Reviewer's report

Title: Practice effects in medical school entrance testing with the Undergraduate Medicine and Health Sciences Admission Test (UMAT)

Version: 3
Date: 27 January 2014

Reviewer: Jon Dowell

Reviewer's report:

>1. Is the question posed by the authors well defined?
Yes, though I read this from the perspective of an admissions lead and overall I think the article might benefit from a slightly less psychometric approach, given the likely nature of BMC readership

>2. Are the methods appropriate and well described?
Yes. I assume the authors did not have access to information about applications/offers/admissions to medicine. If so, then some additional analysis including this would be worthwhile and publishing without it would be remiss I feel as the apparent findings may look different. For instance, and something I wonder if they could do using age as a proxy, is it possible to account for those who may be taking their first test as a deliberate practice/sussing out if good enough? The fact that the age reference group is 16 and below suggests this may be occurring, as it does in the UK with UKCAT. If some were not really trying and had not seriously prepared for a first sitting it would skew the findings. I would suggest the authors at least consider this effect.

>3. Are the data sound?
Most likely.

>4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Reporting seems appropriate. No comment re deposition I could see.

For the non-statistical reader I think it essential that some form of explanation is provided by the tables - column headers are jargon to most.

>5. Are the discussion and conclusions well balanced and adequately supported by the data?
Yes, though I would suggest it is re-ordered and I would encourage the authors to orientate it rather more towards those of us struggling to make sense of these issues rather than yet more detailed (though important and interesting) analysis. After all, selection must happen, despite imperfect tools. I would suggest the paragraph re ‘A final confounding factor’ is given priority and with some more discussion re implications (should only first score be used /
higher scores expected in further sittings or do they not actually know the sitting?). Should removing younger/non-applying first sit scores make much difference then this would also come in here.

> 6. Are limitations of the work clearly stated?
Yes in psychometric terms. I believe they have omitted a significant issue as described above. If they cannot explore this then I think this is a weakness. (Therefore, according to your criteria I think this single aspect should be considered as major compulsory revision)

> 7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Yes.

> 8. Do the title and abstract accurately convey what has been found?
The title suggests they are reporting on 'Practice' whereas the paper is really about 'Repeat testing'. So it could be clearer in this regard as others have and will report regrading time and type of practice employed.

> 9. Is the writing acceptable?
Yes, very clear.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I have received reimbursements from the UKCAT as I am a Board member and Chair of the Research Panel of UKCAT.