Published in final edited form as:

Environ Ecol Stat. 2010 September ; 17(3): 303–316. doi:10.1007/s10651-009-0108-1.

Power Evaluation of Focused Cluster Tests

RC Puett^{1,2}, AB Lawson³, AB Clark⁴, JR Hebert¹, and M Kulldorff⁵

¹South Carolina Statewide Cancer Prevention and Control Program, University of South Carolina, USA

²Departments of Epidemiology and Biostatistics and Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, USA

³Department of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina, USA

³School of Medicine, Health Policy and Practice, University of East Anglia, United Kingdom

⁴Department of Ambulatory Care and Prevention, Harvard Medical School and Harvard Pilgrim Health Care, USA

Abstract

Many statistical tests have been developed to assess the significance of clusters of disease located around known sources of environmental contaminants, also known as focused disease clusters. The majority of focused-cluster tests were designed to detect a particular spatial pattern of clustering, one in which the disease cluster centers around the pollution source and declines in a radial fashion with distance. However, other spatial patterns of environmentally related disease clusters are likely given that the spatial dispersion patterns of environmental contaminants, and thus human exposure, depend on a number of factors (i.e., meteorology and topography). For this study, data were simulated with five different spatial patterns of disease clusters, reflecting potential pollutant dispersion scenarios: 1) a radial effect decreasing with increasing distance, 2) a radial effect with a defined peak and decreasing with distance, 3) a simple angular effect, 4) an angular effect decreasing with increasing distance and 5) an angular effect with a defined peak and decreasing with distance. The power to detect each type of spatially distributed disease cluster was evaluated using Stone's Maximum Likelihood Ratio Test, Tango's Focused Test, Bithell's Linear Risk Score Test, and variations of the Lawson-Waller Score Test. Study findings underscore the importance of considering environmental contaminant dispersion patterns, particularly directional effects, with respect to focused-cluster test selection in cluster investigations. The effect of extra variation in risk also is considered, although its effect is not substantial in terms of the power of tests.

Keywords

clusters; power; small area analysis; spatial statistics

1. Introduction

A range of statistical tests have been developed to detect focused clustering (Lawson, 2006, ch 6, Kulldorff, 2006), which can be defined as the clustering of disease around suspected or known sources of environmental contaminants (Besag and Newell, 1991). To date, limited emphasis has been placed on the selection of a focused-cluster test appropriate to the spatial pattern of environmental exposure. It has been suggested that spatial functions, such as rankings based on inverse distance, can be included in certain focused-cluster tests as proxies for exposure when exact measurements are unknown (Bithell, 1995). However, most focused-cluster tests have been designed and used to detect clusters reflecting a particular spatial pattern, one that centers around the pollution source and declines with increasing distance from the source. In accordance with pollution dispersion equations and models that have been used in various fields such as the study of air pollution, Lawson, (1993) and Lawson and Williams (1994) have suggested that including similar spatial functions could improve the ability to detect populations at elevated risk from exposure to pollution sources. This recommendation is based on the premise that human exposure to environmental contaminants from a known source and consequent disease clusters follow the spatial patterns of environmental contaminant dispersion. Lawson (1993) describes the use of mathematical functions to describe peaked effects, directional effects, the combination of these effects, and the more typical radial distance decline cluster pattern. A study of respiratory cancers in Armadale, Scotland demonstrates the use of these functions in models of bronchitis mortality, and suggests the importance of attention to the directional dispersion of air pollution in focused-cluster modeling (Lawson and Williams, 1994).

Published power evaluations of focused-cluster tests are limited, and interest has focused mainly on score tests; i.e., Stone's Maximum Likelihood Ratio Test (Stone, 1988) and Stone's Poisson Maximum Test (Stone, 1988). Waller and Lawson (1995) examined the power of Stone's Poisson Maximum Test (Stone, 1988), the focused adaptation of Besag and Newell's Test (Besag and Newell, 1991), and Lawson-Waller score tests (Lawson, 1993; Waller et al., 1992) to detect clusters of varying sizes with a range of background disease rates and increases in disease risk. The authors found that the score tests performed well overall, while the Besag and Newell Test improved with rarer diseases. More pronounced clustering and elevated disease rates were necessary for Stone's test to show power equal to the score test. Waller (1996) extended work in this area by evaluating power as related to changes in aggregation levels. His findings revealed that the level of data aggregation is an important consideration with regard to statistical power and that many of the evaluated tests showed poor power overall to detect a small increase in risk. Waller and Poquette (1999) examined power of the Lawson-Waller score tests (Lawson, 1993; Waller et al., 1992) and score statistics proposed by Bithell (1995) with three distance-based exposure values. The authors showed that power of the evaluated tests depends on the shape of the cluster, the shape the test assumes for the cluster and the population density of the geographic area of interest. Tango (Tango, 2002) extended the Lawson-Waller Score Test (Lawson, 1993; Waller et al., 1992) to deal with multiple point sources and peak decline trends. He compares the power of the extended score test against Stone's Maximum Likelihood Ratio Test (Stone, 1988) and Bithell's Likelihood Ratio Test (Bithell, 1995) for detecting clinal

and hot spot clusters. Sun (2002) compares the power and other statistical criteria of Tango's Focused Test (Tango, 1995) and Stone's Poisson Maximum Test (Stone, 1988) through data simulations using a Geographic Information System (GIS). Sun found that Tango's test was more effective than Stone's test based on an evaluation system of accuracy, type I error, power, and consistency. However the author suggests the use of both methods as Stone's test tends to produce lower estimates of cluster size while Tango's tends to produce higher estimates. Sun also reports that Tango's test is preferred with regard to estimating cluster size as long as the clusters are not too weak or too strong. In a comparison of Stone's Poisson Maximum and Maximum Likelihood Ratio Tests (Stone, 1988), Lumley (1995) shows that the Maximum Likelihood Ratio Test has advantages with regard to power.

Though previous evaluations have provided important information that should be considered when selecting a test to detect focused clustering, investigations are lacking regarding the effects of cluster shape on the power of focused cluster tests. The purpose of the current study is to evaluate the power of a variety of focused-cluster tests to detect five different spatial cluster patterns around a fixed focus, which reflect different potential pollution dispersion scenarios. For cluster investigations and other small area analyses, data are most commonly available as counts; therefore this power evaluation study considers only count data. The organization of this paper is as follows. Data simulation methods are discussed in Section 2. Section 3 defines the focused cluster tests evaluated. The results of the power evaluation are presented in Section 4, with conclusions presented in Section 5.

2. Data Simulation

Five spatial patterns of disease clustering were considered under the alternative hypothesis. Figures 1 through 3 provide sample illustrations of some of these spatial patterns and relative risk ranges. However, it should be noted that the simulation is Monte Carlo based; and therefore, ranges of relative risk could vary between simulated realizations. The most complex model represents an anisotropic (asymmetrical) pattern of disease risk where the peak in disease risk occurs at some distance from the source rather than at the pollution source. This peaked distance directional decline model (PDDIR) is defined as:

$$\theta_i = 1 + \exp\{(\alpha_3 \cos(\varphi_i)) + (\alpha_4 \sin(\varphi_i)) + (\alpha_2 \log(d_i)) - (\alpha_1 d_i)\} \quad (1)$$

where $\alpha_1 = 2$, $\alpha_2 = 0.05$ and $\alpha_3 = \alpha_4 = 0.5$. The variable ϕ_i is the angle of the small area centroid measured from the putative source, and $\mu = \pi / 4$ is the assumed mean angle. d_i is the distance from the pollution source to the centroid of region i . The other models are special cases of the PDDIR model in which various parameters equal zero. Table 1 details these models, including an anisotropic pattern or distance decline combined with a directional effect (DDIR). A pattern with increasing disease risk in a particular direction also was considered (DIR). PDD was an isotropic pattern where the peak in disease risk occurs at some distance from the source. The final and simplest model was an isotropic pattern, evincing radial distance decline in disease risk (DD). This spatial distribution of distance decay is most commonly assumed in focused cluster testing.

Let y_i , $i=1, \dots, M$, denote the count of disease in the i th region and $y = (y_1, \dots, y_M)$, further let x_i denote the coordinates of the i th region. In what follows we assume that the basic

model for the counts is $y_i \sim \text{Poisson}(e_i\theta_i)$ where an expected count e_i is modified by a relative risk θ_i . For each simulation, we hold the total number of events fixed (N), since this allows us to focus on the spatial pattern and not the total number of events. Hence we simulate from a multinomial distribution, where the probability of a case of disease in the i

th region is given by $P_i = \frac{e_i\theta_i}{\sum_i e_i\theta_i}$. We choose to simulate 200 events, 500 events and 1,000 events or total cases of disease. The pattern of expected counts was constant for the map, i.e. $e_i = e = N/M$.

We simulated data for five types of spatially distributed clusters representing different alternative hypotheses. We also simulated data under the null hypothesis in the absence of a disease cluster, where $\theta_i = 1$ for all i . Two hundred fifty-six regions of uniform size and shape were simulated. Centroids for each region were placed on a 16×16 unit square grid on coordinates: $0, 1/15, 2/15, \dots, 1$ on the vertical and horizontal axes of the square. The pollution source was located centrally at $x_0 = (0.5, 0.5)$. The counts of events were simulated for each of these regions and assigned to each regional centroid.

The choice of model for the relative risk at location x , $\theta(x, \beta)$ represents the relationship between the pollution source and spatial distribution of disease, for some choice of parameters β . Under the null hypothesis $\theta_i = 1$ for all i . Under the alternative hypothesis θ_i varies by i , where the relative risk in the i th region, θ_i , was determined using functions describing each of the five spatial patterns of disease clustering with the null value used for no effect.

In addition to calculating θ for each of these five spatial patterns of clustering for each of the three N 's, two additional θ 's were estimated for each spatial pattern and total event combination in order to examine the effect of uncorrelated heterogeneity (UH). UH can be defined as extra-variation that is not directly observed but is assumed to exist in the data (Lawson, 2006, ch 5). Sophisticated Bayesian models (Best et al., 2005; Leyland and Davies, 2005) assume that UH exists, therefore UH is also being considered in the models for this study. These models take the form:

$$\theta(x_i, \beta) = \theta^*(x, \beta) \cdot \exp(v_i),$$

where $\theta^*(x, \beta)$ is one of the standard models for the risk of disease in the i th region which is modified by the UH term, v_i . For example $\theta^*(x_i, \beta) = 1 + \exp(-\alpha d_i)$ in the DD model. The term v_i which represents the UH for the i th region is assumed to be normally distributed with zero mean and constant variance $= \sigma^2$. Two variations of UH were added to the models examined here: $\sigma = 0.05$ and 0.5 , representing different levels of UH.

To summarize the overall simulation procedure for data under the alternative hypothesis: five spatial cluster types (DD, PDD, DIR, DDIR, and PDDIR), three variations of UH (none, $\sigma = 0.05$, $\sigma = 0.5$), and three variations of total events ($N = 200, 500, 1000$) were considered. In addition to the simulations conducted to obtain datasets under the alternative

hypothesis, three simulations, one for each total event situation, were performed to obtain data under the null hypothesis of no disease clustering with no UH. Six additional simulations provided data under the null hypothesis including two levels of UH ($\sigma = 0.5$ and $\sigma = 0.05$) for the three variations of total events.

The power of the focused cluster tests examined in the current study was evaluated through computer simulations. This involves computing test statistics for 1,000 data sets simulated under the null hypothesis to obtain the critical value or 95th percentile of these values. Test statistics are then computed from the 1,000 data sets simulated under each alternative hypothesis (i.e., each of the five spatial cluster types). These test statistics are compared to the critical value to estimate power, which is the number of times the null hypothesis is rejected.

3. Focused Cluster Tests

Power was evaluated for Stone's Maximum Likelihood Ratio Test (Stone, 1988), Tango's Focused Test (Tango, 1995), variations of the Lawson-Waller Score Test (Lawson, 1993; Waller et al., 1992), and Bithell's Linear Risk Score (LRS) Test (Bithell, 1995). The decline tests (Stone's Test, Tango's Test, the Radial Score Test and the LRS Distance Decline Test) and the decline and directional effect test (the Directional Score Test) used *a priori* knowledge of the location of the pollution source.

A. Decline Tests

Stone's Maximum Likelihood Ratio Test is based on the isotonic regression estimator (Stone, 1988). It is computed by first binning the regions using bins defined by distance from the source. Stone's Maximum Likelihood Ratio Test is then given by

$$S = \frac{L(\hat{\theta})}{L(\theta_0)}$$

where $\hat{\theta}_i = \min_{s \leq i} \max_{t \geq i} \frac{\sum_{r=s}^t y_r}{t-s+1}$ $i = 1, \dots, K$ is the vector of maximum likelihood estimates under the alternative hypothesis of decreasing risk with increasing distance from the cluster center. θ_0 is the relative risk under the null hypothesis of constant risk, and K is the maximum number of regions.

Tango's Focused Clustering Test for rare diseases (Tango, 1995) (hereafter referred to as Tango's Test) was evaluated with the formulation:

$$T = N^{-1} \sum_{i=1}^K (y_i - e_i) \exp(-d_i/\tau)$$

We selected $\tau = 1, 5$ to represent sharp and slow decreases with distance. These values should be selected *a priori* based on information regarding the dispersion of the exposure leading to the cluster or otherwise expected shape of the cluster of elevated risk..

The Radial Decline Score Test variation of the Lawson-Waller Score Tests as presented by Lawson (1993) was examined:

$$W = \frac{\left\{ \sum_{i=1}^M d_i (y_i - e_i) \right\}^2}{\sum_{i=1}^M d_i^2 e_i - \frac{\left(\sum_{i=1}^M d_i e_i \right)^2}{N}}.$$

Bithell (1995) describes the LRS Tests in detail, using various functions of distance and rank as surrogates for exposure. We applied the distance decline function as described by Lawson (1993) to the formulation Bithell illustrates resulting in an LRS Test for distance decline. Bithell (1995) defined:

$$T = \sum_{i=1}^M y_i \log \theta_{1i},$$

where θ_{1i} is the area-specific relative risk based on the alternative hypothesis, and M is the maximum number of regions. We therefore defined the test statistic:

$$T = \sum y_i \log(1 + \exp(-\alpha_1 d_i))$$

with the parameter α_1 estimated under the alternative hypothesis.

B. Decline and Directional Effect Test

A directional variation of the Lawson-Waller Score Test was examined, as defined by Lawson (1993):

$$W = \frac{\left\{ \sum_{i=1}^M \cos(\phi_i - \mu) (y_i - e_i) \right\}^2}{\sum_{i=1}^M e_i \cos^2(\phi_i - \mu) - \frac{\left\{ \sum_{i=1}^M e_i \cos(\phi_i - \mu) \right\}^2}{\sum_{i=1}^M e_i}}$$

where $\phi - \mu$ is again the angle between the pollution source, regional centroid and mean angle. The mean angle μ is a nuisance parameter and is estimated under the null hypothesis

of no clustering. In simulations, the mean angle is arbitrary and set to $\pi/4$ for convenience. In practice, this angle should be selected based on pollution dispersion principles (e.g., direction of the dominant wind blowing releases from a stack). When considering disease clusters related to environmental exposures, the principles of modeling the dispersion of these environmental exposures should be considered, and the science of environmental toxicant dispersion modeling provides a sound basis for selecting this angle in practice.

4. Focused Cluster Test Results

Tables 2 through 6 present the estimated power of each of the focused cluster tests for detecting the five spatial cluster patterns, simulated with three total numbers of events and three variations of UH. These tables show the results of the computer simulation evaluations to compare the null hypothesis to each of the alternative hypotheses, or spatial cluster patterns. The tables demonstrate the low to modest power achieved overall for detecting each type of spatial cluster; however, these results are related to the choice of relative risk level. As expected, power increases with increasing total numbers of events, and the addition of UH lowers power. The findings of note are revealed by comparing the power achieved among the various tests for detecting the range of spatial clusters.

The decline tests show the highest power for detecting DD spatial cluster shapes (Table 2). Comparison among the tests shows that Tango's Test and the LRS Distance Decline test perform best overall. The decline and directional effect test shows low power across all total event scenarios, and the addition of UH lowered power for all tests.

Table 3 shows low power overall to detect the PDD shaped cluster with the decline tests outperforming the decline and directional effect test, as expected. Among the decline tests, Tango's Test and the LRS Distance Decline Test revealed the highest power overall. Again, the addition of UH decreased the power to detect this type of cluster, and the tests that showed the highest power to detect the cluster also showed proportionally the least loss of power with UH.

Results for the DIR spatial cluster pattern differ from those for the other four spatial cluster patterns (Table 4). The decline and directional effect test was effectively the only test to detect this type of cluster and did so with high power, even with the addition of UH. The decline tests showed poor power overall for detecting the DIR spatial cluster pattern.

The Directional Score Test showed the highest power overall for detecting the DDIR spatial pattern of clustering (Table 5). Power for the decline and directional effect test was equal to that for detecting the PDDIR cluster with 1000 total events and slightly better than that for the PDDIR cluster with 200 and 500 total events. The decline tests also showed high power for detecting the DDIR cluster with 1000 events, with Tango's Test and the LRS Distance Decline Test once again showing the highest power among the decline tests.

Table 6 shows the results of comparing the alternative hypothesis of the PDDIR cluster pattern to the null hypothesis of no clustering. With regard to power, the Directional Score Test was the overall best-performing test, though only modest power was achieved. The

Decline Tests performed almost as well as the Directional Score Test, with the LRS Distance Decline Test and Tango's Test showing slightly higher power than Stone's Maximum Likelihood Ratio Test and the Radial Score Test. The addition of UH lowered the power of all tests to detect the PDDIR cluster pattern. However it is interesting to note that the Directional Score test showed the least loss of power with UH, whereas Stone's Maximum Likelihood Ratio Test and the Radial Score test showed the greatest.

5. Conclusions

The use of pollutant dispersal modeling provides information regarding likely spatial patterns of human exposure, and thus disease. Findings from this investigation reveal the importance of considering this information when selecting a focused cluster test with adequate power to detect a cluster of disease. Power to detect cluster patterns with a directional effect only (DIR), as could occur with an air pollution source and a strong dominant wind to disperse the contaminant, was poor for all focused cluster tests except the Directional Score Test. Among the cluster types examined, the Directional Score Test also appeared to be the best choice overall with respect to power for detecting any cluster type with a directional component. For example, the Directional Score Test would be recommended if a peak in risk is suspected at some distance from the pollution source combined with an anisotropic pattern of risk. This spatial clustering pattern might follow the pollution dispersal of an airborne contaminant from an elevated stack accompanied by a dominant wind. However, it should be noted that the version of the Directional Score Test

examined in this analysis used a prespecified parameter $\pi/4$ for convenience to indicate the angle between the pollution source and regional centroid. The mean angle μ in this analysis was a nuisance parameter and was set to an arbitrary value. Additional variation in this parameter could possibly result in decreased power. As expected, decline tests, particularly Tango's Test and the LRS Distance Decline test, are preferred among the focused tests examined for detecting cluster types with an isotropic pattern of risk, PDD and DD cluster types. Cluster patterns with a peak were more difficult to detect than similarly shaped cluster patterns without peaks. It should be noted that the score tests performed well in each scenario for which the alternative was defined according to the dispersion pattern. Therefore we project that in cases where pollution dispersion patterns can be predicted based on exposure modeling scenarios that account for appropriate variables (e.g. wind dispersion for air pollutants or landcover for groundwater toxicants), power should be improved by selecting a test, possibly defining a score test, which best reflects the appropriate dispersion pattern.

This investigation underscored the difficulty detecting clusters of smaller sizes/rarer diseases as well as the challenges induced with addition of UH. With the exception of the Directional Score Test results for the DIR cluster type, all tests revealed poor power detecting clusters of 200 total events, regardless of the spatial cluster pattern. Similarly, modest power was shown by the best-performing tests in each spatial pattern for detecting 500 total events. This is important to note as many of the conditions that we are asked to analyze for clustering are rare (as is the case for virtually all cancers). The addition of two levels of UH, $\sigma = 0.05$ and 0.5 , lowered the power of all tests for detecting all types of clusters. The

Directional Score Test in cases of an anisotropic pattern of risk as well as Tango's Test and the LRS Distance Decline Test in cases with a DD cluster component handled the addition of UH better than the other tests considered.

Though a number of focused cluster tests and spatial patterns of clustering were examined in this simulation study, additional work is needed to inform the selection of appropriate focused cluster tests. Studies of additional variations in the shape and scale of the cluster patterns in this investigation as well as further types of spatial patterns of clustering should be considered. For example, simulations of clustering in more linear shapes could inform the study of the clustering of health conditions due to traffic exposure. It should be noted that a homogeneous underlying population distribution was simulated, with each grid cell containing the same number of expected events. A more heterogeneous distribution of expected cases could influence the power of focused cluster tests, and should be examined with further simulation studies. Comparisons of these findings with application of these same focused cluster tests to real world data, pollutant dispersal as well as health outcomes, would provide valuable information for cluster investigators.

Acknowledgments

This research was funded by the National Cancer Institute Grant number R01CA095979.

References

- Besag J, Newell J. The Detection of Clusters in Rare Diseases. *Journal of the Royal Statistical Society Series A-Statistics in Society*. 1991; 154:143–155.
- Best N, Richardson S, Thomas A. A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*. 2005; 14:35–59. [PubMed: 15690999]
- Bithell JF. The choice of test for detecting raised disease risk near a point source. *Statistics in Medicine*. 1995; 14:2309–2322. [PubMed: 8711271]
- Kulldorff M. Tests for spatial randomness adjusting for an underlying inhomogeneity: A general framework. *Journal of the American Statistical Association*. 2006; 101:1289–1305.
- Lawson AB. On the analysis of mortality events associated with a prespecified fixed point. *Journal of the Royal Statistical Society Series A-Statistics in Society*. 1993; 156:363–377.
- Lawson, AB. *Statistical Methods in Spatial Epidemiology*. 2nd Ed. New York: John Wiley & Sons, Ltd; 2006.
- Lawson AB, Williams FLR. Armadale - a Case-Study in Environmental Epidemiology. *Journal of the Royal Statistical Society Series A-Statistics in Society*. 1994; 157:285–298.
- Leyland A, Davies C. Empirical Bayes Methods for Disease Mapping. *Statistical Methods in Medical research*. 2005; 14:17–34. [PubMed: 15690998]
- Lumley T. Efficient execution of Stone's likelihood ratio tests for disease clustering. *Computational Statistics and Data Analysis*. 1995; 20:499–510.
- Stone RA. Investigations of excess environmental risks around putative sources: statistical problems and a proposed test. *Statistics in Medicine*. 1988; 7:649–660. [PubMed: 3406597]
- Sun Y. Determining the size of spatial clusters in focused tests: Comparing two methods by means of simulation in a GIS. *Journal of Geographic Systems*. 2002; 4:359–370.
- Tango T. A class of tests for detecting 'general' and 'focused' clustering of rare diseases. *Statistics in Medicine*. 1995; 14:2323–2334. [PubMed: 8711272]
- Tango T. Score tests for detecting excess risks around putative sources. *Statistics in Medicine*. 2002; 21:497–514. [PubMed: 11836732]

- Waller LA. Statistical power and design of focused clustering studies. *Stat Med.* 1996; 15:765–782. [PubMed: 9132904]
- Waller LA, Lawson AB. The power of focused tests to detect disease clustering. *Statistics in Medicine.* 1995; 14:2291–2308. [PubMed: 8711270]
- Waller, LA.; Poquette, CA. The power of focused score tests under misspecified cluster models. In: Lawson, A.; Biggeri, A.; Bohning, D.; Lesaffre, E.; Viel, J-F.; Bertollini, R., editors. *Disease Mapping and Risk Assessment for Public Health*. Chichester: John Wiley and Sons; 1999. p. 257-269.
- Waller LA, Turnbull BW, Clark LC, Nasca P. Chronic disease surveillance and testing of clustering of disease and exposure: Application to leukemia incidence and TCE-contaminated dumpsites in upstate New York. *Environmetrics.* 1992; 3:281–300.

Biographies

Dr. Robin Puett is a Research Assistant Professor with the SC Statewide Cancer Prevention and Control Program and with the Departments of Environmental Health Sciences and Epidemiology and Biostatistics at the Arnold School of Public Health, University of South Carolina. She received a dual Ph.D. in Environmental Health Sciences and Epidemiology from the Arnold School of Public Health in 2004. Her research interests are in environmental and spatial epidemiology.

Dr. Andrew Lawson is a Professor in the Department of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina. He received an MPhil from the University of Leeds in 1987 and a PhD from the University of St Andrews in 1991. He has written/edited 8 books in the area of spatial epidemiology and clustering. His research interests are in epidemiology and spatial statistics.

Dr Allan Clark is a lecturer in Medical Statistics at the University of East Anglia. He received his bachelor's degree in 1996 from the University of Abertay, Scotland, and his PhD in 2002 from the University of Aberdeen, Scotland. His research interests are in the development of statistical methods for spatial epidemiology and, more widely, in medical statistics.

Dr. James R. Hebert is Health Sciences Distinguished Professor of Epidemiology and Founding Director of the Statewide Cancer Prevention and Control Program at the University of South Carolina. He received an MSPH in Environmental Health and Epidemiology from the University of Washington in 1980 and an Sc.D. in Nutritional Epidemiology from Harvard University in 1984. He has published over 230 articles in peer-reviewed scientific and medical journals of which >175 focus on cancer or nutrition or both.

Dr. Martin Kulldorff is associate professor in the Department of Ambulatory Care and Prevention at Harvard Medical School and Harvard Pilgrim Health Care. He studied statistics at Umeå University, graduating in 1984, and received his PhD in operations research from Cornell University in 1989. His main research interest is the development and application of statistical methods for disease surveillance.

Figure 1.
DIR Pattern of disease clustering around a central pollution source simulated under the alternative hypothesis

Figure 2.
PDD Pattern of disease clustering around a central pollution source simulated under the alternative hypothesis

Figure 3.
DD Pattern of disease clustering around a central pollution source simulated under the alternative hypothesis

Table 1

α parameters of the PDDIR equation for four more simple spatial patterns of disease clustering

	Parameters Set to Zero			
	α_1	α_2	α_3	α_4
PDDIR				
DDIR		X		
DIR	X	X		
PDD			X	X
DD		X	X	X

Table 2

Estimated power of tests at $\alpha = 0.05$ level from 1000 trials of simulation to detect clusters with a PDDIR cluster type*

	UH [†] →	200 Total Events			500 Total Events			1000 Total Events		
		0	0.05	0.5	0	0.05	0.5	0	0.05	0.5
Decline Tests	Stone's Maximum Likelihood Ratio Test	0.317	0.209	0.297	0.604	0.308	0.445	0.850	0.381	0.560
	Tango's Test $\tau = 1$	0.387	0.323	0.334	0.701	0.523	0.505	0.929	0.768	0.680
	Tango's Test $\tau = 5$	0.372	0.324	0.336	0.689	0.540	0.535	0.931	0.768	0.683
	LRS Distance Decline Test	0.383	0.327	0.337	0.695	0.529	0.521	0.933	0.767	0.678
	Radial Score Test	0.368	0.321	0.333	0.594	0.323	0.354	0.836	0.460	0.463
Decline and Directional Effect Tests	Directional Score Test	0.469	0.414	0.430	0.887	0.724	0.694	0.997	0.946	0.907

* tabulated values are statistical power; i.e., $(1-\beta) \times 100$

[†]UH = uncorrelated heterogeneity

Table 3

Estimated power of tests at $\alpha = 0.05$ level from 1000 trials of simulation to detect clusters with a DDIR cluster type*

	UH [†] →	200 Total Events			500 Total Events			1000 Total Events		
		0	0.05	0.5	0	0.05	0.5	0	0.05	0.5
Decline Tests	Stone's Maximum Likelihood Ratio Test	0.393	0.276	0.337	0.705	0.422	0.526	0.935	0.549	0.634
	Tango's Test $\tau = 1$	0.457	0.392	0.401	0.797	0.637	0.609	0.972	0.874	0.760
	Tango's Test $\tau = 5$	0.435	0.386	0.390	0.788	0.658	0.629	0.974	0.869	0.758
	LRS Distance Decline Test	0.446	0.390	0.396	0.797	0.648	0.616	0.973	0.872	0.759
	Radial Score Test	0.434	0.384	0.393	0.687	0.442	0.434	0.923	0.616	0.562
Decline and Directional Effect Tests	Directional Score Test	0.513	0.460	0.460	0.921	0.761	0.722	0.997	0.966	0.902

* tabulated values are statistical power; i.e., $(1-\beta) \times 100$

[†]UH = uncorrelated heterogeneity

Table 4
Estimated power of tests at $\alpha = 0.05$ level from 1000 trials of simulation to detect clusters with a DIR cluster type *

	UH [†] →	200 Total Events			500 Total Events			1000 Total Events		
		0	0.05	0.5	0	0.05	0.5	0	0.05	0.5
Decline Tests	Stone's Maximum Likelihood Ratio Test	0.048	0.023	0.067	0.062	0.011	0.059	0.059	0.002	0.062
	Tango's Test $\tau = 1$	0.045	0.030	0.047	0.046	0.020	0.050	0.061	0.009	0.064
	Tango's Test $\tau = 5$	0.043	0.030	0.047	0.044	0.025	0.057	0.064	0.009	0.065
	LRS Distance Decline Test	0.042	0.031	0.048	0.045	0.021	0.053	0.066	0.009	0.065
	Radial Score Test	0.043	0.029	0.048	0.060	0.015	0.054	0.057	0.002	0.053
Decline and Directional Effect Tests	Directional Score Test	0.891	0.848	0.863	1.000	0.997	0.981	1.000	1.000	0.999

* tabulated values are statistical power; i.e., $(1-\beta) \times 100$
[†]UH = uncorrelated heterogeneity

Table 5

Estimated power of tests at $\alpha = 0.05$ level from 1000 trials of simulation to detect clusters with a PDD cluster type*

	UH [†] →	200 Total Events			500 Total Events			1000 Total Events		
		0	0.05	0.5	0	0.05	0.5	0	0.05	0.5
Decline Tests	Stone's Maximum Likelihood Ratio Test	0.283	0.205	0.270	0.528	0.265	0.375	0.807	0.296	0.476
	Tango's Test $\tau = 1$	0.344	0.305	0.305	0.646	0.458	0.455	0.888	0.699	0.613
	Tango's Test $\tau = 5$	0.334	0.304	0.307	0.635	0.479	0.480	0.887	0.700	0.617
	LRS Distance Decline Test	0.338	0.309	0.305	0.645	0.469	0.470	0.891	0.700	0.612
	Radial Score Test	0.326	0.302	0.312	0.512	0.272	0.299	0.767	0.351	0.402
Decline and Directional Effect Tests	Directional Score Test	0.061	0.032	0.059	0.045	0.009	0.054	0.048	0.004	0.060

* tabulated values are statistical power; i.e., $(1-\beta) \times 100$

[†]UH = uncorrelated heterogeneity

Table 6

Estimated power of tests at $\alpha = 0.05$ level from 1000 trials of simulation to detect clusters with a DD cluster type *

	$UH^{\dagger} \rightarrow$	200 Total Events			500 Total Events			1000 Total Events		
		0	0.05	0.5	0	0.05	0.5	0	0.05	0.5
Decline Tests	Stone's Maximum Likelihood Ratio Test	0.334	0.179	0.154	0.633	0.355	0.464	0.884	0.431	0.568
	Tango's Test $\tau = 1$	0.420	0.246	0.233	0.740	0.550	0.549	0.947	0.810	0.719
	Tango's Test $\tau = 5$	0.397	0.242	0.235	0.726	0.573	0.574	0.950	0.805	0.711
	LRS Distance Decline Test	0.406	0.242	0.242	0.736	0.561	0.560	0.949	0.806	0.713
	Radial Score Test	0.393	0.239	0.231	0.621	0.363	0.382	0.877	0.523	0.494
Decline and Directional Effect Tests	Directional Score Test	0.053	0.037	0.042	0.047	0.011	0.056	0.045	0.006	0.061

* tabulated values are statistical power; i.e., $(1-\beta) * 100$

† UH = uncorrelated heterogeneity