Reviewer's report

Title: Feminisation of the health workforce and wage conditions of health professions: an exploratory analysis

Version: 1 Date: 17 Apr 2019

Reviewer: Lindsay Hedden

Reviewer's report:

Thank you for the opportunity to review this article. While exploratory in nature, it offers insights into an important area of workforce development. I have some methodological concerns that I would suggest be addressed prior to accepting this article for publication.

Sex/gender: I appreciated the discussion around the distinction between sex and gender added to the introduction in response to comments made by reviewer 1; however, the nuance of this distinction isn't equally reflected in the methodology. Importantly, the authors are unclear about how gender is assessed in the survey. On line 222, the authors note that WageIndicator data is the only resource that contains sex-disaggregated data, and yet subsequently their analyses refer to gender ratios, and the gender wage gap. There are also several places where "male" and "female" should be replaced by "men" or "women". For example, in the results, line 235 and 326, the statement should read "corresponding to 64.4% participation by men" instead of "corresponding to 64.4% male participation". The authors should review the manuscript and change this throughout. Also the specific wording of the WageIndicator question is important to determine whether participants assume the survey is asking about their gender identity or biological sex. The wording of the question needs to be included explicitly in the manuscript, and the language used in the methodology and results section needs to follow that wording.

A related limitation is the implicit assumption that the trends observed are related to the concept of gender, rather than the interaction between biology (sex) and gender. Women's career choices, trajectories, and wages are affected not only by societal norms but also biological considerations such as menstruation, child bearing, breast feeding etc. The effect of these biological considerations on a woman's career choices/options are likely also variable by country. A discussion of sex effects is warranted and the inability to disentangle sex and gender effects should be cited as a limitation of this analysis, linked again to the specific metric used to ascertain sex/gender in WageIndicator.

The authors do a good of outlining the potential for selection bias in the WageIndicator data and make brief mention of the fact that previous analysis shows that these data deviate from national samples by gender, age and education level. What they do not outline, however, is how the deviation by gender in particular could potentially affect their results. For example, if women have lower levels of education in some countries and are less likely to participate (resulting in
overrepresentation of educated women), the differences they observe are likely to be underestimates.

The authors note that they were unable to conduct any formalized statistical testing because of insufficient sample size. This is a key limitation, and one that deserves a more fulsome discussion whether or not the analyses are painted as exploratory. A small sample, or one with substantial variability, does not negate the ability to perform bivariate statistical testing; it merely reduces the likelihood of finding statistical significance in any differences. Furthermore, because the results are presented solely as figures, it is impossible judge the quality of the data and therefore of the interferences that the authors are making. I would suggest that at a minimum, the authors should include a series of tables that provides, for each comparison the mean and standard deviation for each of their outcomes preferably by year and by low-, middle-, and high-income country. These tables should include appropriate statistical testing. Readers need to be able to get a clearer sense of the variability with the data and whether that changes in that variability over time could be responsible for some of the trends the authors are observing.

The authors note that respondents to the survey ranged in age from seven to 81 years. I would argue that including children in this analysis is not appropriate. They have not yet been affected by the main biological effects of sex, and their experiences as part of any workforce (including wage conditions) are likely to be driven by the fact that they are children more so than by their sex or gender.

The manuscript would be strengthened by the inclusion of a figure that tracks the inclusion/exclusion of individuals and observations. This could help to summarize the text in "country selection and grouping" and lines 268-274 in "health occupations and health worker wages", saving some words.

In addition to an expanded definition of how sex/gender is captured in WageIndicator, explicitly noting how wages were captured would be helpful. In lines 268-271, the authors note that self-reported wages are transferred to gross wages per hour. I am assuming therefore that the way in which rages are reported adjusts for whether or not individuals are working full- or part-time? Including the question wording would make this clear.

The last year of data included is from five years ago. I'm wondering why the data stops at 2014, and whether the authors have any thoughts about what would have happened to these trends in the intervening five years. Could the sample size concerns be addressed with the addition of extra years of data?
In lines 302-304, the authors note that they define gender wage gap gross hourly earnings for women as a percentage of the average for men; however, all the values are between 0 and 1, which is a proportion rather than percentage. Please correct this.

I would like to echo comments made by reviewer 2 about the importance of over/underrepresentation of data by some countries, and the inability to account for horizontal vs vertical wage gaps. Some of these concerns could be addressed using some basic statistical measures.

Given many of the comments I've made, and those made by reviewers 1 and 2 with respect to concerns with methodology and data source, the authors should provide an expanded limitations section. In its current form, it is insufficient. It does not outline the specific selection biases likely to be present in WageIndictator (including over/under-representation of some countries, gender, age and education), nor the effect on results. Nor does it reference the limitations associated with the lack of statistical rigour, limited sample size for some comparisons, inability to distinguish sex and gender effects, or vertical vs horizontal wage inequities.

On a minor note, I found this paper to be extremely lengthy. I would suggest that the Discussion (particularly lines 485-577), and the Results sections are overlong and should be edited down.

Thank you again for the opportunity to review this interesting work.

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 have no competing interests to declare.

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