PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Reliability and Validity of the Japanese Version of the Ocular Surface Disease Index for Dry Eye Disease
AUTHORS	Midorikawa-Inomata, Akie; Inomata, Takenori; Nojiri, Shuko; Nakamura, Masahiro; Iwagami, Masao; Keiichi, Fujimoto; Okumura, Yuichi; Iwata, Nanami; Eguchi, Atsuko; Hasegawa, Hitomi; Kinouchi, Hikaru; Murakami, Akira; Kobayashi, Hiroyuki

VERSION 1 – REVIEW

REVIEWER	Desirée Valera-Gran Miguel Hernández University, Sapin
REVIEW RETURNED	23-Sep-2019

GENERAL COMMENTS	The paper entitled “Reliability and Validity of the Japanese Version of the Ocular Surface Disease Index for Dry Eye Disease” addressed the psychometric properties of a useful tool to detect subjective symptoms of dry eye disease (DED); an important health issue that should be under consideration. However, there are some concerns that the authors should carefully tackle before considering the manuscript for publication. 1) The authors used many acronyms across the text, although most of them were not defined before mentioning them for the first time in the text. It is recommended that the authors check all the text, paying special attention to the Materials and Methods section. 2) The “exclusion criteria” should not be treated as a separate section. This information should be included in the “Study design and participants” section. In addition, the authors should give details about inclusion criteria to be included in the study. 3) The “Dry eye disease diagnosis and classification” section refers to the process of selection of participants. This information should be included in the “Study design and participants” section. 4) A section about the description of the OSDI questionnaire should be included in the Materials and Methods. 5) All information about the instruments used for concurrent validation should be gathered together in only one section, e.g. “Other instruments for DED diagnosis and management”. 6) Both Reliability and Validity sections are part of the statistical analyses conducted in the present study. They should be included in the “Statistical analyses” section. 7) First paragraph of the Discussion should be a synthesis of the main findings obtained in the study. This is the first step before arguing any aspect about the results. However, the authors gave a general description of DED including epidemiological details, which is more appropriate for the Introduction or background. 8) Lines 329-334): the authors mentioned that the vision-related function showed modes internal consistency compared to the other 2 subscales of the J-OSDI. They explained that the internal

	consistency of this subscale is affected by the individual variation for daily activities. Why the authors did not gather information about these behaviours? This information would be of interest to perform a sensitivity analysis. The authors should discuss this issue at greater length and mentioned as limitation of the study. 9) Lines 349-353. The authors mentioned that they reported the cut-off values for the J-OSDI total score and severity categories (i.e. normal, mild, moderate, and severe). However, cut-off values for j-OSDI categories were not provided.

REVIEWER	Zeynep HAZAR KANIK Gazi University, Health Sciences Faculty, Physiotherapy and Rehabilitation, Ankara, Turkey
REVIEW RETURNED	25-Sep-2019

GENERAL COMMENTS	This study investigate "Reliability and Validity of the Japanese Version of the Ocular Surface Disease Index for Dry Eye Disease". It is well written but I have a few concern about the methodology. 1. Abstract, Result section: "Concurrent validity was assessed by Pearson correlation analysis, and the J-OSDI total score was positively associated with the Dry Eye-Related Quality-of-Life Score (γ = 0.829)" The authors should write correlation degree. e very good or good.... 2. Method section: "a forward-backward procedure was applied to translate the OSDI from English to Japanese."which guideline the author prefer? and why five translator? Normally two forward and back translators are enough. Please give detail about translation and cultural adaptation process. 3. Result section: "Test-retest reliability was evaluated in 173 participants, with a median (IQR) period of 119 (81–182) days between the test and retest." In version study we prefer 2 days to 2 weeks time interval for test-retest analysis. Marx et al. also reported that there was no statistically significant difference between the results of test-retest performed at a time between 2 days and 2 weeks. Why the author prefer this time interval? Give reference. 4. result section: line 277-284 please re-write this part. Give correlation degree good, very good.... 5.Please re-write all discussion part. It looks like result section not discussion. compare with the original version and the other version, add comment your result,

VERSION 1 – AUTHOR RESPONSE

Responses to Reviewer #1 comments:

1. The authors used many acronyms across the text, although most of them were not defined before mentioning them for the first time in the text. It is recommended that the authors check all the text, paying special attention to the Materials and Methods section.

Thank you for your suggestion. We have carefully revised the abbreviations throughout the manuscript.

2. The “exclusion criteria” should not be treated as a separate section. This information should be included in the “Study design and participants” section. In addition, the authors should give details about inclusion criteria to be included in the study.

According to the reviewer's suggestion, we have included the exclusion criteria in the Study design and participants section. In addition, we have added the inclusion criteria (lines 125–127).

3. The "Dry eye disease diagnosis and classification" section refers to the process of selection of participants. This information should be included in the "Study design and participants" section.

As requested by the reviewer, we have moved the "Dry eye disease diagnosis and classification" to the "Study design and participants" section.

4. A section about the description of the OSDI questionnaire should be included in the Materials and Methods.

According to the reviewer's suggestion, we have included the description of the OSDI questionnaire in the first paragraph of the Material and Methods.

5. All information about the instruments used for concurrent validation should be gathered together in only one section, e.g. "Other instruments for DED diagnosis and management".

Thank you for your comment. We have combined the examination in "Other instruments for DED diagnosis and management."

6. Both Reliability and Validity sections are part of the statistical analyses conducted in the present study. They should be included in the "Statistical analyses" section.

According to the reviewer's suggestion, we have moved the Reliability and Validity sections to the Statistical analyses section.

7. First paragraph of the Discussion should be a synthesis of the main findings obtained in the study. This is the first step before arguing any aspect about the results. However, the authors gave a general description of DED including epidemiological details, which is more appropriate for the Introduction or background.

We thank the reviewer for this important suggestion. We have revised the first paragraph of the Discussion.

8. Lines 329-334): the authors mentioned that the vision-related function showed modes internal consistency compared to the other 2 subscales of the J-OSDI. They explained that the internal consistency of this subscale is affected by the individual variation for daily activities. Why the authors did not gather information about these behaviours? This information would be of interest to perform a sensitivity analysis. The authors should discuss this issue at greater length and mentioned as limitation of the study.

We appreciate to give us reconsidering the reason why the lower internal consistency of vision-related function; however, it is still < 0.60 indicates satisfactory* (Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. Epidemiol Serv Saude 2017;26:649-59). As requested by the reviewer, we have elaborated on the internal consistency of the subscales as shown below.

The environmental triggers subscale showed good internal consistency and reliability, whereas the other two subscales, ocular symptoms and vision-related function, showed lower internal consistency and reliability compared to environmental triggers. Vision-related function only showed modest internal consistency. Internal consistency denotes whether all items of an instrument measure the same characteristic.[25] In the sensitivity analysis, deleting question item 7 (i.e., night driving) provided the highest ICC value of 0.74 (Supplemental Table 2). This study was conducted in central Tokyo, where the traffic network was developed, and numerous elderly people were included. Therefore, question item 7 on night driving may have affected the internal consistency. This result indicates that the question items included in OSDI need to be adjusted to the changing demands. The ocular symptoms of DED patients have typically varied because of the known fluctuations in the subjective symptoms of DED,[26, 27] thus violating this assumption of reliability.

9. Lines 349-353. The authors mentioned that they reported the cut-off values for the J-OSDI total score and severity categories (i.e. normal, mild, moderate, and severe). However, cut-off values for j-OSDI categories were not provided.

We thank the reviewer for noting this omission. We have revised Supplemental Table 3 to show the sensitivity and specificity of our reported optimal cut-off value and the sensitivity and specificity for the different severity categories: normal (0–12), mild (13–22), moderate (23–32), and severe (33–100).[11].

Responses to reviewer #2 comments:

1. Abstract, Result section: "Concurrent validity was assessed by Pearson correlation analysis, and the J-OSDI total score was positively associated with the Dry Eye-Related Quality-of-Life Score (γ = 0.829)" The authors should write correlation degree. e very good or good....

According to the reviewer's suggestion, we have revised to: "the J-OSDI total score showed a strong positive correlation with the Dry Eye-Related Quality-of-Life Score (γ = 0.829)."

2. Method section: "a forward-backward procedure was applied to translate the OSDI from English to Japanese."which guideline the author prefer? and why five translator? Normally two forward and back translators are enough. Please give detail about translation and cultural adaptation process.

We thank the reviewer for this comment. We have included all committee members. Now, we have revised the descripstion our translation and cultural adaptation process according to the previous guidelines listed below.

・Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417-32.
・Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000;25:3186-91.
・Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth 2017;11:S80-9.

3. Result section: "Test-retest reliability was evaluated in 173 participants, with a median (IQR) period of 119 (81– 182) days between the test and retest." In version study we prefer 2 days to 2 weeks time interval for test-retest analysis. Marx et al. also reported that there was no statistically significant difference between the results of test- retest performed at a time between 2 days and 2 weeks. Why the author prefer this time interval? Give reference.

We appreciate the reviewer's comments. We confirmed that our test-retest period involved variation. Therefore, we have added this information to the Discussion and Limitation sections as shown below. The ocular symptoms of DED patients have typically varied because of the known fluctuations in the subjective symptoms of DED, [26, 27], thus violating this assumption of reliability.

The test-retest method that we used to confirm reliability introduced recall bias due to the required length of the test-retest period between 2 days to 2 weeks.[29]

4. result section: line 277-284 please re-write this part. Give correlation degree good, very good....

We have revised to "the J-OSDI total score showed a significant strong positive correlation with the DEQS (γ = 0.829)."

5. Please re-write all discussion part. It looks like result section not discussion. compare with the original version and the other version, add comment your result,

We appreciated the reviewer's suggestion. We have revised and restructured the Discussion section based on your suggestions.

VERSION 2 – REVIEW

REVIEWER	Desirée Valera-Gran Miguel Hernández University, Spain
REVIEW RETURNED	11-Oct-2019

GENERAL COMMENTS	The authors have addressed properly all the issues required by the reviewer. Only one minor change is required before considering the manuscript for publication. The authors should change the symbol used for Pearson correlation coefficient. They used "γ"; however, the symbol to normally indicate Pearson correlation coefficient is "ρ".

REVIEWER	Zeynep HAZAR KANIK Gazi University
REVIEW RETURNED	21-Oct-2019

GENERAL COMMENTS	Authors made all changes that pointed out by reviewers.