Reviewer’s report

Title: Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the HIV epidemic in Malawi

Version: 0 Date: 25 Oct 2017

Reviewer: Yanni Xiao

Reviewer’s report:

General Comments:

This paper aimed to explain geospatial variation in the severity of the HIV epidemic in Malawi by constructing gender-stratified epidemic surface prevalence maps and risk maps based on HIV-testing and behavior data. The authors found that the geographic variation in the size of the high-risk group can explain a substantial proportion (76% for women, 65% for men) of geographic variation in HIV prevalence. I guess BMC Medicine readers would find this article interesting and the study well-done once a few changes have been made.

Several comments and suggestions:

1. In the Abstract and Results section, the authors stated that the majority of HIV-infected individuals (~75% of women, ~80% of men) and most of the high-risk group (~80% of women, ~85% of men) lived in rural areas, but why prevalence and the size of the high-risk group are highest in urban centers? Whether this means one possibility that most of infected individuals go to work in urban centers in the day but go back home in rural areas in the night? If this holds, how to include the effect of mobility directly in the model?

2. In page 4, line 14, the authors said that Malawi was divided into 26 administrative districts, but Malawi demographic and health survey shows that Malawi has 28 administrative districts, whether the authors omitted two districts or combined some districts? Please check that.

3. In page 4, line 51-53, which kind of kernel, Gaussian kernel or finite extent kernel, is used in the adaptive bandwidth kernel density estimation method? Why the authors choose the smoothing parameter as 200 (the ring size)? Is it the optimal value? If not, what is the optimal value of this smoothing parameter?

4. In page 5, line 40, the authors used Queen's contiguity to assign spatial weights which is based on common edge or vertex, whether a different weight assign method such as distance-related weight will affect the main conclusion? In this case, what is the proportion of geographic variation in prevalence can be explained by the size of the high-risk group?
5. In page 6, line 34, "...the results in Figure 2 show that in Malawi ~66% of HIV-infected women, and ~50% of HIV-infected men, ...", how to obtain the number "66%" and "50%", which is not intuitive. Please briefly explain this calculation.

6. In page 23, Table 1, the results of this Table are for one district or the average value of all districts? If the later holds, please list the similar table for each district in the appendix so that the readers can compare the results for different districts.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.