Reviewer's report

Title: Late Presenters to HIV Care and Treatment, Identification of Associated Risk Factors in HIV-1 Infected Indian Population

Version: 1 Date: 4 May 2010

Reviewer: Dallas Swendeman

Reviewer's report:

This paper has the potential to make an important contribution to the knowledge base of HIV epidemiology in India, with implications for interventions for early HIV identification (testing) and linkages to treatment. Revisions are necessary to bring to publishable quality.

The primary limitation is around how patients “late presentation” to the clinic/laboratory from which the data are drawn is operationalized. The abstract is clear that the focus is on late presentation to CD4 T cell counts at the laboratory. However, throughout the text this focus is conflated with discussions of late presentation to HIV testing, CD4 T cell counts, treatment, and care. The second main limitation involves the apparent lack of truly multivariate (i.e., multiple predictor) statistical analyses. Implications for the methods, results, and discussion are outlined below.

Major Compulsory Revisions:

1. First, in the introduction, some background on the context or systems for HIV testing, treatment, and care in India (or in the region covered by the laboratory’s services) will help readers interpret the biases and limitations of the sample, data, and results.

2. Similarly, in the methods section, more details are needed on the specifics of the patient flow or pathways to the lab for CD4 testing. In the beginning of the methods section the authors state that individuals were coming for CD4 testing for the first time after HIV diagnosis “usually within a week of positive HIV serodiagnosis”. Yet, in the discussion section it states that the “majority of patients are referred here after efforts to treat them at primary and secondary health care levels and also at private settings”, which contradicts the focus on treatment naïve individuals and the recent diagnosis statements in the methods. Clarification is needed. Ideally, the analyses would control for time since testing HIV-positive. Either way, this should be stated more explicitly as a study limitation in conjunction with discussion of the generalizability of the findings for either India as a whole or the region served by the lab.

3. Clarify the statistical methods description and very likely conduct more analysis. Both “multivariate” and “multinomial” logistic regression are mentioned. Multinomial logistic regression is the appropriate description for the 3 level outcome analysis of late and very late presentation compared to not late (i.e.,
CD4 >200). What needs clarification is whether or not the analysis is “multivariate” in that all of the patient factors examined as “predictors” were then included in a single model, which they should be, including some “model building” to eliminate non-significant predictors when “controlling” for all of the factors examined. It appears that the second page of results and Table 3 present a series of what are often inconsistently referred to as “univariate” or “bivariate” analyses, which do not control for confounding of the other factors examined. What is needed is a true “multivariate” analysis in which, at minimum, all patient factors that were statistically significant in Table 3 are then entered into a model together, thereby reducing the confounding effects of the multiple factors. This will give the authors and readers a much more precise indication of the factors that are most significantly associated with late presentation. As it stand now, so many of the factors examined are statistically significant that the authors are not achieving their true goal of identifying patient factors that will effectively guide targeted intervention for earlier testing and linkage to care.

Minor Essential Revisions:

4. Odds ratios should be presented with 95% confidence intervals consistently in the text (as done in Table 3), rather than p-values.

5. The last paragraph of the results section simply presents inversions of the statistically significant odds ratio estimates presented earlier and in Table 3. This is not particularly informative in the results, although it might be useful to incorporate such interpretations in the discussion. Again, however, the focus should really be on multivariate model results, which will significantly impact the content of the discussion.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

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