Reviewer’s report

Title: Task shifting for point-of-care early infant diagnosis: a comparison of the quality of testing between nurses and laboratory personnel in Zimbabwe.

Version: 0 Date: 13 Nov 2019

Reviewer: 2012 Tweheyo

Reviewer's report:

TITLE AND ABSTRACT
Title: Task shifting for point-of-care early infant diagnosis: comparison of quality of testing between non-laboratory and specialised trained personnel
"Proposed alternative title": Task shifting for point-of-care early infant diagnosis: a comparison of the quality of testing between nurses and laboratory personnel in Zimbabwe.
Comment 1 (minor). Title: The present title is not specific as to the study population and study setting. Specifically, non-laboratory trained personnel could refer, for example to "lay personnel" or to any other professional group. I suggest you specify the cadre you involved in your study, as suggested in the alternative title. If you so wish, the term non-laboratory trained personnel would be appropriate in your discussion section.
Comment 2 (minor). Abstract: Authors state that data were uploaded into an Excel-based database for analysis. Would you like to be more specific as to which database this was? For example, your first reported finding of no significant differences in IQC failure rates between the non-laboratory testers and trained lab personnel, p=0.354. In STATA version 15.1, the computation results in a p=0.327 for the stated statistics.

INTRODUCTION/BACKGROUND SECTION
Comment 3 (minor) The introduction/ Background - adequately sets the scene by stating the problem and justifying why the study was conducted. In the last paragraph of the background, please state the objective of this study more clearly than narrating what the data sources were used for (because the latter will be addressed in the methods section).

METHODS SECTION
Comment 4 (Major)
Under the study design section - "We conducted a quasi-experimental secondary dataset analysis". This should be corrected to - We used a quasi-experiment study design. Then go on to describe the approaches and data sources - retrospective study using …. secondary data sources for the analysis etc. To qualify as a quasi-experiment however, you need to more articulately describe the experimental / intervention group, and the comparison group - so that a reader discerns the quasi-experimental nature. Some questions also need to be answered, for example.
What was the defining feature of the intervention, and how does this differ from the comparison?
How many comparisons were selected per intervention and why?
How were comparisons selected to ensure that there was no contamination of the intervention, or even how was contamination assessed, and/ accounted for in the analysis - especially that this was a secondary data analysis? You might consider including this within your study limitations section.
What was done differently in the intervention, compared to the comparison sites (training, data collections etc). Finally, what exactly was the intervention? - This doesn't come out clearly in the existing write up, even if it is a retrospective analysis, it should be made as clear as possible.
If it is not a quasi-experiment, you had better leave out this nomenclature in your definition of methods, and instead call it ("a cross-sectional retrospective study" using a secondary data analysis). A key
definitional feature of an experiment or quasi-experiment is the timeframe through which the intervention should take effect, with our without the randomisation to the intervention, respectively.

Comment 5 (minor) "Hub and spoke model" needs a description and reference.

Comment 6 (minor) "In consultation with the health facility executive, platforms were either placed in the maternal neonatal and child health department (MNCH) or in the laboratory".

What do you mean by platform? Do you mean POC EID testing sites, or health facilities with specific cadres of health care workers, or training facilities? You need to provide more clarity here. Also, how are the hubs and spokes distributed within the platforms?

Comment 7 (Major). You acknowledge that 6 of your sites had nurses, 2 sites had laboratory-trained staff, while the 2 had a combination. Can you present a table that disaggregates how many tests were collected from each of these site categories (nurses only/ lab personnel only/ nurse and lab)? Could you also run a sub-analysis with these sub-categories as it would delineate confounding?, by stratifying your findings within these (naturally occurring) categories.

Comment 8 (Major): Study participants - "We analysed tests conducted by 45 testers (33 non-lab, 12 lab trained)". How were participants distributed between the platforms (hub and spoke testing sites)? I suggest that you take time to describe in detail the sampling process more clearly. If your first stage was sampling at the platform (hub/spoke level), how then did you sample out the next level participants? Perhaps, you should provide the sampling process a sub-heading - to increase the clarity of your methods and to enhance the potential for replicability.

Comment 9 (minor). Measures - "data were downloaded from ….. Detect platform" What is the Detect platform? Perhaps you need to describe the information management software run by the POC-EID devices, and reference this too. You also need to describe the nature of data generated by this device - which you downloaded into excel. For example, is it .xls files, .dbf files, .mdb files etc.

Comment 10 (minor) Analysis. "We conducted another binary logistic regression model after controlling for clustering" - at what level was clustering controlled for? If you consider level of experience as a cluster - isn't this spurious as you would expect number of tests to be determined by the workload at the health facility. A more reasonable clustering effect, which would determine the similarity of participants at baseline should be - the timing of EID training/ or, the recruitment into the program/ or the category of health facility (MNCH facilities vs Lab-only personnel, vs combination of MNCH & Lab personnel).

Comment 11 (Major) Analysis. Considering that this is reported as a Quasi-experimental design. There is no mention of how the time element was dealt with. For example, what set of results is considered as baseline, and what set is considered as endline? Or better still, how many general time points (if more than two were considered) for the program implementation does this analysis refer to? Is this a pre-post design with a comparison, or just a comparative analysis? Please note, this comment is related to comment 4.

DISCUSSION AND CONCLUSIONS

Comment 12 (Major). "Figure 1 IQC failure rate by type of provider over time". If this were a quasi-experiment, I would expect to see this failure rate in groups (intervention/ comparison) shown against each of the time points for which measurements are considered. In its form - this only confirms a comparative retrospective analysis.

Consequently, the discussion and conclusion for a quasi-experiment would base on a regression controlling for baseline findings over the desired endpoints of analysis (e.g model at endpoint time1, time2, time3 etc). This would be more helpful in informing us, given the resource investment of training, supervision etc for EID nurses, how long does it take to obtain EID test results that are
comparable to the test results of lab-trained personnel?

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

An article whose findings are important to those with closely related research interests

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

Needs some language corrections before being published

**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