

# A simulation study of three methods for detecting disease clusters

Geir Aamodt<sup>14</sup>, Sven Ove Samuelsen<sup>24</sup> and Anders Skrondal<sup>34</sup>

<sup>1</sup> Akershus University Hospital

<sup>2</sup> University of Oslo

<sup>3</sup> London School of Economics

<sup>4</sup> Norwegian Institute of Public Health

## Abstract:

Cluster detection is an important part of spatial epidemiology because it can help identifying environmental factors associated with disease and thus guide investigation of the aetiology of diseases. In this article we study three methods suitable for detecting local spatial clusters: (1) a spatial scan statistic (SaTScan) (Kulldorff 1997), (2) generalized additive models (GAM) (Hastie and Tibshirani 1991) and (3) Bayesian disease mapping (BYM) (Besag et al. 1991). We conducted a simulation study to compare the methods. Seven geographic clusters with different shapes were initially chosen as high-risk areas. Different scenarios for the magnitude of the relative risk of these areas as compared to the normal risk areas were considered. For each scenario the performance of the methods were assessed in terms of the sensitivity, specificity, and percentage correctly classified for each cluster.

The performance depends on the relative risk, but all methods are in general suitable for identifying clusters with a relative risk larger than 1.5. However, it is difficult to detect clusters with lower relative risks. The GAM approach had the highest sensitivity, but relatively low specificity leading to an overestimation of the cluster area. Both the BYM and the SaTScan methods work well. Clusters with irregular shapes are more difficult to detect than more circular clusters.

Based on our simulations we conclude that the methods differ in their ability to detect spatial clusters. Different aspects should be considered for appropriate choice of method such as size and shape of the assumed spatial clusters and the relative importance of sensitivity and specificity. In general, the BYM method seems preferable for local cluster detection with relatively high relative risks whereas the SaTScan method appears preferable for lower relative risks. The GAM method needs to be tuned (using cross-validation) to get satisfactory results.

**Keywords:** Spatial epidemiology; disease mapping; Spatial scan statistics; Smoothing techniques.

## 1 Introduction

Different methods have been proposed to locate and identify the clusters dependent on whether the locations of the clusters are suspected or known (focused) or unknown (non-focused). Models for focused clusters are designed for detecting preconceived patterns linked to objects such as power lines or putative sources such as landfill sites (Elliot et al. 2001). Models for non-focused clusters, on the other hand, are designed to estimate the relative risk for each area within the study area. Typically, these models accommodate extra-Poisson variability in different ways (Besag et al. 1991; Clayton and Bernardinelli 2000; Clayton and Kaldor 1987).

In this study we evaluate three different methods with potential for detecting and identifying local cluster patterns for count data, i.e. the number of cases for each area or municipality. The main outcome measure is the estimated relative risk for each municipality or a neighbourhood of municipalities in the study area.

Simulation has been used to evaluate and compare statistical power for different global cluster detection methods (Kulldorff et al. 2003; Song and Kulldorff 2003) as well as for Bayesian methods (Best et al. 2005; Lawson et al. 2000), but little is known concerning their ability to detect different patterns of clusters. The patterns or shapes of disease clusters may vary due to their origin, and it is likely that some methods are more suited to detect specific cluster morphologies than others.

The purpose of this article is to evaluate the performance of three methods for local cluster detection for different types of spatial clusters. The work was motivated by the desire to study the relationship between different disease characteristics, such as the shape of clusters and the risk ratio between high-risk and normal risk areas, and the methods used to detect and depict spatial clusters. The methods differ in the way they detect clusters: SaTScan is based on a likelihood-ratio test to identify areas with increased incidences, whereas BYM and GAM are based on estimation of relative risks. However, all three methods are applicable to detect hot-spot clusters. The ability to detect the different types of clusters is assessed by the sensitivity, specificity and percentage of correct classification.

## 2 Design of the simulation study

Seven different clusters were produced based on the 434 municipalities in Norway: (1) 14 municipalities in a circular pattern (1.1% of the total population) from the middle part of Norway, (2) 6 municipalities along a Norwegian river (1.6% of the population), (3) 15 municipalities (5.2% of the population) along the Norwegian southern coast, (4) 4 separated clusters from 70 municipalities (13.0% of the population), (5) division into an area of high risk in the south (345 municipalities, 89.9% of the population) and

low risk in the north (89 municipalities, 10.1% of the population), (6) a cluster in the southern part of the country (125 municipalities, 32% of the population), and (7) a case of no clusters.

500 simulated datasets were produced for each combination of cluster pattern and magnitude of relative risk. The data were simulated from Poisson distributions. The expected number of cases was 0.2% of the background population, which corresponds to an incidence rate of 200 cases per 100 000 inhabitants. We repeated the simulation procedure with five different values for the relative risk of disease between the high-risk municipalities and the normal risk municipalities. The magnitudes of the relative risks were 1.2, 1.5, 2.5, 4 and 10.

For each combination the sensitivity is estimated as the average percentage of high-risk municipalities that are correctly classified as high-risk municipalities. The specificity is estimated as the average percentage of normal risk municipalities that are correctly classified as normal risk municipalities. We also computed the average percentage of municipalities that are correctly classified as both normal risk and high-risk municipalities.

### 3 Results

The results from our simulations are summarised as sensitivity, specificity, and the percentage correctly classified municipalities for the different cluster types and magnitudes of relative risk. In general, the GAM method shows satisfactory sensitivity but relatively low specificity. For the three small clusters the specificities are around 80-90%, but since the number of municipalities in the clusters is small (6-15), the cluster sizes are overestimated by a factor of two.

The BYM method has higher sensitivity than the SaTScan method for relative risks larger than 2.5 for all cluster types except cluster type 5. SaTScan has higher sensitivity than BYM for relative risks less than 2.5. For a relative risk of 2.5 SaTScan has higher sensitivity than BYM for cluster types 4 and 5), BYM has higher sensitivity than SaTScan for cluster type 3, but no significant differences are observed for cluster types 1, 2, and 6. The mean values of sensitivity change dramatically for relative risks of 1.5 (44%) and 1.2 (18%) as compared to the three higher magnitudes of relative risk (88%, 81% and 74%).

The second cluster, which has a narrow and long shape, is the most difficult cluster to detect for all methods and almost all values of relative risk. The third cluster is, somewhat surprisingly, the one with the highest sensitivity yet high specificity. It is also surprising that the second and third clusters and the first and fourth clusters do not show the same properties in spite of having approximately similar shape. For relative risks larger than 1.5, it is evident that large clusters are more difficult to detect than small clusters, having high specificity but dismal sensitivity. The sixth cluster

includes 89.9% of the population size and exceeds the maximal cluster size for the SaTScan method. Nevertheless, the SaTScan method showed better performance than the BYM model, but worse than the GAM method for this cluster.

We also included a no-cluster scenario. This scenario produces the false alarm rates for the methods. There were no false alarms for the BYM method, but some scattered false alarms for the SaTScan and the GAM methods. The specificity is higher for the BYM method than for the GAM and SaTScan methods for all cluster types and magnitudes of relative risks except for the first cluster type with relative risk equal to 2.5, 4, and 10.

## References

- Besag, J.E., York, J.C., Mollie, A. (1991). Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, **43**, 1-59.
- Best, N., Richardson, S., Thomson, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, **14**: 35-59.
- Clayton, D., Bernardinelli, L. (2000). Bayesian methods for mapping disease risk. In: *Geographical and Environmental Epidemiology: Methods for Small-Area Studies*. Oxford: Oxford University Press.
- Clayton, D., Kaldor, J. (1987). Empirical Bayes Estimates of Age-standardized Relative Risks for Use in Disease Mapping. *Biometrics*, **43**, 671-681.
- Elliot, P., Briggs, D., Morris, S., et al. (2001). Risk of adverse birth outcomes in populations living near landfill sites. *BMJ*, **323**: 363-368.
- Hastie, T.J., Tibshirani, R.J. (1991). *Generalized Additive Models*. New York: Chapman and Hall.
- Kulldorff, M. (1997). A spatial scan statistics. *Communications in Statistics: Theory and Methods*, **26**, 1481-1496.
- Kulldorff, M., Tango, T., Park, P.J. (2003). Power comparisons for disease clustering tests. *Computational Statistics and Data Analysis*, **42**, 665-684.
- Lawson AB, Biggeri AB, Boehning D, et al. (2000). Disease mapping models: an empirical evaluation. Disease Mapping Collaborative Group. *Stat Med*, **19**, 2217-2241.
- Song, C., Kulldorff, M. (2003). Power evaluation of disease clustering tests. *Int J Health Geogr*, **2**, 9.