Reviewer's report

Title: Spatial distribution of the risk of dengue fever in southeast Brazil, 2006-2007.

Version: 1 Date: 24 January 2011

Reviewer: Robin Young

Reviewer's report:

Summary:
The authors use multinomial logistic regression models to determine risk factors for mild and severe dengue fever (each compared to no fever) in southeastern Brazil. The authors also use generalized additive models to determine whether there are differential spatial patterns for mild and severe dengue fever. They find sanitary conditions, including larval density, water storage at home, and frequency of garbage collection among others, to be non-spatial factors increasing the risk of contracting dengue fever. Mild and severe dengue fevers visually determined to have differing spatial patterns. Mild dengue fever occurred more often in the southeastern portion of the study region while severe dengue fever was more often observed in the northern area.

Major compulsory revisions

1) In the Study Design, please give more detail on how the cases and controls were matched. In the sampling design, were the control household randomly selected? What was the desired case-to-control ratio?

2) In the Study Variables section, paragraph 2, the authors describe selecting a random 10% of data to repeat the survey and verify the collected data. How many values were corrected or changed based on this secondary data collection? Please include this information in the manuscript.

3) In the last paragraph of the Discussion section, please clarify the last two sentences. Are the authors discussing “edge effects” that occur in the southeast corner of the map? Please add more description and detail.

4) In the Statistical Analysis section, please specify the type of smoother used with the generalized additive models. Please also specify the smoothing parameter selection method.

In the abstract, methods section, please specify the type of smoother used with the generalized additive model.

5) The authors selected to combine mild and severe cases and examine associations between case/control status and non-spatial variables. Why not use chi-square tests of independence and multinomial regression instead? If the
hypothesis is that mild and severe dengue fevers have different mechanisms and risk factors, should the authors keep the categories of cases separated throughout the analysis?

6) In the Results section, paragraph 5, are the “likelihoods” predicted probabilities of disease? Please clarify what these values represent.

7) In the Conclusions the authors discuss the usefulness of spatial methods; however they make little reference to the results of this study. It may be more informative to make conclusions based on the results of this study and how this study better informs researchers and practitioners about dengue fever.

Minor essential revisions
1) In the abstract, Results paragraph, are the presented odds ratios displaying the effect of severe vs. no dengue fever? Are they from the multinomial model or from the generalized additive models? Please specify the model and whether the odds ratios are from only the severe vs. no dengue fever effect size. If this is the case, please identify variables where similar or different results were observed for mild vs. no dengue fever.

2) In the abstract, the conclusions do not reflect the results that are presented. Please make the conclusions reflecting the material presented in the abstract.

3) The Background, last paragraph, identifies the objectives of the present study. Is another objective the determination of non-spatial predictors? Should this be included?

4) For clarity, please simplify the first paragraph in the Statistical Analysis section. The models can be described as simple and multiple logistic regression models.

5) In Equation 1, what does qi represent? Please specify.

6) What is phi in Equation 6? Is Equation 6 a re-expression of Equation 5? Is it necessary?

7) In the second-to-last paragraph of the Statistical Analysis section, please remove syntax from the manuscript, including “family=multinomial(logit)” and “type=’response’”. Please specify a reference for the VGAM package and for Rv2.7.

8) In general, the Statistical Analysis section can be simplified and the equations can be reduced. Multinomial logistic regression models are relatively straightforward and the detail and equations can be reduced or included in an appendix.

9) In the Results, if the generalized additive model is the model of interest, perhaps the authors could examine how the effect sizes of non-spatial variables differ when accounting for location?

10) In the Discussion, paragraph 4, the authors state that the spatial patterns of
mild and severe dengue fever differ from one another. Was this determined by visual comparison?

11) In the Discussion, Paragraph 1, “crowed” should read “crowded”

12) In Table 1, garbage collection (times per week) has two categories: >2 and #2. Please correct.

13) In Tables 2 and 3, please specify the reference groups for odds ratios.

14) In Figures 2 and 3, what are the risk levels? Are these predicted probabilities of disease? Please clarify.

15) Please make verb tenses consistent throughout the document

16) Please include multiple references at the ends of sentences as [1,2] instead of [1][2]

Discretionary revisions

1) If the authors are interested in identifying differences between risk factors for mild and severe dengue fever, it may be informative to examine predictors of severe dengue fever in comparison to mild fever. The corresponding odds ratios may help to identify differences between the two forms of disease.

2) Can mild dengue fever progress to severe dengue fever? Are the mild and severe cases necessarily different from one another? A further explanation may be warranted.

3) For clarity, the authors may describe the “polytomous regression” as “multinomial logistic regression” models.

**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.