Reviewer's report

Title: Bayesian Spatial Analysis of Demographic Survey Data. An Application to Childhood Diseases in Zimbabwe.

Version: 2  Date: 22 September 2014

Reviewer: Mikiko Terashima

Reviewer's report:

Comments on the paper entitled: Bayesian Spatial Analysis of Demographic Survey Data. An Application to Childhood Diseases in Zimbabwe.

Overall, this is a very well written paper, and it was a pleasure to read it. It has many findings and discussions that are applicable for policy and important to consider for future research. The paper advocates for spatial analyses for childhood diseases using Bayesian approach to better understand geographical (area) factors that are potentially important but not often measured or measureable. I am in agreement with the author’s point that a shortcoming of existing population health studies lies in the fact that they tend to look at supra-individual factors as typologies and not something embedded in specific geographies.

<Major Compulsory Revisions>

The point for overall improvement I would like to suggest is the following. I elaborated in three ways but it is really one issue of how the paper is framed:

1. The paper reads as though it is an excerpt from the author’s Ph.D. thesis. It is too long, with all components of the paper too detailed.

2. The description of the statistical methods is too technical and long. As I understand it, the journal caters to a mix of audience who are familiar with either quantitative or qualitative methodologies. Most readers are unlikely experts in statistics, biostatistics or Bayesian approaches. Even for epidemiologists, this level of description would be challenging to understand. It would be easier for the audience to digest, and hence continue to read, the paper if the author keeps the description of statistical analyses to bare minimum, written in as plain a language as possible. Because it is anticipated that most of the technical descriptions in the methods section will not be included, my comments on the technical description in the below ‘Specific comments’ section are minimum.

3. In short—and this is probably the most important point to consider—the tenet of this paper for this particular audience should be what the author could discover because he used the methodology, and not about the methodology.

<Minor Essential Revisions>
Overall editorial suggestions

There are two editorial points that need to be addressed throughout the paper:
1. The author should make the tense consistent throughout the paper. In the earlier part of the paper, the author used “will be” several times, while this is a reporting of the research the author has already conducted. So past tense should be used throughout.

2. When a sentence includes citation of an author’s name, the name needs to be spelled out in addition to the reference number. For example, “The spatially structured and unstructured variation was modelled using…. Gaussian heterogeneity priors as proposed by Besag, York and Mollie [13].” Alternatively the author can rephrase it (in this case), “… as proposed previously [13]”. It should not be left as “…as proposed by [13].”

These two points are not included in the specific comments below.

Specific comments

Abstract

“social proximate determinants” # Unclear what it means. The audience for this journal may not be familiar with social epidemiological/population health lingos. It will be helpful if the author can clarify this.

“The geographic distribution of ….” # “Understanding the geographic distribution of…”

It would be clearer if the sentence is reorganized as:

“Bayesian semi-parametric regression models were used to quantify the spatial risk of childhood diarrhoea, fever and cough as well as associations between the childhood diseases and a range of factors, after accounting for spatial correlation between neighbouring areas.”

“A higher proportion of male than female children…” # “However, higher proportion of male than female children…”

Related to my overall comments, the abstract should highlight what the author could find because he used the particular method that could not be found otherwise.

1.0 Background

<Line 61>
“…a range of epidemiological variables” It would be helpful to clarify what ‘epidemiological variables’ mean.

“…which are a great deterrent militating against the demographics of under-developed nations” It is unclear what the author means. It needs to be elaborated.

This entire paragraph is probably not necessary in the background, especially once the author makes this paper more oriented toward the particular journal’s audience. Some general description could be kept and included in the methodology section.

“Such research into…” this description is OK if the above paragraph is taken out. Otherwise it is unclear what ‘such research’ is.

Overall population size (for the country at least, and for individual districts somewhere) should be described here or in the result section. Related, because this study is about geographical/spatial effects at the district level, district level population and sample size description (perhaps in a table) should be stated.

It is helpful to clarify what “anthropometric status” means.

Add {?”} at the end of the sentence starting “Has your child had…

“Consider the triple (…)…” The meaning is unclear.

3.0 Results

Overall, this section (stretching from page 17 to the beginning of page 30) is too long and detailed. It will keep the focus of the paper (and the audience’s attention) if the author presents only a few of the most interesting findings.

“Table 4 gives the… corresponding 97.5% credible intervals” Isn’t it 95%?

“…, single parenting, …”. This is curious. It would be nice to see the elaboration of this in the discussion why single parents’ children had a better outcome, which is contrary to our common perception.
4.0 Discussion
This section is also extremely long.

It would be clearer if the author explicitly describe which gender is favoured in which context being described.

"Rational and optimizing parents" #While I understand what the author is trying to say, it would be helpful if this is elaborated.

"..., education in mothers..." #"higher levels of education in mothers..."

This part lacks "so what" conclusion from the discussion. Why would educated parents be unable to reduce the risk of exposure of their children to risk factors? What is a possible explanation to this phenomenon? If it is unknown or not possible to speculate, perhaps the author could state that a further investigation is necessary.

It would be helpful if the author articulate what “the mother’s resources” mean. Breast milk? Time? Other kinds of care (bathing, cleaning, clothing)? Attention?

"This period is when most children are weaned." This sentence should be before the previous sentence starting with “This pattern resembles...”

"...since the late 1990s, mothers have been finding it difficult to feed infants with enough clean...." #Is this a factual statement or the author’s personal observation? Is there any documentation that can support this statement?

What is “poverty datum line”? Poverty line defined in the data?

"Furthermore..." #"However..."?

The sentences here are somewhat iterative.

"In addition, ..., can be used to explain the spatial distribution of..." #"can be an explanation for the distribution of..."
These are interesting insights but it would be helpful if there is any documentation that could substantively support them.

At the end of this paragraph, the author may like to add a sentence or two to state that these should be measured in the future investigations.

5.0 Advantages and Limitations of the Study

“…which may be overstepped by the classical regression models…” # Overlooked?

“…‘borrow strength’ from neighbours to cope…” # “borrow strength’ from information of adjacent neighbours”

“…a major limitation of this study is based on the DIC.” I do not find it as a ‘major limitation’, though it is a limitation. Rather the second (the broadness of the disease definitions used) or third limitation (sample size) should be described as more pertinent issues. Because the acronym DIC re-appeared after a long time, it should be spelled out again here.

Tables and Figures

The unit of “mean” needs to be described. Also how do the numbers of observations translate to the population size of each items (Only if it can be figured out. Does DHS come with a weight variable based on their sampling strategy)?

Please see the comments above on Table 1 about the population size.

At the bottom of the table, the author should indicate variables included in each model. However, this table is not necessary.

It would be helpful to indicate which effect is considered as ‘statistically significant’ in a classical sense. Perhaps putting an asterisk on the ones whose credible intervals did not go over 0 will draw the readers’ attentions to those variables.
These are unlikely necessary. If the author decides to keep them, please include at the bottom of the table explanation of what each of the signs mean. Also, perhaps the authors could combine the three tables in one.

<Figures 2-3, 5, 8, 9>

The authors can briefly describe these in the text instead of showing the figures. Too many figures will detract the attention from the main substantive findings of the paper. As far as I see it, for the audience of this journal, this paper should not be a methodology paper, and these figures are about fairly technical aspects of the study, therefore the paper would not particularly benefit from including them.

<Figures-maps>

Three maps for each disease by three time periods should be organized as one set. The units of posterior mean classifications should be articulated. It seems that the cut-off points are different by map. It will make the comparison across periods difficult. The maps should have the same exact cut off points (at least among the same disease) and how the classification was made should be described at the bottom. All maps should have the district names annotated and the year indicated.

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

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.