Reviewer's report

Title: Spatiotemporal analysis of indigenous and imported dengue fever cases in Guangdong province, China

Version: 5 Date: 16 March 2012

Reviewer: Marc Souris

Reviewer's report:

Major compulsory revision

I do not question the quality of the Chinese CDC. By counting the mobile population, we are talking about an incidence of less than 1/100 000 000, which is very low for dengue, given its transmission cycle. A description of the spatial differences in the monitoring system is essential to ensure the validity of the dataset. Furthermore, introduction talk about 650000 dengue cases for 30 years (78-2008) in whole mainland China. Do you think that 1600 cases in Guangdong province for 6 years is realistic, if Guangdong province account for more than 80 % of all cases? Something seems to be wrong.

Of course, I do not question the search of clusters. But a study on space-time cluster detection with a limited number of cases (eg, 6 cases in 2005, 19 in 2009) aggregated in few spatial objects (124) seems not very interesting. With a so low incidence, this search for clusters loses its appeal if done with groups of cases, integrated in geographical objects.

The study suffers from many defects. The fact of aggregating cases in county is nowhere discussed. The fact of cluster detection on the presence / absence by county is nowhere discussed. Secondary clusters with a single object inside (as shown in table) does not really make sense.

Minor essential revisions

Abstract

« so as to identify high-risk areas of the province and thereby help plan resource allocation for dengue interventions »

This study present observed incidence. High-risk need to be replaced by high incidence all over the paper.

« Descriptive spatiotemporal analyses were conducted, including plotting of seasonal distribution of cases » No seasonal mapping was presented. Only incidence by year. Choropleth maps are not a good way to present ratio mapping. A extended litterature present how to map diseases data.

« The space-time scan statistic was used to determine space-time clusters of dengue fever cases ». Only spatial analysis by county was presented, not on cases.
« This study demonstrated that the geographic range of imported and indigenous dengue fever cases has expanded over recent years ». The conclusion is not demonstrated by this study (no expansion analysis).

p6

650000 cases for 30 years (78-2008) in whole mainland China. Do you think that 1600 cases in Guangdong province for 5 years is realistic, if Guangdong province account for more than 80% of all cases? Something seems to be wrong.

« The mean population per county is about 783,000 people (ranging from 78,800 to 7.1 million) ». Such a difference in density between county represent a very strong problem with statistical analysis and spatial analysis. With such a low incidence as presented with dengue dataset, counties with low density have a very high variability compared to counties with higher density.

p8

« The space-time scan statistic was calculated on the imported and indigenous dengue fever cases to test whether the cases were distributed randomly over space and time and, if not, to locate space-time clusters and determine their statistical significance » Cluster detection is different from random distribution analysis. A spatial distribution can be globally randomly distributed and showing significant clusters.

No need to describe the method used by SatScan. References only are needed.

p9

« In the SaTScan software, a default maximum spatial cluster size of 50% of the population was used, so as to detect the large clusters that tend to have a small relative risk but a high statistical significance. Furthermore, a maximum spatial cluster size of 10% of the population was further employed to detect possible subclusters with smaller size ». Need to specify that « population » means number of counties in a cluster...

With low number of reported cases, integration in large zones weeks spatial analysis. A better way is to analyse cases using their address, not incidence in geographic entities by integration. This is my principal concern in term of methodology.

p10

Temporal analysis must be represented for the whole period (2005-2010) in a graph, giving number of reported cases by month.

**Level of interest:** An article of insufficient interest to warrant publication in a
scientific/medical journal

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests