Reviewer's report

Title: Implications of the HIV testing protocol for non-response bias in seroprevalence surveys

Version: 1 Date: 4 August 2008

Reviewer: Michael Sweat

Reviewer's report:

This is an interesting paper that provides insight on the level of bias found in estimating HIV prevalence using seroprevalence surveys in a hospital setting in Ethiopia. Establishing accurate HIV prevalence estimates is of great interest to many, and there has been considerable controversy over the veracity of such estimates in the literature.

- Major Compulsory Revisions

(1) The manuscript refers repeatedly to “post-test counseling”, which is assumed to mean provision of HIV test results. Standard HIV counseling and testing protocols, such as that provided by the World Health Organization, proscribe both provision of HIV test results, and risk reduction counseling. These are not the same thing. The text should be corrected to clarify whether risk reduction counseling was provided, and differentiate provision of HIV test results from counseling.

(2) It should be clarified whether the study was approved by an ethical committee, and whether study participants received informed consent to participate in the study. Note that consenting to receive HIV testing (which is an outcome of the analysis) is not the same as consent to participate in the study.

(3) A key variable in the analysis is the estimated likelihood of HIV infection associated with IDC diagnostic codes. The methods section does not adequately describe how these proportions were established (as seen in Table 1). The current description reads “For each entry we calculated the proportion of HIV positives and these are used as an indicator of the likelihood of infection.” Much greater detail is needed here on how this was done. Near this reference it also says “We therefore assume that HIV prevalence in each group of admission diagnoses is independent from the willingness to be tested.” This has unclear meaning.

- Minor Essential Revisions

(1) Table 2 – clarify what “col” means.
(2) Table 4, model 2 – both age and age squared are entered into the same regression model, leading to likely multicollinearity as these are monotonically related. Kindly clarify if there is a rationale for including both in the same regression model.

- Discretionary Revisions

(1) Use of “Consent level A, B, C” makes the text a little tedious to follow, requiring that the reader know that “A” represents full consent, “B” represents consent to be tested without receipt of results, and “C” represents declining. It is suggested that some alternative labels that are more descriptive be used.

(2) The manuscript would be more accessible if statistical jargon were reduced, especially in the methods section describing Heckman modeling.

(3) On page 13, the following statement is unclear, “While these findings suggest that most population-based surveys follow a study protocol that contains bias, it does not mean that there is no bias at all.”

(4) Reference to “disclosure” of test results is a little awkward. The term “disclosure” related to HIV testing is widely used to refer to the client disclosing the result. It is suggested that the term be replaced with “provided” or some other similar phrase.

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.