Reviewer’s report

Title: Epidemiology of Human Immunodeficiency Virus-1 and Hepatitis B co-infections and risk factors for acquiring these infections in the Fako Health District

Version: 2
Date: 22 May 2015

Reviewer: Trevor Crowell

Reviewer’s report:

This article presents a cross-sectional survey of HIV and hepatitis B prevalence among a convenience sample of volunteers presenting for voluntary screening at five hospitals in Southwest Cameroon. The prevalence data are of interest to regional caregivers and researchers. The authors also use logistic regression to investigate factors associated with HIV infection and HBV infection, with a focus on whether each individual infection was a risk factor for the other. Their conclusion is that HIV is not a risk factor for HBV and they interpret this as evidence, along with other aspects of their analysis, that sexual transmission of HBV is not a driving force in Southwest Cameroon. This is not a surprising conclusion. The statistical methods used for this analysis are acceptable, but could be bolstered with the inclusion of sensitivity analyses since the authors use a number of relatively arbitrary cutoff points for categorical variables and do not include all collected variables in multivariable regressions. The discussion section of this manuscript could be substantially expanded.

Major Compulsory Revisions:

1. Lines 61-63. The authors state that the prevalence of HIV in Cameroon is 4.3% and the prevalence of HBV is 9%. This contradicts the leading statement in the abstract background that 20% of the population in Southwest Cameroon has one or both infections. If the statement in the abstract refers to data specifically from Southwest Cameroon, rather than country-wide, these data should be highlighted in the manuscript introduction. If the statement in the abstract refers to the present study, it should not be included as background.

2. The authors construct their multivariable models by including only variables that were significantly associated in univariable analysis. Constructing a multivariable model in this method may exclude important variables. Are there any variables the authors would consider including a priori, based on prior knowledge? The authors should perform sensitivity analyses to test the effects of their statistical methods on the interpretation of their data. For example, they could consider performing a sensitivity analysis in which a higher threshold p-value is used to select variables for inclusion in the multivariable model.

3. The authors decided to code age as a binary variable with a cutoff of 32 years. The study population includes volunteers as old as 94. More granular age categories should be explored, at the very least in sensitivity analyses.
4. Lines 243-248. In the discussion, the authors state that the results of this analysis support their initial hypothesis, which had never before been stated. This hypothesis needs to be stated in the introduction. The authors already present information in the introduction supporting this hypothesis. A case needs to be made for why this requires the further investigation described in this manuscript.

5. The researchers administered a 30-question questionnaire, but far fewer variables are modeled. The authors should explain how these variables were selected and others were excluded from the analysis.

Minor Essential Revisions:
1. Line 25. The abbreviation HBV should be defined at first use in the abstract.
2. Line 33. Some words are missing here. Perhaps “…to conduct final model building”? This reviewer recommends dropping the phrase “…and final model building.”
3. Line 33 and throughout. The authors should be careful not to confuse “correlation” and “association”. Correlation describes the strength of a linear relationship between two variables. The authors of this manuscript are generally investigating associations, not correlations.
4. Line 37. The phrase “significantly crude associated” should be changed to “significant crude association”.
5. Line 40. The abbreviations AOR and CI should be defined at first use in the abstract.
6. Line 42. The word “association” should be changed to “associated”.
7. Line 52. The word “cause” should be changed to “causes”.
8. Line 60. “The phrase “is sparse” should be changed to “are sparse”.
9. Line 64. The word “kin” should be changed to “in”.
10. Line 65. The word “rates” should be changed to “rate”.
11. Line 115. The word “or” should be changed to “of”.
12. Line 183. It is not clear why median income is provided as a range of 10,000-50,000 CFA. Median should be one number and, according to the authors’ methods, that one number should define the cutoff for their income categories (currently 50,000 CFA).
13. Lines 185-186. This sentence needs to be reworked. As written, it states that the “no” category indicates a volunteer responded “no to never having a transfusion.” This double-negative suggests that volunteers in the “no” category have indeed received a transfusion.
14. Lines 197-198. As in the abstract, some words are missing from this sentence. This reviewer recommends dropping the phrase “…and final model building.”
15. Lines 201-202. Please define the criteria for determining statistical significance.
16. Line 219. The phrase “a univariate analysis” should be changed to “univariate analyses”.
17. Line 240. The phrase “likely to infected” should be changed to “likely to be infected”.
18. Line 289. “XAF” should be changed to “CAF”.
20. Tables 2 and 3. It is not clear why some results are bolded. For example, statistically significant results in the crude column are not always bolded. Please explain in a footnote.

Discretionary Revisions:
1. Throughout the manuscript, there are instances where it can be confusing to the reader whether the authors are discussing HIV/AIDS and HBV separately or referring to “HIV/AIDS and HBV co-infection,” which is a cumbersome phrase. The authors could consider referring to HIV/AIDS simply as “HIV” throughout the manuscript and referring to co-infection as “HIV/HBV co-infection” (as was done on lines 107-108).
2. The methods section could be shortened substantially.
3. Lines 83-86. These lines restate points in the preceding paragraph. Consider removing or integrating into the preceding paragraph.
4. Lines 86-88. Consider incorporating these lines into the subsequent paragraph.

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.