Reviewer's report

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

Version: 3  Date: 16 December 2013

Reviewer: Don Munro

Reviewer's report:

This is a paper that will be of considerable importance and interest to all medical schools in Australia and NZ, and will have significant implications for other countries or medical schools that use similar selection tests. It is generally well written and clear. The Discussion in particular contains consideration of some very important issues, though it goes a little beyond the findings.

The following changes are suggested to further improve the paper:

1. Discretionary: The absence of commas gives the paper a somewhat 'breathless' feel when read, and in some cases leads to sentences having to be reread to catch the meaning. I think it would improve things if some commas were inserted between clauses and phrases.

2. Discretionary: In Background, p 4, a fuller description of the UMAT tests, especially Test 2, would be useful.

3. Compulsory: There needs to be a justification for the use of percentiles in the paper. Were raw or standardised scores not available? If only percentiles were available, there should be some discussion of the fact that percentiles are essentially ranked scores, applicable to their particular cohort, and therefore there is some risk that comparisons across years could be misleading, even with such large samples. There is also the issue of whether means and other parametric statistics calculated on percentiles can lead to inaccuracies, especially if they involve widely separate parts of the distribution of scores. However, it is unlikely that the results would have been affected in this instance, except in some marginal cases, so I am suggesting a comment be included rather than recalculation.

4. Compulsory: Related to #3, in several places there is reference to UMAT "scores." This should also be clarified.

5. Compulsory: pp 5-6 and Tables 1-4 or 1-7: there is no report of the Ns for the various subgroups. These are of interest, especially when comparisons between subgroups of disparate size are involved. [For the regressions, where there are overlaps between the categories, ideally the numbers involved should also be given, but to report these systematically would be cumbersome].

6. Discretionary: Tables 1-4: although the use of negative Bs is correct, positive
numbers may be clearer to ordinary readers, especially in view of the title wording "from first to second sitting."

7. Discretionary (for Editor): In view of the excellent figures and the clear statements of the results in the text, I believe all the tables could be safely relegated to an appendix.

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I was for some years in the early 2000's a member of the UMAT Technical Subcommittee for another university, and know author AM. Other than that, I have no competing interests.