Emerging Health Threats Journal 2011, **4**:s47. **doi:** 10.3134/ehtj.10.047

www.eht-journal.org

ABSTRACT

Non-parametric intensity bounds for the visualization of disease clusters

FLP Oliveira[1,2], LH Duczmal[1], and ALF Cançado[3]

[1]*Statistics Department, Federal University of Minas Gerais, Minas Gerais, Brazil;* [2]*Ibmec, Minas Gerais, Brazil; and* [3]*Statistics Department, University of Brasília, Brasília, Brazil*
E-mail: fernandoluizoliveira@est.ufmg.br

Objective
Given an aggregated-area map with disease cases data, we propose a criterion to measure the plausibility of each area in the map of being part of a possible localized anomaly.

Introduction
Consider the most likely disease cluster produced by any given method, like SaTScan,[1] for the detection and inference of spatial clusters in a map divided into areas; if this cluster is found to be statistically significant, what could be said of the external areas adjacent to the cluster? Do we have enough information to exclude them from a health program of prevention? Do all the areas inside the cluster have the same importance from a practitioner perspective? How to access quantitatively the risk of those regions, given that the information we have (cases count) is also subject to variation in our statistical modeling? A few papers have tackled these questions recently;[2] produces confidence intervals for the risk in every area, which are compared with the risks inside the most likely cluster. There exists a crescent demand of interactive software for the visualization of spatial clusters.[3] A technique was developed[4] to visualize relative risk and statistical significance simultaneously.

Methods
In this work, we assess the problem of finding error bounds for the delineation of spatial clusters in maps of areas with observed populations and number of cases. A given map with the vector of real data (the number of observed cases for each region) shall be considered as one possible realization of the random variable vector with an unknown expected number of cases. Therefore, the process of identification of possible spatial clusters must take into account this source of variation.

In our methodology, we perform m Monte Carlo replications of the vector of random variables for fixed rates given the observed number of cases in each area. Then the most likely cluster for each replicated map is detected and the corresponding likelihood values obtained in the m replications are ranked. For each area, we determine the maximum-likelihood value among the most likely clusters containing that area. Thus, we obtain the intensity function associated to each area's ranking of its respective likelihood value among the m values.

Results
We apply this tool for three different maps for sharply and diffusely delineated clusters. The intensity bounds found by the method reflect the geographic dispersion of the detected clusters, as in Figure 1.

Conclusions
Our technique is able to detect irregularly shaped and multiple clusters, making use of simple tools like the circular scan. Intensity bounds for the delineation of spatial clusters are obtained and indicate the plausibility of each area belonging to the cluster. This tool employs simple mathematical concepts, and interpreting the intensity function is very intuitive in terms of the importance of each region in delineating the possible anomalies of the map of rates.

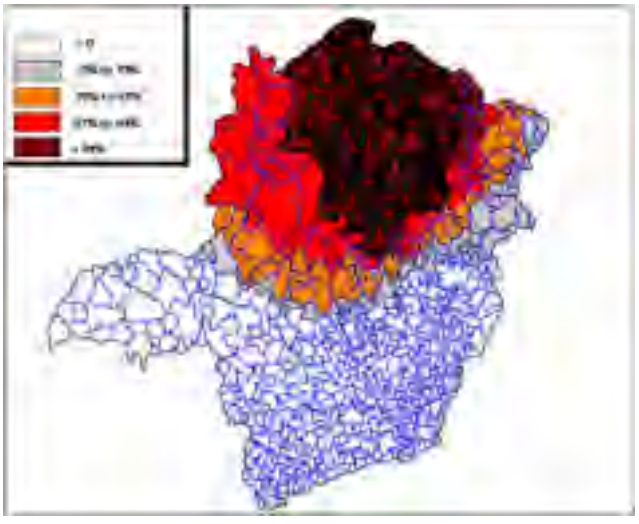

Figure 1 Intensity map quantiles for Chagas' disease data.

The Monte Carlo simulation requires an effort similar to the circular scan algorithm, and, therefore, it is quite fast. We hope that this tool should be useful in public health decision making of which areas should be prioritized.

Acknowledgements

This paper was presented as an oral presentation at the 2010 International Society for Disease Surveillance Conference, held in Park City, UT, USA on 1–2 December 2010.

References

1 Kulldorf M. Information Management Services Inc: SaTScan v7.0: software for the spatial and space-time scan statistics. *Tech Rep* http://www.satscan.org/ 2006.
2 Rosychuk RJ. Identifying geographic areas with high disease rates: when do confidence intervals for rates and a disease cluster detection method agree? *Int J Health Geogr* 2006;**5**:46.
3 Hardisty F, Conley J. Interactive detection of spatial clusters. *Adv Dis Surveillance* 2008;**5**:37.
4 Boscoe FP, McLaughlin C, Schymura MJ, Kielb CL. Visualization of the spatial scan statistic using nested circles. *Health Place* 2003;**9**:273–7.