Reviewer's report

Title: Fitness to Practice sanctions in UK doctors are predicted by poor performance at MRCGP and MRCP(UK) assessments: data linkage study

Version: 0 Date: 17 Jul 2018

Reviewer: Matt S. Homer

Reviewer's report:

This is a very interesting paper and I would recommend publication. I have a few suggestions for improvements - mostly to do with clarity of presentation and details of methods - as I don't know the context particularly well. These are detailed below.

Abstract: The last sentence of the abstract is a bit complex - it might be better split into two.

Page 3: USMLE step 2 is 'controversial' - could you just characterise why this is the case - without the reader having to look it up?

Page 4: Does the Norcini paper show or claim a causal connection - maybe not, but it might be useful to be very explicit here either way? Also, line 5 - two lots of 'study'.

Page 5 - 'equated to PACES' - could you be a bit more explicit on what this involved (e.g. IRT or something less complex)?

Page 6 - the 2.5th percentile vs the 97.5th - this is quite interesting, and I personally haven't seen it before - so maybe a little more justification of why this is a good way to present odds ratio results? Also, the multiple imputation - what assumptions does this rely on, and are they OK - at least, in simple terms? Also, the AUC stuff is not really justified - why does this help (e.g. to allow comparison of models??)

Page 7: 'odds=1.004' - should that be 'log(odds)=1.004'? Also, do you need the 'x' at the end in the bracket? Also, the AUC statistics at the bottom of the page - what does this mean (i.e. help the less stats-savvy reader really make sense of what this is telling you)?

Bottom of page 8 - 'midway' - should that be 'between' or is it exactly 'midway'?
Page 9 - the commentary on the linearity (also in discussion) - is this having it both ways? High scoring candidates are less likely to have sanctions whilst, simultaneously, at all levels there is a linear relationship between log(odds) of sanctions and performance. Maybe it's just me, but I am not sure what the 'theoretical' implications are. Maybe one way to make this clearer is to imagine what other possible shapes the Figure 1 could show and what else this would then mean? Also, Table 4 is missing!? This is pretty important as it (apparently) shows the data for the potential confounders.

Discussion - I wonder if you could draw a little more on the literature here - there are not many references in this section (e.g. compared to the introductory sections).

Top of page 13 - 'examinations are...important evidence for the predictive validity of examinations' - is there are missing word before the first instance of 'exams'?

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

Statement on potential review bias
Please complete a statement on potential review bias, considering the following questions:

1. Did you co-author any publication with an author of this manuscript in the last 5 years?

2. Are you currently or recently affiliated at the same institution as an author of this manuscript?

If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria”. If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments
which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal