Reviewer's report

Title: Can HIV incidence testing be used for evaluating HIV intervention programs?

Version: 1 Date: 15 March 2010

Reviewer: Raymond Scott McClelland

Reviewer's report:

There has been substantial interest in the potential for using HIV incidence assays to determine the rate of new HIV infections in populations using cross-sectional data. The ability to evaluate the effect of intervention programs using cross-sectional surveys would be an important development. In the present study, the authors ask, “Can HIV incidence testing be used for evaluating HIV intervention programs?” The article demonstrates that the effect of adult male circumcision in the Orange Farm trial could have been estimated using cross-sectional data from incidence assays performed at the final follow-up time point.

Major Compulsory Revisions

1) Use of incidence assays to estimate the relative risk in two arms of a randomized trial is a unique use of these assays. The RCT design addresses both recognized and unrecognized sources of bias by randomizing and comparing between the study arms. Using incidence assays in this scenario, misclassification due to the test characteristics (i.e. the incidence assays) should be randomly distributed between the two arms. As a result, a risk ratio defined by using the assays in both arms of the trial provides a reasonable approximation of the effect of the intervention. Similarly, as the authors point out, it is not necessary to know the window period precisely when the assays are used to estimate the RCT effect.

While use of the assays in an RCT provides a reasonable estimate of the risk ratio that would be obtained by longitudinal survival analysis, the same would not necessarily be true in an intervention program where there is no randomization. For example, if the tests were compared before versus after an intervention (e.g. circumcision), changes in factors such as the incidence to prevalence ratio, the ratio of different HIV subtypes, and the use of ART might result in different operating characteristics for the incidence assays at different time points. Similarly, if men who chose circumcision were compared to men who did not choose circumcision at a single time-point, differences could exist in the incidence to prevalence ratios, ages, etc. of the populations.

If the authors feel that this type of testing could be used for a cross-sectional study of a single-arm intervention, more explanation of how they feel that the RCT results can be generalized to non-RCT settings is essential, as it is currently not clear. Alternatively, if the intent is to say that this would be an intervention
that would be useful in RCT settings, this needs to be made clearer in both the title and the manuscript. Statements like, “Such results imply that HIV interventions may be assessed using HIV incidence assays on samples obtained from a cross-sectional survey by calculating incidence rate-ratios (Discussion – 1st paragraph),” may otherwise mislead readers because the text does not clarify that the conclusion only applies in RCT settings.

2) It is curious that overall, correction for misclassification has produced results that are, for the most part, farther from the “true” estimate provided by longitudinal analysis. This should be addressed in the discussion.

3) On page 5, second paragraph, the paper states, “We calculated the proportion of false long-term seroconverters and the proportion of false recent seroconverters among those with long term infection.” The former category is difficult to interpret; who would be classified as a false long-term seroconverter among those whose true status was long-term HIV-infected?

Minor Essential Revisions

1) The abstract talks about the correction for misclassifications, but gives no explanation of what is involved. This makes it difficult to interpret the findings that are presented in the abstract.

2) Table 3 “NE: Not Calculable” should be “NC: not calculable”

Discretionary Revisions

1. The authors point out that, in contrast to another study that they cite, this analysis found only moderate agreement between the two assays. It would be helpful to also comment on the possible reasons for this.

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:

'I declare that I have no competing interests'