Author's response to reviews

Title: The Spatial Epidemiology and Clinical Features of Reported Cases of La Crosse Virus Infection in West Virginia from 2003 to 2007

Authors:

Andrew D Haddow (adhaddow@gmail.com)
Danae Bixler (Dee.Bixler@wv.gov)
Agricola Odoi (aodoi@utk.edu)

Version: 2 Date: 27 September 2010

Author's response to reviews: see over
To The Editor:

Please accept the revisions for the manuscript entitled “The Spatial Epidemiology and Clinical Features of Reported Cases of La Crosse Virus Infection in West Virginia from 2003 to 2007” by A.D. Haddow, D. Bixler, and A. Odoi for publication in the journal BMC Infectious diseases. We feel we have addressed the reviewers comments and suggestions.

Sincerely,
Andrew Haddow
Danae Bixler
Agricola Odoi

**Reviewer 1:** “In the methods section, page 6 line 116-117, the authors should explicitly describe the difference between confirmed and probable cases and also present how many cases were regarded confirmed and possible. They might also want to evaluate the chance that a patient with meningitis or encephalitis due to causes other than LACV has LACV antibodies due to earlier asymptomatic LACV infection, and thus could be wrongly included in the study.”

**Response:** We have added the CDC case definition to the methods section to clarify the difference between confirmed and probable cases. We have also added how many cases were confirmed and probable. The most likely cause of meningitis or encephalitis would typically be herpes encephalitis, another arbovirus, or a bacterial infection all of which would have been ruled out under the CDC case definition. Additionally, IgM would not have been detectable in the acute phase serum.

**Reviewer 1:** “Also in the methods section, the authors should add how many reported LACV cases during the study period were excluded due to missing data on clinical symptoms and location, or due to the fact that they were not <=15 yrs of age. They should also assess (in the discussion section) the possibility of variation in reporting clinical symptoms and/or location data between different health departments, leading to bias in the analyses results.”

**Response:** We have added information on the number of cases and those that were 15 years and younger in the methods section. We have already noted in the limitations section that: “Clinical data were collected from multiple health care providers by local health department personnel for the purpose of case ascertainment. Due to limited resources, verification of complete and accurate reporting of laboratory and clinical variables was not possible”. We added an additional sentence concerning the limitations of the available clinical data. We have also added a sentence clarifying the method by which location data was collected in the methods section.

**Reviewer 1:** “For spatial analyses of the incidence risk per region (either county or census tract), the authors use spatial empirical Bayesian smoothing, and present maps
with both the unsmoothed and smoothed incidence risks, both showing higher risk in the south of the state. Then the authors use Global Moran’s I test to confirm that there is significant spatial autocorrelation in the data, and the Moran Local Indicators of Spatial Association (LISA) to confirm the location of significant high-risk clusters. The statement of the authors on page 14 line 304-306 that for this LISA analyses they did not correct for multiple comparisons seems controversial. Especially for the census tracts (n=466) it seems likely that statistical significance will occur by coincidence. Although they cite a publication that proposes not to correct for multiple testing, other publications do propose correction. The authors could discuss the effect of adjusting on their analyses results: how many significant clusters are still detected after adjusting for multiple comparisons? Another option could be to consider methods for spatial cluster detection that automatically adjust for multiple testing like the spatial scan statistic published by Kulldorff. An additional benefit of this and possibly other alternative methods is identification of clusters of various sizes, so also clusters that contain several census tracts, instead of one at the time. This prevents that such “multiple-census-tract-clusters” are missed if the census tracts are separately not significant.”

Response: As per the reviewer's suggestion we have replaced the Moran Local cluster analysis with Kulldorff's spatial scan statistic which does not have the limitation of multiple comparisons associated with Local Moran's statistic and additionally allows for detection of clusters of different sizes. We thank the reviewer for his/her insight.

Reviewer 1: “In the results section of the abstract the authors indicate the number of detected high-risk clusters as “n=4” and “n=30” (on county and census level respectively). This is confusing, as some readers might think that “n” indicates the total number of counties or census tracts. The authors should therefore rephrase this.”

Response: We have rephrased this sentence.

Reviewer 1: “Regarding the environmental risk factors the authors acknowledge in the discussion section that they did not perform a case-control study, but only gathered information of (known) risk factors for reported cases. For this reason they should rephrase line 39-41 on page 2 (abstract), as some readers might think that they investigated what the environmental risk factors are for LACV cases, whereas the authors actually describe the results of gathering data on apparently already known environmental risk factors for reported cases.”

Response: We have rephrased this sentence.

Reviewer 1: “In addition to comment 5 above: Page 15, line 323-324 “The presence….LACV”. Some readers might think that this sentence at the end of the article also suggests that this study proofs that wooded areas and containers are a risk factor for LACV, whereas this is not a case-control study, as mentioned before. Please rephrase.”

Response: We have rephrased this sentence.
**Reviewer 1:** “In the discussion section, page 10 line 213-216, the authors mention an increase in incidence, whereas in this paper they describe spatial variation in the incidence of LACV in 2003-2007, and not an increase in time. Please rephrase or clarify.”

**Response:** We have rephrased this sentence.

**Reviewer 1:** “The authors already described in an earlier study (reference 37) that spatial analysis on data with higher spatial resolution (census tract level vs county level) leads to more precise results on the spatial distribution of disease incidence. It would therefore greatly improve the readability of the current paper, if only the census tract level results would be extensively described. The comparison between results on county and census tract level could then very shortly be described in one paragraph, just to confirm that the census tract level is the preferred spatial resolution for these analyses.”

**Response:** This is a new concept in the study of LACV infections and this further illustrates the differences found between different geographic scales as such we prefer to present both spatial scales.

**Reviewer 1:** “Page 4 line 79-80 The authors might want to rephrase this sentence. Probably they intend to say that LACV meningitis and/or encephalitis has been emerging in West Virginia, and not that LACV has been found to be the cause of an emerging disease of formerly unknown cause?”

**Response:** We have rewritten this sentence.

**Reviewer 1:** “Some minor details:
Page 7 line 141, shouldn’t it be “the number of cases per 100,000 persons per YEAR”? 
Page 10 line 221 What is the “block level”? Same as county level? Please clarify. 
Page 11 line 237-238 repeats statement made in line 227-228.”

**Response:** We have added the definition of incidence risk to clarify the difference between an incidence risk and incidence rate. We have rephrased the line 221 for clarity. We have removed the similar sentence.

**Reviewer 2:** “The study is rather interesting, focusing on a very hot topic (i.e., the emergence of La Crosse virus infection in west Virginia). However, the paper has several weak points. First of all, the spatial analysis may be quite difficult to understand for the average reader of the journal and should be better explained. For example, the meaning of terms as census and county tracts may be obscure for many biomedical scientists, especially for those who are from countries other than US.”

**Response:** An explanation of the meanings of ”county” and ”census tracts” has been added to the methods section of the manuscript (under ”Population, geographic and environmental data”). We have also rewritten several sections to improve clarity.
Reviewer 2: “Furthermore, it is not clear to me whether the incidence risk is a measure of association or a frequency measure (actually it seems a sort of incidence density or maybe a risk and not an association measure. Actually the term "incidence" risk is not commonly used in epidemiology, where they use incidence or risk, depending on the method); this should be better explained.”

Response: We have defined incidence risk. Contrary to the reviewer's comments, incidence risk (or cumulative incidence)---which is different from incidence density----is a standard term in epidemiology (see, for instance, pages 75-76 of Veterinary Epidemiologic Research, 2nd Edition, by Ian Dahoo, Wayne Martin and Hendrik Stryhn. Incidence risk is a measure of disease frequency (not a measure of association)---please see pages 73-89 of the above text.

Reviewer 2: “Secondly, it is not clear whether the study is a methodological exercise and whether it is possible to provide a clear take-home message (i.e., geographical variation of La Crosse virus infection can be detected only at the census level or, for example, there wide variation both at the larger - county - and smaller -census - level, etc.).”

Response: This study not only determined those areas at the highest risk for contracting LACV which will be used to guide intervention strategies, it further demonstrated that the use of a smaller spatial scale provides a finer or more appropriate method of detecting spatial clusters i.e. those areas of highest risk than a larger geographic scale, but it also provides the most current data on the clinical presentation of reported cases and contributing environmental factors found at case residences.

Reviewer 2: “Thirdly, clinical data analysis is, to some extent, out of context in this paper, unless signs and symptoms are presented at the beginning of the results to describe the study population.”

Response: This paper aims to give an overview of the “current” epidemiology of LACV infections in West Virginia, and therefore it is not only appropriate to present available clinical data, but also necessary. We feel that the current location of the clinical results within the manuscript is appropriate.

Reviewer 2: “Finally, environmental analysis is very poor. In absence of a control group, no inference can be made from the frequency distribution of environmental variables (i.e., risk factors may be identified exclusively from a case-contro study).”

Response: The goal of presenting the environmental factors was not to investigate for association with LACV infections. Rather, it was purely descriptive to provide the reader with information on observations that were made by the health professionals when they made home visits. We have presented the available data and the study’s results in an appropriate manner and clearly stated the study’s limitations. Additionally, we have removed the word “risk” to avoid any confusion.
Reviewer 2: “Information on clinical and environmental variables was incomplete”

Response: Again, we have presented those available data that could be extracted from data collected by the West Virginia Department of Health and Human Services from 2003-2007. Furthermore, we have presented the results in an appropriate manner and clearly stated the study’s limitations.

Reviewer 2: “Definition of confirmed cases is not reported in the text”

Response: We have added the CDC case definition.

Reviewer 3: “This is a nice paper on an important topic. It was a pleasure to read it. It is well written, its conclusions are well and moderately formulated, its limitation properly stated. It uses appropriate and up-to-date statistical techniques in correct way to reach reasonable interpretations. Spatial maps (both empirical/unsmoothed and empirical Bayes smoothed) presented in the Figure 1 are certainly of substantial practical interest and should be communicated to a broad public audience.”

Response: We appreciate the reviewer’s kind comments.