Reviewer's report

Title: Cross-validation of MRCGP and MRCP(UK): a database linkage study of 2,284 candidates taking both assessments

Version: 2  Date: 28 February 2014

Reviewer: John C Mclachlan

Reviewer's report:

1) This paper addresses a particular aspect of post graduate performance – effectively, the correlation of performance of those doctors who undertook both MRCP and MRCGP assessments in the assessments themselves. The analysis is relatively simple in nature, but appropriate, and generates a surprisingly large amount of interesting information! The paper is well written and makes its central points clearly. It represents a valuable, if specific, contribution to the literature. In my view it is worthy of publication. All my comments are Discretionary.

2) The title uses the term ‘cross validation’. But ‘validation’ is a complex term, with both competing definitions and sub definitions. It could be argued that this is a kind of concurrent validation, although this generally implies that the events take place closely linked in time. From another perspective, it could be argued that this study is not really a ‘validation’ of either set of tests. Perhaps it would be less complicated to title the paper simply: “A database linkage study of 2,284 candidates taking both MRCGP and MRCP(UK) assessments”. But I leave this up to the authors.

3) The Abstract mentions the ethnicity findings in the Conclusions, but not previously, although these findings are likely to be of considerable interest to a number of parties. I would include mention of the relevant methodology in the Abstract (Methods), indicate briefly the outcome in Results, and include ethnicity as a key word, for the particular benefit of readers who initially have access only to the abstract and/or computer key word searching.

4) In the Abstract (Conclusions), the authors use the phrase “cognate knowledge”. This phrase on its own is sometimes used to describe word matching in language acquisition. The terms ‘cognitive knowledge’ or ‘declarative knowledge’ are alternatives, as would be the phrase ‘knowledge cognate to both assessments’. I can see the original phrase has been used, since in the Conclusions (2) the authors use the phrase “assessing cognate areas of relevance to medicine”, where ‘cognate’ is indeed the most appropriate term.

5) The authors state “That they are in fact substantially linked supports the idea that both are assessing cognate areas of relevance to medicine. Of course that alone cannot demonstrate validity, but, as has been emphasised earlier, the demonstration of validity requires information from multiple strands. The correlations shown here certainly support the validity of both MRCGP and
MRCP(UK)”. As I read this, the last sentence contradicts the first clause of the previous sentence. This also lead to the major substantive comment I would wish to make about the paper, below.

6) As the authors indicate, the concept of ‘validity’ is complex: as I have said above, it has various definitions, each of which may have sub definitions. To my mind, the most relevant version in medical training environments is that of predictive validity, with the framing “Does performance on this test predict how effectively the candidate will perform in real clinical settings?”. Most medically related assessments are indeed intended to grant or decline candidate access to positions of greater clinical responsibility. From this view point, the article, while interesting and significant, does not advance our knowledge of predictive validity of either test (nor was it intended to, of course). If predictive validity data is subsequently identified for either set of tests, then the findings become even more valuable.

As a consequence, for my taste the authors are a little too positive about having confirmed the ‘validity’ of the tests. It remains possible that both sets of tests are measuring the same construct but that that construct is not the one we really intend to measure. I would have stated my conclusions in rather more cautious terms with regard to validity alone. However, I appreciate that my view is personal, and possibly idiosyncratic, and I propose this as a consideration for the authors (and the editor) rather than as a requirement.

**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 declare that I have no competing interests.