Reviewer's report

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

Version: 2 Date: 29 July 2014

Reviewer: Eran Bendavid

Reviewer's report:

The article “Strategies for cost-efficient monitoring and evaluation of resource-limited national antiretroviral therapy programs” by Haneuse and colleagues present a stylized example of a potentially useful approach to improve the quality of the monitoring data in ART programs in resource-limited contexts. The premise is intuitive and appealing: most monitoring data comes as facility-level data rather than individual-level data; facility-level data enables low-cost monitoring, but may lead to ecological biases in monitoring events like loss to follow-up. They propose a stratification approach for sampling individual-level data, guided by the clinic-level data.

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. 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.

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.

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.

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.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

None