Author's response to reviews

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

Authors:

Sebastien Haneuse (shaneuse@hsph.harvard.edu)
Bethany Hedt-Gauthier (bethhedt@gmail.com)
Frank Chimbwandira (frankchimbwandira@gmail.com)
Simon Makombe (simond.makombe@yahoo.com)
Lyson Tenthani (ltenthani@gmail.com)
Andreas Jahn (drandreasjahn@gmail.com)

Version: 3
Date: 1 February 2015

Author's response to reviews: see over
Dear Editor:

On behalf of my co-authors, please find attached a revision of our manuscript titled “Strategies for cost-efficient monitoring and evaluation of resource-limited national antiretroviral therapy programs” which I hope you will consider for publication in the BMC Research Methodology. Also, attached below, please find point-by-point responses to the two reviewers.

We have very much appreciated the additional time that you have given us to complete this revision. As mentioned in an email dated 01/07/2015, reviewer #2 raised a very important point regarding the analyses. We have been working on methods that address the reviewers concerns but, unfortunately, are not in a position to disseminate them. We have responded to the reviewer and written a detailed paragraph in the Discussion on this point. For this manuscript, however, we believe that this issue does not affect the key message of this paper but rather reflects the need for additional methods. We hope you agree.

If you have any questions regarding this submission, please do not hesitate to get in touch.

Sincerely,

Sebastien Haneuse, PhD
Associate Professor of Biostatistics
shaneuse@hsph.harvard.edu
Reviewer #1

The authors make a compelling case for the pitfalls of ecological data using the cross-sectional audit of all person-level data in Malawi and showing the differences in the predictors of a negative outcome. The approach is compelling and the application is potentially useful and could improve efficiency.

We thank the reviewer for the time taken to review the manuscript and are pleased that they found the paper interesting.

In terms of the article, a few comments are worth noting:

1. There is insufficient detail given on the proposed stratification scheme. The first phase is described as “phase I would correspond to a stratification of the entire population on the basis of outcome status (as in a case-control design) and some combination of the known aggregated quarterly-clinic cohort data.” What does “some combination mean”? How is stratification based on the outcome status possible without a complete cross-sectional census? These are unclear and lacking in detail.

We appreciate that the reviewer pointed out the lack of clarity in this section. Indeed, this description was one of the most challenging aspects of writing the manuscript. Part of the challenge lies with the flexibility of the two-phase design; any way in which one could use the group-level information on the entire population could, in principle, be used to form a phase I stratification. We hope that the revisions in the manuscript, along with the additional examples, has clarified the use of the two-phase design in this setting. Finally, we have added a reference to Wakefield and Haneuse (2008) to this paragraph that provides an alternative but similar context and, we hope, will help the reader.

2. There is no theoretical and mathematical framework to justify the authors’ proposed approaches. The precise mathematical frameworks (with applications) are absolutely needed here to enable generalizeability and to ensure that all steps logically follow from one another.

We understand that the manuscript is brief on the theoretical aspects of the methods. Indeed, we have been purposely brief in this regard because most of the existing literature on two-phase designs is theoretical. As such, the intended audience for this paper is a combination of statistical analysts and substantive researchers who work “in the field”, with overarching goal of the manuscript is to highlight the utility of the two-phase design in the ART M&E setting. Notwithstanding this, we do note that references to two key papers that lay the theoretical foundation for the methods (Scott and Wild (1997) and Breslow and Holubkov (1997)) and would anticipate that readers interested in the theoretical details could use them as a start.

3. The example using the ecological data is misrepresented. The authors had choices when aggregating to the clinic-level. The choices made for some odd comparisons with the person-level data. For example, age aggregation (first section of Table 1) would not intuitively be about the portion of clinics with mean age <30, 31-35, etc. A legitimate comparison with the data presented in the individual-level data would be about the mean % in the age group 16-25, 26-35, etc in the various clinics. Throughout it feels like the authors made decisions that would make the ecological comparison fare poorly. The article would be strengthened if the ecological analysis were as strong as possible, and still biased.

We respectfully disagree that the example is a misrepresentation. In practice, as with any dataset (be it individual-level or aggregated), researchers have many options on how to proceed with an analysis. We agree with the notion that it would have been possible to cherry-pick our analyses to make the ecological results biased. We did not do this and proceeded in a manner that we believe we would have done in a real analysis. Either way, we want to emphasize that the purpose of showing bias is not to say that there will always be bias. Sometimes there will not be bias but the key point is that without any individual-
level data one would never know. Hence the importance of collecting individual-level data.

With respect to age in Table 1, we emphasize that the group-level information is the average age among the patients enrolled in the quarterly-clinic cohort; that is, one would only have a single number for each quarterly-clinic cohort. As such, one couldn’t report the mean % in each of the age groups within the various clinics. We also note that the group-level age variable was only categorized for the purposes of Table 1. In all regression models we left age in its continuous form.

*In the end, this is a potentially useful stop-gap for ART programs, but the core issue here is that the routine monitoring data available from ART programs is unacceptably poor. It is utterly puzzling why patient-level data cannot be compiled and analyzed beyond the clinic. If the clinic reports are put together from pieces of paper, why can’t these pieces of paper move up the data chain and get entered into a database that allows patient-level monitoring. The costs for streamlining that process, if anything, should be lower than current effort and time spent on creating ecological data reports with unknown data quality.*

We agree with the sentiment that data that is routinely available is often poor. The reality, however, is that Malawi is very resource-limited and it has taken considerable work and effort to build the ART program to what it is today. We hope that the program will eventually shift to an all-electronic system which should give researchers and policy-makers at the Ministry of Health with better quality data and mitigate the potential for ecological bias.
Reviewer #2

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.

We appreciate the time the reviewer took to review the manuscript and are glad that they found the paper 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.

This is an excellent point and we very much appreciate it. The reviewer is quite right to point out that in the complete data setting one would need to acknowledge potential cluster-correlation to ensure that inference is valid. Interestingly, to our knowledge, no methods for two-phase designs have been developed in the cluster-correlated data setting. Indeed, as far as we can tell, no methods for the standard case-control design in the cluster-correlated data setting have been developed. This suggests, to us at least, that any case-control study reported in the literature and for which the underlying patient population is naturally clustered may be not correctly analyzed! In the revised Discussion we comment on the lack of methods although do not go so far as to comment on the applied literature more generally.

For this paper, we have decided to leave the results as they are but, as mentioned above, acknowledge the point in the Discussion. We are currently working on developing new statistical methods for the analysis of cluster-correlated case-control and two-phase studies. Unfortunately, given the timeline given to us by the journal we are not in a position to report on those methods. We understand this is non-standard but even though we should, strictly speaking, account for the correlation we believe that the results we report in the paper remain valid for the key points we are hoping to convey. This is, that ecological bias can be alleviated by collecting individual-level data and that the two-phase design can provide efficiency gains relative to the case-control design.

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?)

These methods would indeed make sense and represent valid approach to collecting individual-level data. The Discussion mentions such methods (and includes two additional references). We have retained a focus on the two-phase design because of its flexible, explicit use of the aggregated data.
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.

The reviewer raises an excellent point. We agree that data quality is likely a huge issue although, as we believe the reviewer understands, one that is outside the scope of this manuscript. With regard to the use of the two-phase design, the reviewer is correct in that the two-phase design is still applicable with the overall validity depending on the quality of the phase II data. Efficiency gains by stratifying are still possible but will likely be somewhat diminished if the phase I data lack full validity.

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.

We understand the reviewers point and have changed to sentence to read “…the designs are flexible and can often permit the investigation of patient-level outcomes/associations with detailed information on only a fraction of patient registrants.”

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.

We have re-written this paragraph, expanding the description in a way that we hope the reviewer will find the text clearer.

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).

We do apologize for the mistakes in the first submission; for the specific one pointed out by the reviewer, the text should have been “were female”. We have been through the numbers in the text and believe that they now all correspond with those reported in the tables.

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.

The review is correct in that we took “cost-efficient” to be a reflection of the lower costs that would results from the smaller sample sizes that are needed under the two-phase design. We have modified the title to remove “cost-effective”. We have also modified the Conclusions sections in both the Abstract and the Discussion.