Author’s response to reviews

Title: Forecasting Imbalances in the Global Health Labor Market and Devising Policy Responses

Authors:

Richard Scheffler (rscheff@berkeley.edu)

James Campbell (campbellj@who.int)

Giorgio Cometto (comettog@who.int)

Akiko Maeda (Akiko.MAEDA@oecd.org)

Jenny Liu (Jenny.Liu2@ucsf.edu)

Tim Bruckner (tim.bruckner@uci.edu)

Daniel Arnold (danielarnold@berkeley.edu)

Tim Evans (tgevans@worldbank.org)

Version: 1 Date: 07 Dec 2017

Author’s response to reviews:

We thank both reviewers for their helpful comments. The reviewers’ comments are reproduced below. Our responses follow after RESPONSE.

Reviewer reports:

Reviewer #1: The article provides a clear explanation of key differences between the different concepts and the resulting estimates. The discussion supports the Global Strategy and the recommendations of the Commission. I believe that while the article advances the debate it could be made stronger with a short paragraph on how the priority actions could be operationalized.

RESPONSE: Thank you for this suggestion. We agree that a method for operationalization should be included in the paper. We have added the following paragraph to end of our discussion in our revision.

"These priorities could be operationalized by a global strategy that involves not only the health ministers of each country, but also ministers of finance. Leadership at the president or prime minister level, especially in low-income countries, is key to success. Creating closely monitored
and agreed upon benchmarks which measure the progress of eliminating shortages by 2030 is an important way to ensure that the global workforce strategy is implemented."

Reviewer #2: This study estimates the global need, demand and supply of health workers using data from 165 countries. The figures are then used to contrast gaps between need/demand on the one hand and supply on the other hand in 2013 and 2030. The findings indicate that there are both need-based shortages and demand-based shortages in the availability of health workers. While the need-based shortages are larger in developing countries, the demand-based shortage is projected to be higher in developed countries. This finding has important policy implications. For example, it suggests that the need-based shortage in developing countries could be exacerbated if developed countries draw on health workers from developing countries to address their demand-based shortage of health workers.

The present study draws extensively on previous work conducted for a WHO Report and a paper published in this journal. The contribution of this study is to re-estimate the models using a common dataset and to contrast the results. As noted above, the findings are very interesting and have important policy implications.

However, there are certain aspects of the paper that could be improved before publication.

Major points:

As noted above, the study essentially replicates the methods used in two previous reports/papers. As I understand it, these two previous studies were used as background information for policy recommendations. Thus, it seems reasonable that the aim of this paper is not to innovate or improve on the methodology, but rather to replicate the methods and contrast the results with each other. However, a careful discussion of the strengths and limitations of the methodological approach still seems warranted, after all the credibility of the results hinges on whether the chosen methods were appropriate. In the current manuscript, you simply state that an overview of the limitations can be obtained by a "careful read" of the previous two studies. I think this is not sufficient and the current manuscript should at least provide a short discussion over the most important strengths and weaknesses of the methods used in this study.

For example, when reading the paper I wondered why the median attainment level was chosen as the benchmark for the estimates of health worker need. While I am unfamiliar with the SDG tracer indicators, an attainment of 25% does not seem to be particularly high. However, I could not find a justification for this analytical choice either in the current manuscript nor in the previous report. At the same time, this choice has of course important implications for the results - as far as I can see, setting the desired attainment level at a higher rate would mean the estimated health worker need is likely to be even higher.

RESPONSE: The SDG tracer condition attainment of 25% was selected by an advisory commission that included individuals affiliated with WHO and The World Bank, as well as various academic institutions. See the acknowledgment section of the WHO report for the complete list of advisors.
You are correct about the implications of the choice – a higher attainment level would imply higher levels of health worker need.

Another example is the estimation of health worker supply. This explicitly assumes that past trends will continue into the future. However, it is well acknowledged that in most developed countries population aging will lead to higher retirement rates in the future, which may result in skill shortages in certain sectors. Unless trends in the healthcare sector systematically deviate from the overall trend, this implies that the supply for developed countries is likely to be lower than estimated.

RESPONSE: This model was used to make a basic prediction of the supply of health workers in the future. It is true that population aging in developed countries will lead to higher retirements in the future, meaning our supply projections are likely overestimates for these countries. This now noted as endnote 1.

As stated above, I agree that it is reasonable to simply replicate the methods from the earlier reports in order to contrast the findings. Nevertheless, I think you should provide a short overview over the strengths and limitations of the methods, given that these directly relate to the strengths and limitations of their current study.

RESPONSE: The strengths and weaknesses of the model can be summarized as follows. We’ve included this complete list in the appendix of our revision.

Demand model:

**Strengths:**

Builds on previous models of HW demand (see Scheffler et al. 2008) and adds in 2 factors not previously accounted for but that can have a large affect on demand: out-of-pocket expenditures (i.e., as an indicator for the generosity of health insurance, and thus social protection against catastrophic healthcare spending) and the size of the population aged 65 or over (as an indicator for population aging, which is likely to drive greater demand for care).

Inclusion of country fixed effects to account for time-invariant unobservable heterogeneity across countries (i.e. differences in baseline characteristics between countries).

Sensitivity analyses showed that the final specification yielded the best predicted values in terms of having the lowest mean error compared to alternative specifications using differently calculated input variables (i.e., percentage of the population aged 65+, out-of-pocket expenditures as a percentage of total per capita health expenditures).

Predicted values were also relatively stable using alternative estimates of future values of GDP per capita and size of the population aged 65 or over.

**Weaknesses**
Cannot take other factors into account that also affect demand, such as changing epidemiology of disease (e.g., epidemiological transition from infectious diseases to NCDs in many lower income countries), increased productivity (through technological advance, which can also affect skills mix), changes in the organization of health care delivery

Only had sufficient data for physicians to do predictions, had to make global assumptions about nurses/midwives and other types of health care workers. Can build these other cadres in as more data become available.

Supply model:

Strengths:

Maximizes use of historical data for each country to predict future trends

Assumes a conservative linear relationship in yearly growth rather than a more aggressive exponential relationship.

Weaknesses

Assumes no change in either the entry or exit of workers into the market, which may be unrealistic (as reviewer mentions example about the aging workforce). Linear trend assumes rate of aging will stay the same when it actually may be increasing, resulting in an overestimate of future supply.

Other factors, such as changes in certification/licensing rules, migration, and education capacity could also affect the supply. In many countries, these policy and programmatic changes are enacted to augment the production of health workers, some of which may actually become employed in service delivery. Ignoring these factors would likely result in underestimating future supply.

In the end, it is not known what the net effect of these various factors might be on supply without more detailed data on graduation rates, in/out migration of health workers across countries, and retirement to include in the model.

Minor issues:

I think in some instances you should provide a bit more details for the interested readers. For example:

- In the appendix, it should be stated which GLM was used to estimate the demand regression equation.

RESPONSE: Good point. We have now included the following in the appendix.
"GLM with a normal distribution and identity link function was used to fit a linear regression using a maximum likelihood estimator. Predicted values of logged physician densities from this model were then transformed with an antilog and multiplied by a correction factor ($e^{\sigma^2/2}$) to account for the skewed distribution."

- The choice of the lags in this equation seems somewhat arbitrary. I noticed that in the previous study it was stated that these were chosen using a model selection procedure. I think it would be helpful to state this in the appendix.

RESPONSE: We agree. We have added note of the selection procedure in the appendix as follows.

"To avoid endogeneity, GDP per capita, OOP spending per capita (OOPPC), and the size of the population aged 65 or over (Pop65) were all lagged up to five years to allow time for such factors to work through the economy and affect the labor market, as other authors have done in previous projection exercises (Getzen 1990; Scheffler et al. 2008). A stepwise approach was used to select the specific combination of year lags that maximized the predictive power of each variable. Lagged variables that achieved a minimum 1% level of significance after repeated iteration were kept within the model."

- It would be helpful to provide a link to the report in Reference #1

RESPONSE: Good idea. We have included a link in the revision.