Reviewer's report

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

Version: 3
Date: 15 January 2014

Reviewer: Daniel McLaughlin

Reviewer's report:

This is an interesting study of importance in the field of selection of medical students. The study refers to one specific admissions test, the Undergraduate Medicine and Health Sciences Admissions Test (UMAT), which is used for selection of medical students in Australia. However, as this test shares some common features with other tests (e.g. the UKCAT in the United Kingdom), the findings will be of relevance to a wider, international audience.

The question posed by the authors was well defined and the research methods are appropriate and well described. The findings are presented in the context of previously published literature and critically analysed in a substantial 'Discussion' section, in which the limitations of the study are clearly stated.

There are some issues with the display of data in figures and the amount of information given in figure legends. However, these should be relatively simple to correct. On one or two occasions sentence structure could be changed to improve overall comprehension. I feel that the manuscript could benefit from the following changes before publication:

Major Compulsory Revisions

1. Section on 'Predictors of Re-sitting the UMAT': On line 13 of this section, the authors state that likelihood of re-sitting the UMAT is predicted by being from a European language background, but the P value is 0.05, not P < 0.05. In fact this is the least statistically significant finding in the study, which was not borne out when they analysed data on candidates who re-sat UMAT more than once. For these reasons the authors should consider whether they wish to present this as a statistically significant finding, and amend the Abstract accordingly. Alternatively, this finding should be discussed in more detail towards the end of the manuscript to justify its inclusion.

2. Section on 'Predictors of Re-sitting the UMAT': The sentence that begins "The pattern of predictor variables..." on line 22 of this section should probably say that the pattern was "largely unchanged" as not all predictor variables remained significant (e.g. the 'UMAT-2 quartile' effects and 'School type' effects).

Minor Essential Revisions

1. Methods section, paragraph 2: The second sentence might be better phrased
as "They were further evaluated by dividing the cohort into quartiles of initial test performance and analysing the upper and lower quartiles in relation to...". A cohort can't be divided into two quartiles as the current wording suggests.

2. Methods section, paragraph 2: Needs to make reference to 'country of origin' as one of the socio-demographic indices, as this is later mentioned in the 'Statistics' section and throughout the manuscript.

3. Figures 1-3: Error bars are shown on these line graphs, but it is not indicated in the 'Methods' section or in the Figure legends whether these show standard deviation, standard, error, etc. This information should be added.

4. Figures 1-3: The data points on some line graphs do not have error bars. It should be made clear why this is the case. For example, are the errors too small to be seen alongside the mean? Whatever the reason, please state this in the figure legend.

5. Section on 'Predictors of Magnitude of Practice Effects': The sentence that begins "Separate models..." and runs from line 5 to line 10 of this section is difficult to understand because the order in which the data and the section of UMAT being referred to are not presented in a consistent manner. Consider revising to "Separate models (data not shown) indicated that for those in the lowest quartile of total UMAT score at initial testing, the predicted increment on re-testing was 18.8 percentiles (CI 18.0, 19.7) for total UMAT score...".

6. Discussion, paragraph 1: When comparing the findings in the current study with others in this area, the authors write that "For the UMAT <they> now report comparable results". Whilst the results are qualitatively similar, they are not especially quantitatively similar. I would favour a rewording here to say "For the UMAT we now report qualitatively comparable results.

Dicretionary Revisions

1. Background: For completeness and comprehension by non-Australasian readers, it would be worthwhile explaining (a) how much time a 'standard' candidate has to complete the UMAT and (b) whether UMAT is a computer-based or paper-based test.

D P McLaughlin
15th Jan 2014

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

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