Reviewer’s report

Title: Standard setting: Comparison of three methods

Version: 1 Date: 15 March 2006

Reviewer: David J Solomon

Reviewer’s report:

General

This study compared several standard setting approaches for a multiple choice examination given to 4th year medical students. The writing was clear and concise. The organization of the manuscript was appropriate for this type of research study.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. The most notable weakness of this manuscript is the inadequate review of the literature. Many of the references are appropriate but quite dated. There are more up-to-date pertinent references and a number of relevant studies addressing the reliability and consistency of Angoff and other standard setting procedures in medical education. I strongly suggest the authors perform a more comprehensive review of the literature and discuss their findings in relation to similar studies.

2. I suggest the authors delete the discussion of what they term the “grading method” of standard setting from their analysis. I suppose the approach is technically standard setting but it is totally arbitrary and I have never seen it discussed as a serious method of standard setting though unfortunately the authors are correct in that it is commonly used method of assigning grades. The pass rate using the grading method is tied to the difficulty of the items while both a norm-referenced and Angoff methods are sensitive to the difficulty of the items. This would tend to make the comparison very specific for their particular data set. Also, given the arbitrary nature of the “grading method” I cannot see the relevance of comparing it to a more defensible method of standard setting such as the various Angoff procedures.

3. I am not sure what value the discussion of the pass rates for different norm referenced standards provides e.g., the percentage passing at one SD or ½ SD. In a roughly normal distribution of scores the proportion of the curve one or ½ standard deviation below the mean is set and can be found in any statistical textbook. Knowing the results in the author’s sample may be of value to them but has really no relevance for anyone else reading their manuscript since it is tied to the particular shape of the distribution of scores found in their sample of students. I would drop this discussion.

4. The component of the study that in my view would be of most general interest is the reliability studies of the Angoff method. There have been a number of other studies looking at this issue though possibly not among undergraduate medical students in Great Britain. Again, I strongly suggest the authors reference some of these studies and discuss their findings in relation to the findings of other researchers.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the
author can be trusted to correct)

None

Discretionary Revisions (which the author can choose to ignore)

None

**What next?:** Reject because too small an advance to publish

**Level of interest:** An article of insufficient interest to warrant publication in a scientific/medical journal

**Quality of written English:** Acceptable

**Statistical review:** No

**Declaration of competing interests:**
I declare that I have no competing interests