Reviewer's report

Title: Predictive validity of the UK Clinical Aptitude Test in the final years of medical school: a prospective cohort study

Version: 2  Date: 8 January 2014

Reviewer: Boaz Shulruf

Reviewer's report:

Manuscript: Predictive validity of the UK Clinical Aptitude Test in the final years of medical school: a prospective cohort study

This manuscript addresses an important issue related to medical student selection: the quality of selection tools. In particular the manuscript aims “to examine the predictive validity of the UKCAT and compare this to traditional selection methods in the senior years of medical School”.

The manuscript followed the conventional methodology that has been used in previous studies, measuring the correlations between the selection tool scores and grades in the target year, in this study the last years in the medical programme, followed by regression analysis. The findings of this study are well in line with findings presented in previous studies. Thus the question is what is new? To my understanding this study does not add significant knowledge to what we already know. However, this by itself does not disqualify such a manuscript from being published if the editors believe that it is important to present additional evidence suggesting that aptitude tests are not effective selection tools for medical programme. Overall, the data presented in this study suggest that the UKCAT scores explained 6-13% of the variance in Y4-Y5 assessments. That means that about 90% of the variance was not explained by the UKCAT. To my view, the summary above better interprets the results than what was suggested in the manuscript.

Generally the manuscript is well written, however, there are a few points which I believe need to be addressed:

1. The authors rightly argued in the introduction that “Unfortunately, measures of achievement, such as A levels and highers, while widely used in selection decisions, have been shown to be skewed by socioeconomic class”. This is an important point and a very good justification of undertaking their research, but unfortunately, the analysis did not include variables of socioeconomic class.

2. Variables for the regression model were selected if they significantly correlated with Y4-Y5 assessment scores. This process means that the decision was data-driven and not theory-driven. There is nothing wrong with that as long as it is explicitly stated. Also there is a need to list all the variables that were initially
considered for that analysis. Note that if the zero-order correlation between a predictor and the outcome is statistically significant it is very likely that the respective regression coefficient would be statistically significant as well.

3. The results include some errors. P10 “UKCAT scores explained between 6 and 22 percent of the variance in each exam alone or in combination with students’ gender”. This is incorrect. The correct sentence should read ‘UKCAT scores explained between 6 and 13 percent of the variance in each exam. The final regression models explained between 10-22% of the variance in each exam’. There is a need to make similar corrections to the sentence “Furthermore, the UKCAT was the most substantial predictor of medical school academic performance in the multiple regression models, accounting for between 6 and 15 percent of the variance alone or in combination with gender” (P11). It should read ‘Furthermore, the UKCAT was the most substantial predictor of medical school academic performance in the multiple regression models, accounting for between 6 and 13 percent of the variance”.

4. To my view, the sentence “The results from the Aberdeen OSCEs are particularly reassuring as the clinical situations representing these assessments are more similar to what is seen in clinical working practice” (P11) is an overstatement. How one can be reassured that a selection tool which explains not more than 13% of a critical outcome is effective? 13% variance explained is not negligible but nonetheless, it is relatively small. Thus, I believe that the conclusions made in this manuscript are overstated.

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

**Quality of written English:** Acceptable

**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