Reviewer’s report

Title: Application of machine learning models in predicting length of stay among healthcare worker in underserved communities in South Africa

Version: 0  Date: 12 Jul 2018

Reviewer: Antonio Ribas

Reviewer’s report:

Very interesting insight on trying to develop a quantitative and prediction tool for this matter. I must confess I was a bit surprised by its effectiveness, considering all the factors that have influence in the length of stay. I do have some comments, some very few and minor observations/recommendations, that will not interfere in the article approval, though.

1) In line 10 of the Introduction (p.2), you mention the loss of health care workers in South Africa, and use a reference from an article about Ghana (reference number 4). I did not read the article, thus I am not saying it is placed by mistake, but it was something I noticed, and perhaps you could doublecheck;

2) In the first three lines of page 3, you wrote: "To make matters worse there are no alternatives for the population to seek health from the private sector or next health facility as these may be hundreds of kilometers away or too expensive". There is no reference for this sentence, and I believe it is an affirmation that demands at least one. I am suggesting this because this might not be true in 100% of the cases, even though one could point in this direction;

3) In line 38 of page 4 you wrote "However, from our study we recognize that demographic factors, such as nationality or marital status, could significantly influence the length of practice of foreign practitioners". I believe to be other previous studies (not in Africa, necessarily) that found similar results. Maybe it is better to rethink how to write this sentence without leaving the impression that this is something that only your study discovered;

4) In the Statistical Analysis (p. 6, line 26) you described very clearly all the steps. However, it seemed to me that the number of entries is not such an important information to put in the Abstract, for instance. I would leave it only in the Statistical Analysis section.

Level of interest
Please indicate how interesting you found the manuscript:

An article of importance in its field

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

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