Author’s response to reviews

Title: Reliability and validity of a questionnaire for self-assessment of complete dentures

Authors:

Yuriko Komagamine (y.komagamine.gerd@tmd.ac.jp)
Manabu Kanazawa (m.kanazawa.gerd@tmd.ac.jp)
Yoshinori Kaiba (rinamiryu@ac.auone.net.jp)
Yusuke Sato (y.sato.gerd@tmd.ac.jp)
Shunsuke Minakuchi (s.minakuchi.gerd@tmd.ac.jp)

Version: 2
Date: 18 December 2013

Author’s response to reviews: see over
Dear Ms. Magdalena Morawska,

We are most grateful to you and referees for the helpful comments on the original version of our manuscripts. We have taken all these comments into account and submit a revised version of our paper.

We have addressed all the comments by referee 1, 2, 3 and 4 as indicated on the attached response letters.

We hope that the explanations and revisions of our work are satisfactory, and the revision of our paper is now suitable for publication in BMC Oral Health. We are looking forward to hearing from you at your convenience.

I have a question. Please see the following. I received the comment from the Referee 4. 'Another problem is that results were not well described in appropriate section. The results of statistical analyses were rather written in Discussion section than Results section. Several sentences expressing real values should be transfer to Results section'

I described the results of statistical analyses including the sentences expressing real values in the Results section.

Should I transfer the description of the statistical analyses to the Discussion section? Please let me know your opinion.

Yours sincerely,

Yuriko Komagamine

Manabu Kanazawa
Corresponding Author
Gerodontology and Oral Rehabilitation, Graduate School of Medical and Dental Science, Tokyo Medical and Dental University
1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8549, Japan
Tel: +81 3 5803 5563  Fax: +81 3 5803 5568
E-mail: m.kanazawa.gerd@tmd.ac.jp
We are grateful to Mr Peter Robinson for the critical comments and useful suggestions, which have helped us to improve our paper considerably. As indicated in the responses that follow, we have taken all these comments and suggestions into account in the revise version of our paper. The line numbers provided with the following responses correspond to the number of lines from the top (1-24).

Responses to Comments Mar. Peter Robinson

1. The authors never define what the PDA measures. Is it a proxy for a normative assessment or is it a subjective assessment? Some of the items look like a condition specific quality of life measure, whereas others are very vague. Any subjective assessment is inevitably vague, which is why they should be tied to an explicit theoretical model.
Response
The description of what the PDA measures has been added in the forth paragraph of the revised Background (page 5, line 12 – page 6, line 4).

2. The authors' critique of OHIP and OHQoL in general are ill-founded. Measures of impact can show an improvement if the impact is diminished. This critique by the authors again suggests they have an atheoretical approach.
Response
The description of the limitation of the OHIP has been added in the third paragraph of the revised Background (page 4, line 15 – page 5, line 11).

3. Allied to the above, how can you measure expectations of a new denture after it has been fitted? This lack of intuitive sense perhaps explains the low ICC for this item.
Response
Expectations for new dentures involve patients’ prediction about their new dentures. The ICC for the subscale was calculated using the ‘Before-1 PDA’ and ‘Before-2 PDA’ values. Thus, ‘expectations of a new denture after it has been fitted’ don’t cause the low ICC for the subscale. The discussion for the low ICC for this item is described in the third paragraph of the revised Discussion (page 10, line 24 – page 11, line 9).

4. Item 21 does not seem to be related to the importance of the denture, which challenges the face validity of the measure.
Response
We agree what you mentioned.
The question item #21 is related to denture maintenance rather than importance.
Actually, from the beginning of our original plan, this subscale was named ‘importance and maintenance’. However, the subscale name seemed to be too long, so we named the subscale ‘importance’ with the thought that ‘importance’ contains the meaning of ‘maintenance’. In the previously published article (reference #18), we already described the subscale name as ‘importance’. Thus, we described ‘importance’, not ‘importance and maintenance’ in the article to keep consistency with the previously published article.

5, Tables 4 and 6 in the paper could easily be incorporated into the text
Response
I know what you say. However, we thought that Tables 4 & 6 shouldn’t be incorporated into the text because that would be easier to be understood.

In summary, I think these authors need to link up with a social scientist or health psychologist who will help them understand the need for theory in this kind of work. A measure and a manuscript with that kind of insight would be much stronger
Response
Thank you very much for your advice. In the study, we would like to demonstrate the reliability and validity of the initial PDA at first. In the future study, we would like to modify the PDA by benefits from a social scientist’s and a health psychologist’s expertise.
We are grateful to Dr. Yasuhiko Kawai, for the critical comments and useful suggestions, which have helped us to improve our paper considerably. As indicated in the responses that follow, we have taken all these comments and suggestions into account in the revise version of our paper. The line numbers provided with the following responses correspond to the number of lines from the top (1-24).

Responses to Comments from Dr. Yasuhiko Kawai

1. The purpose and clinical implication of this questionnaire should be written down in the text. Is PDA is useful for diagnosis, measuring treatment effect? Etc.
Response
The description of the purpose and clinical implication of the PDA has been added in the forth paragraph of the revised Background (page 5, line 12 – page 6, line 4).

2. These questionnaire results are influenced by past denture experience. However, you have not shown on this factor. What are the participants’ characteristics of experience to past dentures: ie number of the denture they had worn? Your discussion in the second paragraph in p11 should be discussed with those data.
Response
The Table that shows the participants’ characteristics of experience to past dentures has been added to the revised article (Table 1) and the content of the discussion in the second paragraph in p11 of the previous Discussion has been changed to the content of the fifth paragraph of the revised Discussion (page 11, lines 14 – 23).

Background:
1. In the third paragraph, you have mentioned, “some patient satisfaction evaluation methods depend on only one question [4, 5]”. But most of the VAS ratings are questionnaire is composed of several questions and asking top and bottom dentures respectively. What are the advantages of your PDA compared to those questionnaires?
Response
The description of the advantages of the PDA has been added in the forth paragraph of the revised Background (page 5, line 12 – page 6, line 4). Simply stated, the advantages of the PDA are the following. The PDA allow multidimensional measurements for patients’ consciousness and feeling of their dentures, detections of both positive and negative changes of patient’s perceptions, confirmed reliability and validity and assessment specific to dentures in themselves.
2. Please refer the VAS complete denture specific questionnaire such as Feine et. al or Kawai et.al or others.
Response
We have referred the reference whose title is ‘Do traditional techniques produce better conventional complete dentures than simplified techniques?’ The reference has added to the revised reference list (reference #9).

3. Please enter “the” before PDA. At page 5 line 18
Response
We have changed the sentences of the paragraph in the original article. So, the sentence that you pointed out was deleted. However, we have entered “the” before PDA in the revised article.

Methods and results:
1. Why “test-retest” was carried out by randomly selected 33 patients out of 122 patients? If you can write the reasons, please do so.
Response
Test-retest reliability is estimated by interclass correlation coefficients. So, we calculated the sample size for correlation analysis of the test-retest reliability. According to standard statistical criteria (α=0.05, β=0.2), 30 was calculated as a sample size. With a computer-generated random number sequence, we assigned participants and selected.

2. You have excluded the person of dementia. How did you assessed them?
Have you done any testing such as MMSE?
Response
Certainly, I agree with you.
We didn’t perform a test of diagnosis of dementia. Thus, we deleted the description ‘and showed no evidence of dementia’ from the revised article.

3. Please enter the ethical approval number.
Response
We added the approval number (approval number 232) in the first paragraph of the section titled ‘Participants’ in the revised Methods (p 6, line 22)
4. In the “Development of the PDA” section p7 line 3, “Questions items that were similar to other factors were eliminated and few new questions……”. I cannot catch “similar to other factors” please explain.
Response
The details are described in the reference #18.
Question items whose context matched more than one subscale were eliminated. For example, the question “How often does your tongue stumble for your dentures?” was eliminated after twice factor analysis with Promax rotation because this question item seemed to be related to ‘aesthetics & speech’, ‘lower denture’, and ‘function’.

5. In the clinical measurement section, the period between 1-PDA and 2-PDA could be identified as “two months” instead of “about two month”. Even if that is 62 or 59 days, it does not influence the results.
Response
We have deleted “about’ in the revised article (page 8, line 3).

6. Please explain if there is a criteria or cut off of “responsiveness”. Your result shows the value of 0.97, which is hard to decide this usefulness to the readers.
Response
The criteria of the “responsiveness” have added in the sixth paragraph of the revised Discussion (page 11, line 24 – page 12, line 9).

Discussion:
1. As told earlier, you should make clear the purpose of this questionnaire, diagnosis, prediction or effect measuring and discuss.
The description about the purpose of the PDA, diagnosis, prediction or effect measuring has been added in the forth paragraph of the revised Background (page 5, line 12 – page 6, line 4).

2. Why do you still keep “importance” even with low alpha level and prone to ceiling effect? Please make it clear.
Response
I agree with you.
However, how much patients perceive their dentures as important seems to play a crucial part in prediction of prognosis. By my supposition, high scores of the ‘importance’ subscale before treatment may carry a poor prognosis, which is helpful in
using the PDA as an instrument for prognosis. Moreover, in the study, participants were recruited from the university hospital, not respondents going to general practitioners. Extrapolation of the PDA to respondents going to general practitioners might produce different results.
We are grateful to Reviewer 3, Dr. Suguru Kimoto for the critical comments and useful suggestions, which have helped us to improve our paper considerably. As indicated in the responses that follow, we have taken all these comments and suggestions into account in the revise version of our paper. The line numbers provided with the following responses correspond to the number of lines from the top (1-24)

Responses to Comments from Dr. Suguru Kimoto

Methods
1. The authors stated that 100-mm VAS was anchored by two words, which are representing the worst situation at the left extremity of the scale and the best situation at the right extremity of the scale. However, the scores of the negative question items such as # 1, 4, 5, 7, 8, 9, 12, 14, 16, and 18 increased after new denture delivery. This means that the word anchored at left side and at right side in negative questions might represent the best situation and the worst situation respectively.
Response
The words anchored at the left side and at the right side aren’t “no” and “yes”, respectively. For all question items, the words of the left side represent the worst situations and the words of the right side represent the best situations. For example, in the case of question item #1, the word at the left is “Painful” and the word at the right is “None”. In the other case of question item #18, the left is “Always” and the right is “Never”.

2. The authors selected 33 participants from 93 participants to confirm test-retest reliability of the PDA. Please describe the method to randomly select 33 participants from 93 participants in detail.
Response
Test-retest reliability is estimated by interclass correlation coefficients. So, we calculated the sample size for correlation analysis of the test-retest reliability. According to standard statistical criteria ($\alpha=0.05, \beta=0.2$), 30 was calculated as a sample size. With a computer-generated random number sequence, we assigned participants and selected.

3. The interval between Before-1 PDA and Before-2 PDA was about two months. Please mention about the interval between the Before-2 PDA and After PDA.
Response
The interval between Before-1 PDA and After PDA was about one month.
4. The interval between the test and retest was set 2 months. To confirm the reliability of the questionnaire, the conditions of dentures and patients should be kept constantly during test-retest period. In other word, if the conditions changed at the retest by some reasons such as denture adjustment and/or pain on mucosa, participants would give their dentures different score from the first test. Please add some information how to manage dentures and patients during the 2 months in order to keep a constant condition for test-retest.
Response
The operators were instructed not to perform adjustments as possible during the test-retest. However, we couldn’t leave patients experiencing pain on mucosa. We adjusted dentures to some patients, so we couldn’t perfectly manage dentures and patients during the period to keep a constant condition.

5. The reliability of the questionnaire was assessed by the internal consistency and Cronbach’s alpha. Please explain for readers why the authors used 2 methods.
Response
Only Cronbach’s alpha is insufficient to confirm the internal consistency for the following reason; all question items should be consistent but different in one subscale. Cronbach’s alpha can estimate consistency. On the other hand, the average inter-item correlation can check whether question items overlap or not.

6. The table 2 involves the questionnaire and the means and SD. Please separate the two factors because the questionnaire is for the method section and the mean and the SD is for result section.
Response
We have separated the questionnaire and the means and the SD of the tables in the revised article. The questionnaire items are shown in the Table 3 and the means and SD are shown in the Table 4 in the revised article.

7. Generically, a table is supposed to have the horizontal line at only top and bottom and has no vertical line. I suggest that all tables are revised because tables have too many lines.
Response
We have corrected all the tables in the revised article.
We are grateful to Reviewer 4, Dr. Hisatomo Kondo for the critical comments and useful suggestions, which have helped us to improve our paper considerably. As indicated in the responses that follow, we have taken all these comments and suggestions into account in the revise version of our paper. The line numbers provided with the following responses correspond to the number of lines from the top (1-24).

Responses to Comments from Reviewer 4, Dr. Hisatomo Kondo

1. A major problem with this manuscript is little of explanation of important key words, such as PDA and number of edentulous patients, in the beginning of Background section.
   Response
   The description of the PDA has been added in the forth paragraph of the revised Background (page 5, line 12 – page 6, line 4). With regard to the number of edentulous patients, we deleted the description of future demand for complete dentures in Japan from the first paragraph of the revised article.

2. Another problem is that results were not well described in appropriate section. The results of statistical analyses were rather written in Discussion section than Results section. Several sentences expressing real values should be transfer to Results section.
   Response
   We will check whether the results of statistical analysis should be included in ‘Result section’ or ‘Discussion section’ to the editor, now. Please wait for the answer from the editor.

Specific comments
1. Abstract: According to “Survey of Dental Diseases” conducted by Dental Health Division of Ministry of Health, Labour and Welfare in Japan, the number of edentulous people have not significantly increased. Please explain clearly why authors claim demand for complete denture treatment is expected to rise over several decades in Japan.
   Response
   Certainly, I agree with you.
   As you said, there was no description about Japan in the reference #2. Therefore, we have deleted the description, ‘in Japan as well’ in the revised article.
2. Background: Since Patient Denture Assessment (PDA) seems to be not so common but considerably important in this article, authors had better explain it with additional sentences and references in the beginning of Background section.
Response
The description of the PDA has been added in the forth paragraph of the revised Background (page 5, line 12 – page 6, line 4).

3. P4, line 15: “Patient satisfaction” should be “Evaluation of patient satisfaction”
Response
We have the changed the sentences of the paragraph in the original article. So, the sentence that you pointed out was deleted.

4. Table 2 does not contain ‘Before-2’ data. Comparison between ‘Before-1’ and ‘Before-2’ on Table 2 might give us interesting information although sample number was different between those.
Response
As you say, the sample number of the ‘Before-2 PDA’ was different from the ‘Before-1 PDA’ and ‘After PDA’, Thus, we didn’t contain the data of the ‘Before-2 PDA’ in the Table 2 in the original article (Table 4 in the revised article). Both the ‘Before-1 PDA’ and ‘Before-2 PDA’ values were used for calculating the ICCs. The ICC values were based on correlation coefficients between the measures at the two different moments. Correlation analysis is not comparison between the ‘Before-1’ and ‘Before-2’, but assessment of association rather than comparison seemed to be appropriate for the analysis of the ‘Before-1 PDA’ and ‘Before-2 PDA’.

5. P12, line 23: Conclusion is not clear, and authors should ask someone expertise to correct description in English.
Response
We apologize the grammatical errors.
We have changed the original description to the description, ‘When used among edentulous patients requiring new complete dentures, the PDA demonstrated good reliability by assessing internal consistency and test-retest reliability. In addition, the PDA demonstrated good validity by assessing construct validity and discriminant validity’ in the revised Conclusions (P13, lines 4 – 7).