Author's response to reviews

Title: Standard setting: Comparison of two methods

Authors:
Sanju George Dr (sanju.george@talk21.com)
Sayeed M Haque Dr (m.s.haque@bham.ac.uk)
Femi Oyebode Prof (Femi_oyebode@msn.com)

Version: 2 Date: 17 August 2006

Author’s response to reviews:

Dear Editor,

We thank all three reviewers for their useful and constructive comments. We have now incorporated all their comments in the revised version of our manuscript. Please see below our responses to specific comments:

Response to Reviewer I

Comment: “The most notable weakness of this manuscript is the inadequate review of the literature. Many of the references -------------------------------------------------------------”

Response: In response to this, we have now done a much more comprehensive review of the relevant literature (please see updated reference list-pages 16,17,18,19 &20). As you will note, the revised reference list includes numerous recent papers in the field, and we have discussed our findings in relation to these studies. Whereas the first version of our paper had only 9 references, we now have 29 references.

Comment: “I suggest the authors delete the discussion of what they term the grading method of standard setting from their analysis ------------------------------------------”

Response: We agree with this comment and it has now been incorporated. We have now deleted the discussion of the grading method, and hence the revised version of our paper compares only 2 standard setting methods-the Angoff and the norm reference methods - also see change in title from 'standard setting: comparison of three methods' to 'Standard setting: comparison of two methods.'.

Comment: I am not sure what value the discussion of the pass rates for different norm referenced standards provides ------------------------------------------

Response: In light of this comment, we have now dropped the discussion about pass rates for different norm reference standards such as mean minus 1 SD or 1/2 SD.

Comment: ” The component of the study that in my view would be of most general interest is the reliability studies of the Angoff method ------------------------------------------”

Response: We have now carried out a comprehensive review of the literature (please see updated reference list) and other reliability studies of the Angoff method have been identified. We have also discussed our findings in relation to previous research ion the field. (see discussion section-pages 12, 13 &14).

Response to Reviewer II

We are very grateful for these positive comments.
Response to Reviewer III

Comment: "In the methods section, the first question is whether different standard setting methods result in different standards ----------------------------------"

Response: We agree with this comment in that this area has been adequately researched before, although not so in the field of undergraduate medical examination (MCQ examinations) in the UK. We felt that it was worth attempting to replicate such a study within our setting, and it was comforting to note that our findings were similar to previous research.

Comment: "One of the standard setting methods used is the grading method --------"

Response: In response to this comment and those of reviewer I we have now deleted the section discussing the grading method. Hence the revised version of our paper now discusses comparison of only two standard setting methods: the Angoff and the norm reference methods.

Comment: " A clear statistical paragraph in the methods section is missing -----------"

Response: This change has now been made - see pages 8 and 9 for the separate paragraph in the methods section on statistical analyses. We have also referenced some of the statistical methods used in our analyses (see references 19 and 20) and further emphasise that the statistical analyses were carried out by a statistician (M.S.H).

Comment: "The test-retest section is interesting -------------------------------------------"

Response: We acknowledge that the number of raters in the Angoff method of standard setting is small - see discussion section on pages 10 & 11. Although only 3 raters were used for the test-retest reliability calculations, 5-7 raters were included for the inter-rater reliability estimation.

Comment: "The conclusions described by the authors cannot be based ----------------"

Response: In light of this comment, we have made changes to the conclusion section (see page 14). We have also made considerable changes to the discussion section- in fact, we have re-written this section. We believe that now we have discussed our findings in relation to recent previous research in this field.

Comment: "The reference list should be updated."

Response: Please see updated reference list (Pages 16,17,18,19 & 20). We have now included more relevant and more recent references.

Thanking you,

Yours sincerely,

Dr Sanju George
Dr Sayeed Haque
Prof Femi Oyebode