuscript as a single, final, clean version that does not contain any tracked changes, comments, highlights, strikethroughs or text in different colors. All relevant tables/figures/additional files should also be clean versions. Figures (and additional files) should remain uploaded as separate files.
Response 6. I will follow your instructions.

We had our manuscript edited by a native English speaker.