Reviewer's report

Title: Predictive validity of a new integrated selection process for medical school admission

Version: 1
Date: 21 December 2013

Reviewer: Barbara Griffin

Reviewer's report:

This research adds to the body of work necessary to evaluate the selection of students into an undergraduate medical degree. Although only from one Medical School, a strength of the study is that the data come from four cohorts. Multiple criteria is also a significant strength. The title and abstract are clear. I have a number of comments regarding the paper that I hope the authors will find useful.

Major Revisions

1. Because this is framed at testing the predictive validity of a new system, then that system needs to be explained in more detail than is currently provided. How is shortlisting for interview made? How is the interview conducted? On a number of occasions the authors mention that tests were used as "filters" - this needs more detail

2. the claim the student demographics "should be considered" when interpreting predictive validity needs more explanation. Why would this be so? Rather that just controlling for these factors, perhaps they actually moderate the validity of the three selection tests?

3. The participants section was a little unclear - I was unsure why the 2003 cohort was compared to the 2002 cohort but then the 2004/5 cohorts used for the remainder of the testing. Why not combine the 2003 cohort with the other two where the new selection process was used?

4. Related to 1) above, the issue of direct and indirect range restriction needs greater attention. Most importantly, if there were "filters" used in the new process then arguably there was less range restriction for the TER than there was in the 2002 cohort, which would entirely explain the stronger correlations seen in the 2003 data. A table with means and SDs for all tests for all cohorts is essential for readers to assess range issues. Another possibility for assessing the utility of the new method would be to compare drop out rates from the two cohorts.

5. I do not agree with the author's ranking of schools where selective schools are placed as "moderate" educational advantage. They may have lower SES than private schools, but given that students from selective schools have overwhelmed the top ranking students in NSW for many years, and are even in the majority of selected students in this study, suggests that they are actually the most educationally advantaged.
6. Not sure I understand why coefficient alphas are used to examine divergent validity - this is not common practice. A factor analysis of all components would be a better option or just observation of the correlations between the tests.

7. It does not appear that the authors have corrected for unreliability of the criterion (despite referring to the need for it in the text). Given that knowledge tests are likely to have much greater reliability than assessments of clinical skills, correction would give a better comparison. Furthermore, in light of the likely significant range restriction of all predictor variables, the size of the correlations is actually quite good and if interpreted as per other selection research (see authors such as Hunter & Schmidt, Lievens, Ones), probably have more utility than the authors indicate.

8. The Discussion would benefit from reference to the need for construct matching when undertaking selection research. In other words, there is no conceptual or theoretical reason for an interview that assesses the constructs this one does to have any relationship to knowledge tests or WAM in phase 1 for instance. It might even be argued that there is less reason for cognitive ability/TER to relate to performance of clinical skills.

Minor Revisions

1. P.5 first para, "the UMAT ...was designed at UNSW" - might be misleading for some readers less familiar with the context.

2. I think TER was no longer used after 2000, when it was replaced by the UAI.

Level of interest: An article whose findings are important to those with closely related research interests

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 have received two grants from the UMAT consortium to investigate issues related to the UMAT test.