Reviewer's report

Title: Seroprevalence and risk factors of herpes simplex virus type-2 infection among pregnant women in Northeast India.

Version: 1 Date: 7 January 2011

Reviewer: Christine Johnston

Reviewer's report:

This is a well conducted study of HSV-2 seroprevalence in antenatal clinics in Northeastern India. It provides needed data about HSV-2 seroprevalence and elucidates risk factors for HSV-2 seropositivity in this region. The methods for determining seroprevalence and the statistical methods require some clarification; otherwise a solid paper which answers an important question.

Major Compulsory Revisions:

1. Methods: specimen collection: I have some concerns about using the Focus ELISA in this low prevalence population, particularly because it can have a low positive predictive value (PPV) when index values are <3.0 (Golden STD, 2005). Higher index values may have better specificity and PPV (but potentially at the cost of sensitivity). Therefore it is important to know what was used as a cut-off in this study. What cut-off was used to determine a positive test on the Focus ELISA? How were equivocal samples handled? If a higher cutoff was used how would your results change? Please add how you determined a positive value and consider a paragraph in the discussion about the use of the Focus ELISA.

2. Methods: Statistical analyses: I have several questions about the univariate and multivariate analysis
   a. Please be specific about how odds ratios were calculated (ie logistic regression). Please calculate p-values for groups, especially when using continuous variables (age) rather than for each category.
   b. The univariate analysis includes many fewer variables than collected in the questionnaire. How did you decide to include only these variables? Were all variable tested? For instance, it seems that FSW or exchange of sex for money would be important risk factors to consider, but these are not included.
   c. Multivariate analysis: I am unclear whether the multivariate model includes only age or if it includes all categories. Please clarify in methods. If only age was included in the multi-variate model, why? Were there attempts to build a more complex multivariate model using all significant co-variates? Were there attempts to look for interactions (for instance, state and religion). I don't think the multivariate model, as it is now, alters any of the results from the univariate analysis, and is not that informative.

3. Results: Give numbers in addition to percentages (ie literacy rate: xx (84.5%) etc. A Table 1 with n(%) of general demographic and OB history of your
population may be considered.

4. Results: Seroprevalence: could give 95% confidence intervals around the overall seroprev and the seroprev per state

5. Results: Seroprevalence: The difference in seroprevalence among the states is very interesting; in this study “state” may be the most significant risk factor for HSV-2 seropositivity. Are there differences in these states (religion, male circumcision rates, education) that could explain the differences (see comment 6 in minor essential revisions)-could include in results or discussion?

6. Results: Seroprevalence: Age: Multiple epidemiologic studies have shown increasing HSV-2 seroprevalence with increasing age. The results in this study that HSV-2 “peaks” in the 22-25 year age group seems artificial, and the lack of significance in the >29 yr group seems to be due to lack of power. How were the age strata chosen? Would it make more sense to keep age as a continuous variable? If you keep it as is, use an overall p-value for the category.

7. Income group: Few in high income, would it make sense to combine with middle income for more power?

8. Results: seroprevalence: give numbers for comparisons/95% CI or p-values for 1st compared to 2nd/3rd trimester and GUD compared to no GUD

Minor Essential Revisions:

Abstract:
Background: 1st sentence Do not capitalize Simplex Virus Type
Methods: Last sentence: verb: were instead of was…risk factors associated with HSV-2 seropositivity.
Results: 1st paragraph: second sentence: Missing verb: was
Include comparison group for “Middle income group” (compared to low income group)
Last sentence: genital is misspelled.
Discussion/reference: Ref 25 could be updated

Discretionary Revisions:

1. Abstract: conclusions: Consider removing statement that “routine antenatal screening for HSV-2 may be helpful for safe pregnancy outcomes”. This is a controversial statement and there is no data in the current study regarding this issue.

2. Background: The HSV-2-HIV link is mentioned in several parts of the first paragraph. To tighten the paper, I suggest discussing seroprevalence of HSV-2 in the first paragraph (and perhaps including more information about HSV-2 seroprevalence studies in India), and the HSV-2/HIV connection in the second paragraph.

3. HIV prevalence rates in Northeast India, particularly in antenatal clinics, would be helpful to include, if available, to give the reader a sense of the burden of HIV in this population.
4. Results: Risk factors: The authors state throughout the paper that women with “multiple” sex partners have higher risk of HSV-2. Rather than multiple state “more than one in past year”. It would be most helpful if you had lifetime sex partners-is this available?

5. Discussion, 2nd paragraph: Results of this study are different from US, in that in India HSV-2 is associated with higher SES rather than lower. Care to speculate about these differences?

6. Discussion, paragraph 4: As above (comment 3), the difference in seroprevalence rates by state is intriguing and it may be interesting for the authors to tease out these differences in the multivariate model and/or speculate on reasons for these differences.

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

**Quality of written English:** Needs some language corrections before being published

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

**Declaration of competing interests:**

I declare that I have no competing interests.