Reviewer's report

Title: Strategies for cost-efficient monitoring and evaluation of resource-limited national antiretroviral therapy programs.

Version: 2 Date: 26 September 2014

Reviewer: Benjamin Johns

Reviewer's report:

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

1. Is the question posed by the authors well defined?

Yes. The authors are arguing that a two-phase design is an affordable alternative to case-control or population based designs to answer questions that may be relevant to HIV (especially anti-retroviral therapy) programs, especially in developing countries. Considering that I have never seen a two-phase design in a developing country context (although I am admittedly not an expert on the area), I think that drawing people’s attention to this method (which is more accepted in developed countries) is important and relevant.

2. Are the methods appropriate and well described?

Yes, although I had two questions that were not addressed.

A. It seems like the data is multi-level, in that there are some independent variables that measure facility level characteristics (e.g., ownership) and some that describe patient level characteristics. Normally, I would approach such an analysis with a multi-level model in mind in order to correctly estimate standard errors etc. Is this concern valid for two-phase design studies? The authors do not discuss & I am not sure of the answer.

B. In cases such as this, where a relatively large number of patient records need to be collected (minimally 500), it sometimes makes more sense to collect data in clusters. That is, rather than collecting one patient record at facility X and two records at facility Y, it might be better to collect 25 at facility X and 25 at facility Y (even though in the end that means collecting far more than 500 patient records) because it minimized the number of facilities visited. Such methods are common for household surveys. Would they make sense in this type of study as well? (if so, what are the implications?)

3. Are the data sound?

To the extent possible. They do not address the quality of the data they gathered; this is not really the main topic of their paper. They have the best available data from Malawi, I would think. In my experience, however, the type of M&E data can suffer greatly from poor reporting systems (I am not familiar enough to say if this
is the case for Malawi HIV/AIDS treatment data or not, although the authors make brief reference to it in the conclusion). I do wonder if having first phase data that lack full validity would affect the authors’ recommendations or not, but do not have the expertise / data to follow through on the thought. My feeling is that it would not affect the two-phase design differentially compared to the case-control, but that it might affect the sample size needed (e.g., if the number of drop-outs from a facility is poorly reported, you may need to adjust the sample size ‘on the fly’). As noted above, this is not the main topic of the paper, and the authors can raise these issues in the discussion at their discretion.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

Yes. The author state that the study did not use data with patient identifiable information, and was exempted from need for ethical review by 2 IRB board reviews. Data are reported with appropriate confidence intervals, following accepted practice.

5. Are the discussion and conclusions well balanced and adequately supported by the data?

Yes, in general. At a few points I quibbled with their language, but their overall points are well taken. For example:

Page 12 line 14: they state the two-phase design is possible ‘with only a small amount of patient level data’. I agree it is small relative to sampling the entire pool of patients, but the field work & effort involved in collecting 500 to 2000 patient records is not trivial, especially when the individual files could be scattered all over the country. It is, obviously, preferable to the alternatives, but claiming it is small is a bit misleading.

6. Are limitations of the work clearly stated?

Yes.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

Yes.

9. Is the writing acceptable?

Yes. I struggled with the paragraph on page 5 lines 11 to 23, which I found a bit unclear. Hopefully the authors can re-work this paragraph. Otherwise it is very nicely done.

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

The numbers on page 10 line 2 and 3 do not match those in the corresponding table (e.g., if think it should be ‘were female’ instead of ‘were male’ or the table rows are incorrectly labeled).
8. Do the title and abstract accurately convey what has been found?

Yes, although the title mentions ‘costs’, which are not discussed, The authors discuss methods are sampling or data efficient, and this likely leads them to be cost efficient as well but this last issue is not directly assessed in their paper. I would prefer they be precise in their language.

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

None.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests