Peer Review File

Reviewer A

The authors have previously developed and reported the OSCAR rubric for trainee evaluation for various ophthalmic surgeries.
This paper is an OSCAR rubric for evaluating the surgical skill of pterygium surgery especially for resident trainees, and as disclosed in the discussion, the authors and other groups have suggested similar evaluation criteria.
Overall, the evaluation criteria presented by the authors are reasonable.
I also agree with the authors' opinion that evaluation criteria for pterygium surgical skill of resident trainee are necessary and helpful.
The manuscript is well written and I have no major concerns.

 Thank-you

Reviewer B

Minor grammatical mistakes, but otherwise well done!

Thank-you

Reviewer C

The authors describe the development of a surgical competency assessment rubric for pterygium surgery that they developed with the help of 10 experienced cornea specialists from different countries on five different continents. A very similar rubric has been previously published in 2020 in the journal Ocular Surface. However, there is value to having different versions of OSCARs for the same surgeries published, as this could stimulate further study in this area to advance the standardization of resident surgical skill evaluation. The major weakness that I see in this manuscript is the lack of scientific evaluation of the validity of this rubric. The OSCAR described herein has not been tested on trainees to get a sense of the expected distribution of scores, or ability to objectively assess improvement in skills.

Specific comments are as follows:

Pg 1: Muhammad Moin is designated as corresponding author; however, Karl Golnik is listed as the corresponding author on page 2, line 36.

Karl Golnik is listed on the manuscript as corresponding author – this correct.

Pg 2-3: Keywords do not match in list on title page and list that follows abstract.

This has been corrected.

Line 61: Language about this OSCAR being a "validated" tool should be removed. There is a difference between a tool being validated, meaning there is observed evidence showing that it works, and having "face validity" (seeming like it should work).

We have changed this language

Lines 72-74: Needs reference(s).

This statement has been changed to reflect our 25 -ear experience in residency program development and training the trainer programs around the world. We have observed this trend.

Line 90: Please briefly describe The Ophthalmology Foundation. If funding was obtained for this work, it should be noted.

Added a statement

Line 111 (and line 57): I think you mean to say that they content of the rubric was refined after expert panel review. The content and face validity of the OSCAR is what it is, and if the experts questioned the content or face validity, the rubric would have been modified to improve the content and face validity.

Made changes to reflect this

Line 112: In the abstract, you mention "multiple rounds of critique." Please specify the number of rounds of critique, and/or whether all 10 panelists came to a final consensus around the presented version of the OSCAR. I myself have multiple questions about the OSCAR itself (see Table 1 comments), and find it hard to believe that 10 experts completely agreed on the final result presented here.

Language has been modified. We do not mean that the 10 panelists came to a final consensus – the authors did.

Line 144-145: You should also assess your OSCAR for inter-rater reliability, or else the question of whether this is really a useful way to rate residents remains unanswered.

We agree and this is being done.

Line 168, 171-173: Again, avoid calling your OSCAR "validated," despite the panelists agreeing that the content included in the OSCAR is inclusive of the necessary steps/skills for pterygium surgery (content validity), and that it seems that it should be able to measure what it is intended to measure (face validity).

Language changed

Table 1:
Step 2. I don't see the difference between the description of the advanced beginner and competent. There seems to be no objective difference. Also, it is more appropriate to say "if applicable," rather than "if apply," for the optional steps.

There is a difference between the advanced beginner and competent – an advanced beginner may create an adequate incision without using appropriate technique. This would need to be corrected.
Changed apply to applicable.

Step 4: again "if applicable"

Changed to applicable

Step 8: The part about removal of 2mm strip from conjunctival edge does not seem like it would universally apply to the way different surgeons perform pterygium surgery.

This was the consensus of the authors after reviewing panelists comments.

Step 10: Mitomycin-C application (and concentration) is likely to vary between surgeons. This should be "if applicable," and the exact concentration of mitomycin-C should be left out because that is likely not standard. It would still be ok to include the part about residents needing to know the general concentrations/durations most widely used, but allow for variability based on attending surgeon preference.

Changes made
There should be instructions on how this rubric should be used, given the multiple descriptions in each level. For instance, if a resident meets criteria described in 2 different levels, which level score should apply?

The different levels are cumulative – in other words, if they meet a higher level, they would already have met a lower level. A sentence was added in the discussion 4th paragraph.

It seems that this OSCAR still needs some work to be refined and put into practice, and sending it out to a wider group, or testing it on a sample of residents, would help to increase its validity.

We agree but all the other OSCARs have been published wit this degree of work. We have found that we get feedback from users after publication and this is leading to further on-going validation studies