Reviewer’s report

Title: Spatial heterogeneity of environmental risk in randomized prevention trials: consequences and modeling

Version: 0 Date: 13 Feb 2019

Reviewer: T. Alex Perkins

Reviewer’s report:

Summary

This paper examines how spatial heterogeneity of environmental risk (proximity and importance of mosquito breeding sites) can bias efficacy studies. Specifically, the authors utilized data from a malaria vaccine study to show that estimating the intervention effect of the malaria was biased, especially in situations when the proximity to breeding sites is of high importance to disease risk.

My Review

While this paper does not present any new methodology or strategies for examining intervention efficacy in situations of spatial heterogeneity of environmental risk, it does quantify how drastic this can be in certain cases. All sections besides the results section are presented clearly and logically. However, I think that the results section needs a lot of tuning up. Numbers are thrown in all over the place and it is difficult for the reader to draw big-picture conclusions. Different measurements of the bias are used at different times, and sometimes inappropriate statistics (such as minimum/maximum/range of bias) are used. This section could also be much shorter, and the tables the authors provide say a lot about the specific numbers if a reader is really interested in the details. This paper could be a useful starting place for further analysis on how to conduct efficacy trials when faced with spatial heterogeneity.

Specific Critiques

The author mentions specifics about malaria a lot when the details are not unique to malaria. I think malaria should be brought on later in the paper for the example and not integrated early.

Seems unnecessarily pessimistic at times- saying we will never be able to measure environmental risk but it biases the studies. Where are the suggestions on how to improve this? Also, an explanation of what "important but not measurable" environmental effects are would be helpful. Also. Contrast this with random spatial effects.

In general, the results section is too numerical and wordy. Hard to draw meaningful conclusions out of. The graphs are more helpful. Metrics like maximum/minimum/range bias are hard to
generalize, mean bias would be easier. I would move the tables with long charts of values to the supplement. MSE might not need to be reported for efficacy, rather SR and CR are enough.

The simulation structure is only shown in the one figure. I think it needs more explanation within the paper itself. Also, this is person controlled and not cluster controlled which I think should be mentioned.

In general, clarification and definition of certain epidemiological terms should be more prevalent. For example, treatment effect could either be percent of cases averted or cases that occurred despite the intervention when compared to no intervention. Same with population density, age brackets, protective effect of age, censoring time, etc.

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

No

**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