Peer Review File

Article Information: https://dx.doi.org/10.21037/atm-21-6946

First Round of Peer Review

Reviewer A

The authors presented a very interesting study analyzing the performance for the early detection of the Ocular Surface Disease Index and other clinical parameters on chronic ocular GVHD. The manuscript is absolutely with merit and the findings are worth reporting to predict chronic ocular GVHD by patients by themselves and medical providers other than ophthalmologist and to provide the standard for referring the patient to the specialized ophthalmologists. However, the authors should revise the manuscript and address the following minor comments before publication:

Comment 1: Line 74-75. Generally, ophthalmic examination by ophthalmologists is better to be performed before and after HSCT. Examination before HSCT is recommended to avoid the existing dry eye disease before HSCT.

Reply 1: Thank you and we added this important point in the revision.

Changes in the text: The sentence was rephrased as "Generally, ophthalmic examination by ophthalmologists is recommended to be performed before and after HSCT. Examination before HSCT is to recognize and avoid the existing dry eye disease before HSCT. After allo-HSCT, routine ophthalmological screening evaluation should be started at 3 months, no later than 6 months, and re-screening annually" (see Page 5, line 75-78).

Comment 2: Line 79. The authors described "Qui observed". It should be written as "Qiu Y., et al observed".

Reply 2: Thank you and we revised as suggested.

Changes in the text: "Qui observed" was written as "Qiu Y., et al observed" (see Page 5, line 83).

Comment 3: Line 183. The term, “transplantation” should be replaced with “HSCT”.

Reply 3: Thank you and we revised as suggested.

Changes in the text: “Transplantation” was replaced with “HSCT” (see Page 11, line 193).

Comment 4: Line 191. Please discuss the possible reason why the proportion of HLA matched donors are larger in the co GVHD group and that of HLA unmatched donor is larger in the non coGVHD group in the discussion session.

Reply 4: Similar results have been reported by several studies. The correlation between matched donor and cGVHD has been reported in ocular GVHD (PMID: 26088932) and sclerotic skin cGVHD (PMID: 23547053). Post-transplantation cyclophosphamide is now the most widely used GVHD-prevention strategy in haploidentical transplantation, however, it is less used in patients with matched donors. (PMID: 34724567)

Changes in the text: We added that “Our results suggest an association between HLA matched donor and coGVHD. Similar results have been reported by previous studies (27, 32). Post-HSCT cyclophosphamide is now the most widely used GVHD-prevention strategy in haploidentical transplantation, however, it is less used in patients with matched donors (33). ” in the discussion (see Page 16-17, line 310-314).

Reviewer B

The authors present an interesting evaluation of the performance of the OSDI score in patients with cGVHD in comparison to the ST and BCT.

Comment 1: While the issue is clearly of interest the approach of the validation is somewhat difficult. The authors appear to compare a cohort of coGVHD patients with estalished cGVHD with a cohort of patiens earlier after transplantation without cGVHD. Therefore, the conclusion of the authors is not fully covered by the data. I would strongly recommend to provide the onset of cGVHD and the onset of ocular symptoms in relationship to the assessment. It would be more interesting to check, whether the OSDI can prognose subsequent ocular cGVHD and focus the analysis on this aspect.

Reply 1: Thank you for raising this interesting and valuable point of view that whether OSDI can be used as a predictor of coGVHD. Present cross-sectional study is not able to answer the question. Another well-designed prospective cohort study is warranted. Nonetheless, we collected and compared the available information of the patients in the study and found an association between the onset of oGVHD and the onset of ocular symptoms ($r=0.618$, $p < 0.001$). Patients with ocular discomfort tend to seek timely ophthalmic consultation and therefore their coGVHD may be diagnosed earlier.

Changes in the text: We added that "A well-designed prospective cohort study is warranted to evaluate whether OSDI can be used as a predictor of coGVHD." in the discussion (Page 15, line 280-281).

Comment 2: I would also ask the authors to check for grammar together with an native English speaker.

Reply 2: Thank you and we have checked the grammar carefully.

Changes in the text: we have checked the grammar carefully and some typos and grammar errors were corrected (see Page 6, line 91, 100; Page 7, line113, 119; Page 9, line 158; Page 11, line 188, 199; Page 12, line 222, 224; Page 15, line 270, 276).

Comment 3: Line 179 The patient are "excluded" from the analysis

Reply 3: We corrected this in the revision.

Changes in the text: We have changed the "moved" into "excluded" (see Page 11, line 189).

Comment 4: Line 183: "months"

Reply 4: We corrected the spelling in the revision.

Changes in the text: We have modified the spelling (see Page 11, line 194).

Comment 5: Line 266: poor reproducibility"

Reply 5: We corrected this in the revision.

Changes in the text: "Poor repeatability" was rephrased as "poor reproducibility" (see Page 15, line 285).

Comment 6: Please provide also NIH grading of the ocular cGVHD.

Reply 6: Thank you, we have added the NIH grading of the ocular cGVHD in the revision. The NIH eye score was significantly higher in patients with coGVHD (1.9±0.72 vs 0.17±0.38, $P <$ 0.001).

Changes in the text: We added the NIH eye score in the methods, results and table 2 (see Page 7, line 125-126; Page 12, line 215-216. Page 15, line 271; Table 2).

Reviewer C

I read with interest the paper of Jing Yang et al. on sensitive tests for coGVHD. In this study the authors assess the validity of OSDI questionnaire for primary screening of coGVHD, and they results suggest that OSDI is more sensible and specific than ST or BCI.

Although I think the idea is very interesting and could help clinicians to rapidly address transplanted patients with ocular symptoms to ophthalmologists, I think that the paper needs a major revision.

Comment 1: In fact, if we consider the group with coGVHD, the authors say that an OSDI score> than 19 is sensible and specific for the diagnosis. However, the study lack of a dry eye disease group, we do not have baseline values for both groups, and we cannot exclude that the coGVHD group had a higher OSDI score at baseline. Further, we do not know if patients had dry eye disease at baseline, or if there were more DED subjects in the coGVHD group, and so we cannot exclude a selection bias. Finally, the difference in the cutoff of DED patients (> or equal to 13) and coGVHD group (> or equal to 19) could be due to the ocular damage caused by GVHD. However, the study lacks of a dry eye group, so I think we cannot interpret the difference as due to the GVHD damage.

Reply 1 : We agree that a dry eye group will be more helpful. Dry eye is one of the most common manifestations of coGVHD and has been recognized as an important complication after HSCT(PMID: 19282026, 24305504). The higher OSDI cutoff in coGVHD (> 19.4) may indicate coGVHD associated dry eye is more symptomatic than primary dry eye without system

disease. We deleted the sentence “This higher cutoff compared to the dry eye without systemic diseases (OSDI ≥ 13) may be attributed to the stronger ocular surface damage caused by coGVHD”.

Since the patients were referred from transplant centers, we did not have ocular data before HSCT, although most patients denied history of dry eye or symptoms. An interesting study by Giannaccare G., et al showed that with or without pre-transplant ophthalmological examination had little impact on the diagnostic accuracy using ICCGVHD criteria (Inter-rate agreement coefficient Kappa was 0.933, PMID: 30194377).

Changes in the text: We have added the above points in the discussion (see Page 17, line 322-328). The sentence “This higher cutoff compared to the dry eye without systemic diseases (OSDI ≥ 13) may be attributed to the stronger ocular surface damage caused by coGVHD” was deleted in the revision (see Page 15, line 281-283).

Comment 2: I would also suggest to present data with means + SD instead of only with medians. It is interesting to see the median, but I would report also the means.

Reply 2: Thank you and data were presented as means ± SD and median (interquartile range) in the revision.

Changes in the text: We added the means ± SD in tables (see Table 1,Table 2).

Specific Comments:

Comment 3: Line 88-91: could you please add a reference for these lines? Do you have data about "a large number of transplanted patients is unable to access the ophthalmological evaluation" ?

Reply 3: A survey from Germany found that only about half of the post-HSCT patients underwent ocular examination. The author mentioned that the healthcare structure was not sufficient to treat all patients suffering from ocular GVHD.(PMID: 31075805). As mentioned in page 5, line 84-86, Qiu Y., et al also reported that delayed eye examination was commonly found in half oGVHD patients in China. (PMID: 30159166)

Changes in the text: We have added that “A survey from Germany found that only about half of the post-HSCT patients underwent ocular examination. The author mentioned that the

healthcare structure was not sufficient to treat all patients suffering from ocular GVHD (14).” in the revision (see Page 6, line 95-98).

Comment 4: Line 105-110: what about Dry Eye Disease? it can affect the OSDI score.

Reply 4: We did not have pre-HSCT ocular data. Most of patients denied history of dry eye before HSCT, however we need to admit that it would not exclude dry eye by recalling history. We have added this limitation in the discussion section. Interestingly, a prior study reported that with or without pre-transplant ophthalmological examination had little impact on the diagnostic accuracy of ICCGVHD criteria (PMID: 30194377).

Changes in the text: We have added the limitation of lacking baseline in the discussion section (see Page 17, line 322-328).

Comment 5: Line 141: I would suggest to use a different grading system for corneal staining like the NEI system (5 quadrants, stain scored 0-15 for each eye). The system that you are citing is mainly used in China, I would suggest a more reproducible grading score.

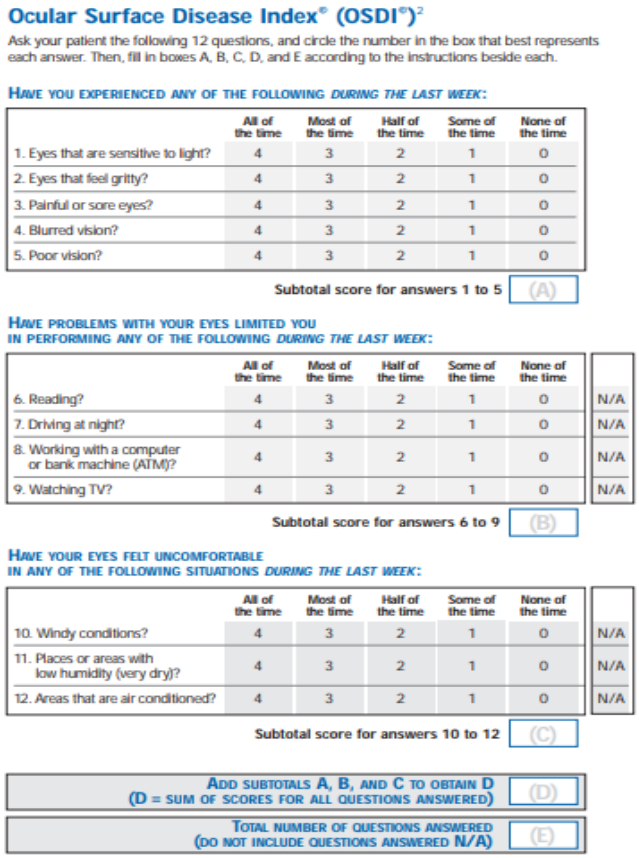
Reply 5: Thank you for your comments. We have re-evaluated the corneal staining using the NEI system based on slit-lamp photograph and prior records.

Changes in the text: The methods and results sections have been modified accordingly (see Page 8, line 147-149; Page 12, line 210; Table 2).

Comment 6: line 160: You are saying that you included the eye with higher score based on ICCGVHD severity score. However this scoring system use also OSDI, but I think you patients filled only 1 OSDI questionnaire, not 2 questionnaires (one for each eye).

Reply 6: The OSDI questionnaire describes the overall subjective symptoms without focusing on which eye specifically (please see the questionnaire attached below).

OSDI questionnaire:



Ocular Surface Disease Index® (OSDI®)[2]

Ask your patient the following 12 questions, and circle the number in the box that best represents each answer. Then, fill in boxes A, B, C, D, and E according to the instructions beside each.

HAVE YOU EXPERIENCED ANY OF THE FOLLOWING *DURING THE LAST WEEK*:

	All of the time	Most of the time	Half of the time	Some of the time	None of the time
1. Eyes that are sensitive to light?	4	3	2	1	0
2. Eyes that feel gritty?	4	3	2	1	0
3. Painful or sore eyes?	4	3	2	1	0
4. Blurred vision?	4	3	2	1	0
5. Poor vision?	4	3	2	1	0

Subtotal score for answers 1 to 5 (A)

HAVE PROBLEMS WITH YOUR EYES LIMITED YOU IN PERFORMING ANY OF THE FOLLOWING *DURING THE LAST WEEK*:

	All of the time	Most of the time	Half of the time	Some of the time	None of the time	
6. Reading?	4	3	2	1	0	N/A
7. Driving at night?	4	3	2	1	0	N/A
8. Working with a computer or bank machine (ATM)?	4	3	2	1	0	N/A
9. Watching TV?	4	3	2	1	0	N/A

Subtotal score for answers 6 to 9 (B)

HAVE YOUR EYES FELT UNCOMFORTABLE IN ANY OF THE FOLLOWING SITUATIONS *DURING THE LAST WEEK*:

	All of the time	Most of the time	Half of the time	Some of the time	None of the time	
10. Windy conditions?	4	3	2	1	0	N/A
11. Places or areas with low humidity (very dry)?	4	3	2	1	0	N/A
12. Areas that are air conditioned?	4	3	2	1	0	N/A

Subtotal score for answers 10 to 12 (C)

ADD SUBTOTALS A, B, AND C TO OBTAIN D (D = SUM OF SCORES FOR ALL QUESTIONS ANSWERED) (D)

TOTAL NUMBER OF QUESTIONS ANSWERED (DO NOT INCLUDE QUESTIONS ANSWERED N/A) (E)

Changes in the text: N/A.

Comment 7: line 182: Do you have baseline evaluation of these two groups? in terms of OSDI, corneal and conjunctival staining, ST.

Reply 7: We agree that baseline assessment is important. As mentioned above, we didn't have baseline evaluation for these patients. We have added this limitation in the discussion.

Changes in the text: We have added the limitation of lacking baseline in the discussion section (see Page 17, line 322-328).

Comment 8: line 198-202: I would suggest to present these data in the results. I know that you added a table, but I would summarize these findings in the results. And as mention above, I would use mean and sd too. Again, do you have baseline values?

Reply 8: Thank you and we summarized these data in the results, and presented these data as means ± SD.

Changes in the text: We added the suggestive data in the results (see Page 12, line 209-218).

Comment 9: line 230: statistically speaking, the AUC value of OSDI+ST was not lower than OSDI alone, you say that p is equal to 0.56, so no significant difference. The p value just say that there is not difference, you cannot say that is lower.

Reply 9: Thank you and we rephrased the sentence as suggested.

Changes in the text: The sentence was rephrased as "However, there were no significant differences in AUC value between combination of OSDI and ST, and OSDI alone (0.917 vs 0.931, *P*=0.565)." (see Page 13, line 245-247).

Comment 10: 263-264: in these two lines, you are speculating about the OSDI score results by saying that the higher OSDI scores found in your study could be related to the damage caused by coGVHD. However, you did not demonstrated this, you do not have a dry eye group to compare, and you do not know how many DED patients there are in each of your groups (with and without ocular GVHD). So, for me these part should be reviewed because it could lead to a bias.

Reply 10: We deleted the sentence "This higher cutoff compared to the dry eye without systemic diseases (OSDI ≥ 13) may be attributed to the stronger ocular surface damage caused by coGVHD". As mentioned above, we have added the limitation of lacking baseline in the discussion section.

Changes in the text: The sentence "This higher cutoff compared to the dry eye without systemic diseases (OSDI ≥ 13) may be attributed to the stronger ocular surface damage caused by coGVHD" was deleted in the revision (Page 15, line 281-283).

Comment 11: line 268-270: I do not agree with your analysis about ST. You are looking for a screening test, so you need a test with high sensitivity. You want to screen patients, and to refer those with a suspected coGVHD to the ophthalmologist. So ST test could be a good test of screening.

Reply 11: The accuracy of a diagnostic test is characterized by its sensitivity and specificity. However, the sensitivity and specificity of a test vary according to the level that is chosen as the cut-off point. The ROC curve is widely accepted as a method for selecting an optimal cut-off point

for a test and for comparing the accuracy of diagnostic tests. The AUC provides a measure of the overall performance of a diagnostic test(PMID: 29538535). In this study, the AUC for OSDI was significantly higher than ST (0.931 vs 0.826, P=0.010). Although ST has a high sensitivity (91.7%) but the specificity is low(59.7%). The OSDI shows not only high sensitivity(89.3%) but also specificity(89.6%). Moreover, ST is a contact examination and large variability. We also discussed the limitation of ST in the discussion (Page 15, line 284-287).

Changes in the text: N/A.

In conclusion, the idea related to this paper is very interesting, and it could be good to have a screening test for coGVHD for non-ophthalmologist clinicians. However, it needs significant improvement to be considered for publication.

Second Round of Peer Review

The authors improved the manuscript but some issues still remain.

Comment 1: Line 80: dry eye disease before transplant can not be "avoided" – the patients just have it.

Reply 1: Thank you and we agree that the dry eye disease before transplant cannot be "avoided". We have deleted "and avoided" in the revision.

Changes in the text: The sentence was rephrased as "Examination before HSCT is to recognize the existing dry eye disease and establish a baseline" (Page 5, line 84-85)**.**

Comment 2: Line 109: This statement belongs into the methods section

Reply 2: According to the submission checklist, we indicated that at the end of the Introduction section of Main Text. This statement is now indicated in the methods section as well.

Changes in the text: We added: "This study was reported in accordance with the STARD guidelines" in the methods section (Page 7, line 121).

Comment 3: Line 122: please use the past tense (who had active ocular….)

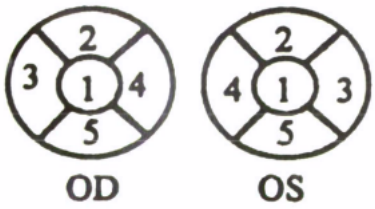
Reply 3: Thank you and we revised as suggested.

Changes in the text: The sentence was rephrased as " who had active ocular infection" (Page

7, line 126).

Comment 4: Line 152: the statement if 5 zones is somewhat unclear because directly afterwards 4 zones are deleted – please explain

Reply 4: In the original submitted version, 4 zones were used. However, 4 zones were replaced by 5 zones according to the suggestion of Reviewer #3 in the R1. Differences of CFS between the two groups were unchanged. According to the NEI system, the cornea was divided into 5 zones: central, superior, temporal, nasal, and inferior.



Changes in the text: We added the sentence "The cornea was divided into 5 zones based on the National Eye Institute grading scheme: central, superior, temporal, nasal, and inferior" to clarify the assessing method in the methods section (Page 9, line 158).

Comment 5: Line 195: "due to" post silicon

Reply 5: Thank you and we revised as suggested.

Changes in the text: "within" was rephrased as "due to" (see Page 11, line 198).

Comment 6: Line 198: you can not diagnose "without" cGVHD – please change the statement to "and 77 patients hat no cGVHD.

Reply 6: We changed the statement as suggested.

Changes in the text: We have changed **"**and 77 without" into "and 77 patients had no coGVHD" (see Page 11, line 202).