Reviewer's report

Title: Effect of Integrated Infectious Disease Training and On-site Support on the Management of Childhood Illnesses in Uganda: A Cluster Randomized Trial

Version: 2

Date: 3 March 2015

Reviewer: Paul Mullan

Reviewer's report:

• Discretionary Revisions = DR (which are recommendations for improvement but which the author can choose to ignore)

1. You mentioned a few task shifting articles (references 10-12) related to inpatient HIV testing, pharmacy refills, and once on HIV primary care roles. References #10 and 11 are less related to your article’s content than #12. Consider adding an article in addition to #12, from a more resource constrained setting (than South Africa) that relates to your article’s topic, such as Monyatsi, 2011 (PMID 22273135).


3. Methods line 296 – this limitation could be in the limitations section.

4. Results. Line 383 – consider adding a line that states something about “The proportions of each type of provider per treatment arm were not significantly different (p<0.05)” to allay reader fears that the samples were not equal (by eyeball calculation, they look approximately the same).

5. Line 409 – “described”

6. Line 490 : into # in

7. Line 504 – unclear sentence.

8. Line 497 – Consider adding a reason as to why you believe treatment and pt/caregiver education did not improve significantly.

9. Line 521 – insured # ensured

• Minor Essential Revisions = MER (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. Introduction p 5, line 79-80 – please expand on how the supervision visits were “effective” and how supervision “may help maintain health worker skills.” As described here, it leaves the reader wondering just how effective these interventions were, and these interventions form the basis of this manuscript.

2. Methods – line 324. Six sets of categories were tested at 0.05 significance, giving approximately a 0.30 chance of a type I error. Consider either adjusting for the fact that you had six categories (eg. Bonferroni method: 0.05/6 = 0.008; if no a priori hypothesis existed for which of the 6 would be significant) or choosing one as your primary outcome for your hypothesis (if one existed a priori to your experiment). Or – expand/explain this “caveat” and its effect on significantly
increasing the chance of a type 1 error.

3. Methods – lines 112-116 – This 3 sentence paragraph is confusing and its content either need to be incorporated further down in the methods section after the outcomes are described, or integrated elsewhere. Revisions to a tool are being discussed before the tool is discussed.

4. Methods – line 118 – Please clarify that these trainees are the MLPs that were discussed earlier.

5. Methods line 252 – remove “the”

6. Results line 386 – this was not an exclusion criteria mentioned earlier. Either add this exclusion criteria to the methods or include this MLP in the analysis.

7. Line 454 – Non significance is not “weak evidence.” It is “not significant.” Likewise line 473, if the CI truly includes 1.0, then one can not say it is statistically significant (if this is a “round upwards” from 0.9958, then you can’t; if it is a round downwards, from 1.0024 then you can).

8. Line 491 – Repetition is also what engrains bad habits of practice. Remove this sentence or the first half of it as the data from this study don’t necessarily prove how the art of learning is achieved.

• Major Compulsory Revisions = MCR (which the author must respond to before a decision on publication can be reached)

1. Intro – line 89-90; Please expand this sentence in 1-3 sentences to describe the effects on these variables by the intervention. This helps the reader understand the clinical impact of these interventions without needing to read 3 more articles.

2. Methods, line 97; describe what a mid-level practitioner’s training entails (e.g. 4 years of nursing school, etc.).

3. Results – line 423 – This sentence sounds contradictory. “no difference in practice… except lower appropriate lab tests ordered and higher % of correct diagnoses.” Please correct this sentence; and, please put some data (percentages, p values, something…) to let the reader know what “lower” and “higher” really mean. The validity of this intervention is largely based on the question “Were these two groups the same at baseline?” This must be described in detail where data are available. Clarifying this will also help line 458’s validity.

4. This was described as a continuous quality improvement project, but traditional statistics were used for all of the analyses. I was surprised not to see any of standard QI-related statistical process control (SPC) charts or run charts to describe any of the QI findings from the study. SPC charts would account for some of the changes over the course of the intervention rather than looking at the combined statistics for each time periods and drawing conclusions. With over 300 observations per group, you could easily draw proportion charts (p charts) over each of the three month periods for arms a and b so that the reader could see whether the effect of IMID on A and B was significant in and of itself over the three month period, and whether the endline period had any similar trends. This would be much more revealing to what actually was happening patient to patient.
I took a picture (might not transfer in BMC upload version but this book is likely owned by authors or easily accessible) from p. 16 and 17 of Provost,Murray “The Health Care Data Guide, 2011” to illustrate this issue. The results from this study mimic figure 1.5 showing a before and after change standard model. Even though you don’t have continuous data from Jan 2010 to Feb 2011, you do have two continuous time periods (Jan 20 to Mar 10) and (Dec 10 to Feb 11) from which you could look at the phenomena of the training itself, and importantly how differences were seen from the beginning to end of each time period. Were the findings from arm A similar to Figure 1.6 Case 5 where the improvement in Arm A happened before the OSS intervention (in the latter half of the first period), but such an effect didn’t happen in Arm B? Or one of many other scenarios. If you don’t have the capability to do an SPC chart for these data, then it should at least be mentioned in the limitations that an SPC chart would account for the time based nature of the data more predictively for the “next patient.” The first author is at Baylor College of Medicine (where I trained as well), where SPC charts are routinely used in QI projects – the Chief Quality Officer there, Dr. Eric Williams, could either help or refer the first author to a QI specialist who could determine if p charts could be successfully used for this paper’s analysis.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests