Author's response to reviews

Title: The Human Resources for Health Effort Index: A Tool to Assess and Inform Strategic Health Workforce Investments

Authors:

Alfredo Fort (alfredofort@imaworldhealth.org)

Rachel Deussom (rdeussom@fhi360.org)

Randi Burlew (RBurlew@philliberresearch.com)

Kate Gilroy (kate_gilroy@jsi.com)

David Nelson (dnelson@intrahealth.org)

Version: 1 Date: 09 Jun 2017

Author’s response to reviews:

Reviewer reports:

Reviewer #1: The Human Resources for Health Effort Index: The tool developed and tested by the authors is promising as it intends to measure the neglected HRH status and progress in low and middle income countries qualitatively. The following points may be considered for a better yield from the tool.

Overview

The score needs further analysis and refinement to become a tool for assessing and informing strategic health workforce investments. The perception of respondents has been considered leading to scores aggregating in the mid level for all dimensions. To use the scale for judging progress may be difficult after strategic interventions have been made.

RESPONSE: We note the reviewer's comment about the central tendency of responses with the HRH Index. However, this is similar to what has been found in several other indices, given their perceptive nature. For example, with the Family Planning Effort Index, made up of 30 measures embedded in 4 dimensions (Policies, Services, Evaluation, Access), total scores also tended to aggregate around the mid-point. Scores were 46% (of total) in 1989 (92 countries), 48% in 1994 (97 countries), and 49% in the latest 2009 application among 81 countries (see http://www.track20.org/pages/data/FPE). However, it is in changes over time within individual countries, as well as changes found in individual dimensions, which provides the necessary insight for policy recommendations and produces those differences among countries after interventions. In the case of the HRH Index, the most important variations are found within each
dimension (e.g., scores between 3.8 and 7 in Burkina Faso for the Finance dimension; scores between 4.5 and 7.1 in Ghana for Education & Training). This indicates the need to pay special attention to scores between and within dimensions, in each country. We have added this consideration in the Discussion section.

Objectives:

Are not mentioned separately nor clearly specified.

RESPONSE: We thank the reviewer for this input. Because of word count limitations, and in order to adhere to the sections of the manuscript, a separate "Objectives" section was not included in the manuscript. However, we have now added more clearly specified objectives at the end of the "Background" section.

Methods:

Method of selection of respondents needs to specify why the proportion of Government employees selected is higher and males predominant (more males in health sector particularly in 3 African countries?). Table 1 clearly shows the disproportionate representation of positions held by respondents. This makes inter-country comparison questionable.

RESPONSE: The reviewer notes an important point that was not sufficiently explained in the text. By necessity, the distribution in the sample favors greatly the government and the public sector. This is because several dimensions of the Index deal with aspects of national interest, such as governance, policies, national deployment of health workers, etc. Also, the gender balance of respondents often favored male officers because that reflects the actual composition found in ministries and other bodies in the countries investigated. We have added appropriate explanation to the manuscript.

It is mentioned that three formats were applied for data collection but the criteria for selection of the format to be used are not mentioned. It is important to specify this in view of the fact that scores in each dimension are averaging to midpoint, and it has been observed by researchers that if respondents are briefed well the scoring in scales is more specific. The results using manual filling of hard copies handled by a consultant (One of the methods used in the study) may have yielded better results. This information is important in view of the perception bias observed.

RESPONSE: The differing formats were only chosen to facilitate completion of the survey in varying contexts. In all cases a local trained consultant was hired, or the project assisted locally, to contact prospective respondents and to ensure instructions were provided consistently and understood by respondents. We have added this information to the manuscript. In an additional analysis of an indirect measure: attrition (average difference in respondents answering items in each dimension), we found that the average difference of respondents answering items (or not) was 2.8 overall, with the largest difference among countries being Mali (4.0), where completion was done through hard copies after individual instructions. Ghana, where 85% of forms were
responded to via SurveyMonkey, had the lowest attrition of respondents to items and dimensions, at 1.6; the largest difference in attrition among dimensions was tied between Finance and Education/Training (4.2), and not M&E, as may have been suspected. This indicates that there is variation that does not seem to depend, at least in part, on form of completion, and that manual completion, as assumed, does not necessarily yield "better results."

Results

The sampling method has not been specified. It appears that purposive sampling has been done and as a result the observation is "As expected majority of respondents were managers or directors "(Line 190). This however would not fulfill the objectives of the study.

The observation that prominence of HRH within the MOH scored highly could also be attributed to the fact that majority were managers or directors and 1/3rd respondents were MOH employees in the 3 African countries.

RESPONSE: This is an important observation. Again, as explained above, we admit sampling was purposive, given the content of the Index (e.g., national HRH policies) and the need to identify HR experts. We feel that future applications of the tool will probably tap a similar range of respondents. It is important to state that an effort to increase female participation in some countries may be tantamount to "over sampling" and may not represent the actual gender distribution. Though slowly changing, in many developing-country contexts of high level authorities in ministries, academia and even civil society, men continue to be over-represented. Of course, future applications should consider such context in their sampling strategies.

Discussion

The limitations have been accepted and discussed well. The authors have stated in their objectives that "Further by not knowing the importance of inputs, processes and outputs on HRH interventions, their effect on health outcomes is not known" (Line 102). This implies that some measurable output indicators should be mentioned and measured at baseline, and the changes in these indicators after interventions would enable measurement of the progress made. No such indicators have been mentioned.

RESPONSE: We have added examples of output indicators.

An implied objective (Line 119 -121) was that" though good management of HR role in ultimate delivery of health care services is undeniable, the status and inputs in countries that influence other dimensions such as leadership, finance or policy is less known". This objective does not seem to have been fulfilled by the scale as the leadership dimension scored well but was not tied to low scoring area of financial resources. This has been accepted by the authors (lines 365-373). In order to use the tool for assessment of HRH and measure progress after intervention this point may need to be sorted out by checking on selection criteria for respondents to assess bias. Respondent bias in answering any one dimension may lead to erroneous overall score.
RESPONSE: The reviewer is correct at stating that respondents' bias may lead to incorrect findings and conclusions. This might especially be the case if a respondent feels that the Index will judge their individual performance or results. Given that this phenomenon will be difficult to avoid, two strategies will be needed: first, to ensure instructions confirm the confidentiality and neutrality of the questionnaire (i.e., not intended to find fault, but to assess current status); second, that sampling strategies ensure the widest and most varied representation of respondents. We have added text to the manuscript accordingly.

Conclusions

The recommendations are based on the responses to the scale only and do not consider the outputs expected by following these.

Recommendation 1 (line 387-391) is a general statement and cannot be concluded from the results of the scale responses. A comprehensive approach to strengthening health workforce cannot be inferred from the results as all dimensions are scoring about midline and overall scores are not conclusive. The items however are scored differently so a brainstorming session including various stakeholders would be needed to formulate any meaningful intervention.

Recommendation 4 (line 401-404) "Countries should invest in developing greater decision making capacity among HRH leaders " cannot be recommended as the scores captured in the leadership dimensions are high and indicate no deficiency necessitating such intervention.

RESPONSE: This is correct, and we have toned down the first part of the recommendation, however keeping the latter part, which advocates for safe environments, rewarding good performance and providing career progression opportunities.

It is mentioned in the results that the Monitoring and evaluation dimension was responded to by a small number of respondents which was attributed to its specialized nature by the authors. This dimension clearly needs a reanalysis to see whether it has universal applicability or should be restricted to specified respondents only.

RESPONSE: Thanks to the reviewer for pointing this out, which made us look at the data again. Although in total there were fewer respondents addressing this M&E dimension, the average "attrition" number was actually lower (4.0) than the average experienced with the Finance (4.2) and Education/Training (4.2) dimensions. This may be for different reasons, from specialization to the difficulty of the questions, but this tells us that--again--the way to get sufficient responses to all dimensions is to have a highly diverse and sufficient sample of respondents for each application. We have modified that section to tone down the statement about "fewest respondents" having completed it.

Reviewer #2: The topic is interesting and useful for health service development. The key questions/comments to this papers are as follow:
Research question: In the background section, the authors pointed out clearly the importance of human resource for health, the limitations of the existing/conventional quantitative methods of assessing the effect of HRH on health services, and reasons why there was a need to develop a tool to measure HRH status and progress in low- and middle-income countries. However it was not clear what was the research question of this paper - whether the authors want to demonstrate the validity of the HRH Effort Index that was developed or want to describe the HRH in the four selected countries using the HRH Effort Index that was developed, etc. Was the research question answered?

RESPONSE: The reviewer is right at pointing the need to be more specific about the objectives of the study. In that sense the research question is more the former than the latter, i.e., to test the validity of the Index under different contexts/countries. Obviously, in doing so, we applied the Index in different "real-life" country scenarios, which also provided some variations, which were explained in the manuscript. We have added more wording on the objectives of the study to the Background section.

Methods: More details of the process to develop and assess validity the HRH Effort Index are needed.

RESPONSE: Although we understand the reviewer’s desire for more details about the process, this is challenging given word count limitations. We have added some more details in the sampling, Results, and Discussion areas, which we hope better explains some issues with the Index.

Validity of the findings: In the discussion section, the HRH Effort Index was claimed to provide new and standardized information on the inputs and processes in the area of HRH, gain insight into critical areas of HRH. However, the sample in each country was quite small and the authors was also aware of the limitation of the subjective nature of the responses. How valid of the policy recommendation based on these findings?

RESPONSE: We are aware of the relatively small sample size of each application (i.e., from 16-27). However, we feel that, taken altogether, the 108 total sample size provides some confidence with the initial findings. However, because of this and other limitations mentioned by the reviewer, we have added to the discussion of the limitations of this study, plus the recommendation at the end that there should be further applications of the Index to refine it and correlate findings with other output/outcome measures.

Different techniques were employed to survey the opinions of the respondents, does it affect the validity of the findings?

RESPONSE: This is a valid question and we have explained further that the differing formats were only to facilitate completion among respondents in varying contexts. In all cases a local trained consultant was hired, or project assistance provided locally, to contact prospective respondents and to ensure instructions were provided consistently and understood by respondents. We have added this information to the manuscript. In an additional analysis of an indirect measure: attrition (average difference in respondents answering items in each
dimension), we found that the average difference of respondents answering items (or not) was 2.8 overall, with the largest difference among countries being Mali (4.0), where completion was done through hard copies after individual instructions. Ghana, were 85% of forms were responded to via SurveyMonkey, had the lowest attrition of respondents to items and dimensions, at 1.6; the largest difference in attrition among dimensions was tied between Finance and Education/Training (4.2), and not M&E, as may have been suspected. Thanks to the reviewers for prompting us to carry out additional analyses.

Tables: Small adjustments of the details/figures in the tables and consistency in presenting the figures will ease readers.

RESPONSE: Thank you. We have improved titles of tables and formatting (e.g., bolding of totals).

What is the "average respondents" at the last row of each table (Table 2-8)? How it was calculated? What was the use of these figures?

RESPONSE: We wanted to have average scores for each dimension and country, as well as a total score for each dimension and overall. To complete the table we also realized that in each country different respondents had answered individual items within dimensions, so there was a need to get at an "average" number of respondents for each country, as well as for the whole dimension. This has now been used to obtain the "attrition" of differences between all respondents and fewest respondents answering each dimension per country.