STATISTICS IN MEDICINE
Statist. Med. 2001; **20**:2977–2987 (DOI: 10.1002/sim.917)

A weighted average likelihood ratio test for spatial clustering of disease

Ronald E. Gangnon[1,*,†] and Murray K. Clayton[2]

[1]*Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, 1300 University Avenue, 6720 Medical Sciences Center, Madison, Wisconsin 53706, U.S.A.*
[2]*Department of Statistics, University of Wisconsin-Madison, 1210 W. Dayton Street, Madison, Wisconsin 53706, U.S.A.*

SUMMARY

We consider methods proposed for detecting localized spatial clustering. We propose a new test statistic, the weighted average likelihood ratio test, as an alternative to the spatial scan (maximum likelihood ratio) test statistic. Two different types of weights are considered. We propose an unbiased cluster selection criterion and evaluate the bias of the tests through simulation. We also examine the power of the tests through simulations and apply the methods to the well-known New York leukaemia data.

1. INTRODUCTION

Disease clustering studies are typically approached as hypothesis testing problems. The null hypothesis of no clustering, that is, a common rate of disease across the study region, is tested against an alternative hypothesis of clustering. Characteristics of the clustering hypothesis have been used to categorize testing procedures. Besag and Newell [1] distinguish between tests for focused clustering and tests for general clustering. Kulldorff [2] further distinguishes, among general clustering methods, between tests for global clustering and tests for localized clusters. For our purposes, we divide clustering methods into three categories: (general) tests for global clustering; focused tests for localized clusters, and general tests for localized clusters.

In a test for global clustering, clustering occurs when cases are closer to other cases than cases are to non-cases. This global clustering can be detected with statistics that measure the average distance between cases [3–6]. In a test for localized clusters, clustering is defined as an elevated rate of disease in a small portion of the study area, which is then called the cluster. For a focused test, the location of a potential cluster is prespecified, models for the clustering process can be developed, and cluster risks can be estimated [7–9]. For a general

*Correspondence to: Ronald E. Gangnon, Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, 1300 University Avenue, 6720 Medical Sciences Center, Madison, Wisconsin 53706, U.S.A.
†E-mail: ronald@biostat.wisc.edu

Received January 2000
Accepted November 2000

test, the location of the potential cluster or clusters cannot be specified in advance, and the goal of the study is to determine whether the disease rate is elevated in one of a large number of potential clusters. In the remainder of this paper we will discuss only general methods for detecting localized clusters.

Openshaw *et al.* [10] proposed the geographical analysis machine (GAM) as an exploratory cluster detection method. With GAM, one considers numerous overlapping circles of various radii as potential clusters. If a circle contains at least two cases and an unusually high rate (nominal p-value $\leqslant 0.002$), its circumference is drawn on a map of the study region. To control the overall error rate, Openshaw *et al.* suggest using a test based on the number of nominally significant (that is, drawn) circles.

Turnbull *et al.* [11] and Besag and Newell [1] proposed alternatives to the GAM based on 'circles' of fixed 'population radius' and 'circles' of fixed 'case radius', respectively. These 'circles' are formed by aggregating adjacent cells. These methods are designed to be more tractable versions of the GAM, since the circles are, in some sense, directly comparable. Significance levels are assessed by permuting the case assignments. Since the correct radius is unknown, one applies either procedure using several different radii.

Kulldorff and Nagarwalla [12] generalize the previous procedures to allow for an arbitrary collection of clusters. Using their approach, one avoids artificial constructs such as population or case radii and still controls the overall type I error rate. For every potential cluster, one calculates the likelihood ratio statistic for testing the null hypothesis of a common rate in all cells against the alternative of two rates – one rate for cells outside the cluster and a second higher rate for cells inside the cluster. The potential cluster associated with the largest likelihood ratio is the estimated cluster. Significance levels can again be evaluated through simulation.

In this paper we propose a new general testing procedure for localized clusters. Our approach is similar in spirit to that taken by Kulldroff and Nagarwalla [12], but attempts to use the information about potential clusters more fully. In Section 2 we propose an average likelihood ratio test for cluster detection. In Section 3 we use simulation results to evaluate the bias and power of the average likelihood ratio test and the maximum likelihood ratio test of Kulldorff and Nagarwalla [12]. In Section 4 we apply our new testing procedure to the well-known New York leukaemia data set. In Section 5 we provide some brief concluding remarks.

2. WEIGHTED AVERAGE LIKELIHOOD RATIO TEST

We begin by defining some terminology and a basic statistical model. We consider situations in which the study region is divided into N subregions, or cells. For each cell i, we observe O_i, the number of cases of disease, and n_i, the population at risk in cell i. We assume a Poisson model for the data, that is, $O_i \sim \text{Poisson}(\rho_i n_i)$, where ρ_i is the disease rate in cell i. This Poisson model allows for the inclusion of covariate effects by replacing n_i with the expected case count E_i and letting ρ_i denote the relative risk in cell i (a situation we will not consider here).

We are interested in testing the null hypothesis of a constant rate of disease across all cells, that is, $H_0{:}\rho_1 = \rho_2 = \cdots = \rho_N$. For our purposes, the alternative hypothesis is defined in terms of a collection of potential clusters. To make this discussion concrete, we will discuss

a specific set of potential clusters; none the less, the methodology can easily be adapted to arbitrary collections of potential clusters.

In particular, we consider circular clusters centred at the cell centroids as potential clusters. The radius of the circles varies continuously from zero up to a fixed maximum radius, $r_{\max}$. If the centroid of a cell falls within the circle, then the whole cell is included in the cluster. Since there are only a finite number of cells, there will only be a finite number of clusters about each cell centroid. To identify these clusters, let $0 = r_{i,(1)} < r_{i,(2)} < \cdots < r_{i,(m_i)} \leqslant r_{\max}$ be the ordered distances from the centroid of cell i to the centroids of all cells, truncated at $r_{\max}$. (If two or more centroids are equidistant from the centroid i, the common distance is only listed once.) Then, the distinct potential clusters about cell i are circles of radii $r_{i,(1)}, r_{i,(2)}, \ldots, r_{i,(m_i)}$. We refer to the cluster centred at the centroid of cell i of radius $r_{i,(j)}$ as cluster i,j for $j = 1, 2, \ldots, m_i$ and $i = 1, 2, \ldots, N$. Kulldorff and Nagarwalla [12] define a similar collection of clusters, but instead of using a single maximum radius, they define a separate maximum radius for each centroid so that the largest circle includes no more than 50 per cent of the population. In practice, the choice of a particular set of potential clusters (population-based or geographically-based, the choice of maximum radius, etc.) should be driven by the application.

For potential cluster i,j, we follow Kulldorff and Nagarwalla [12] and consider the two-parameter model $O_i \sim \text{Poisson}(\lambda_1 n_i)$ if cell i belongs to cluster i,j and $O_i \sim \text{Poisson}(\lambda_0 n_i)$ otherwise. We then calculate the ratio of the likelihood under the alternative hypothesis $\lambda_1 \neq \lambda_0$ to the likelihood under the null hypothesis $\lambda_1 = \lambda_0$. The likelihood ratio for cluster i,j, denoted by $\text{LR}_{i,j}$, equals

$$\left(\frac{O_{i,j}/n_{i,j}}{O_t/n_t}\right)^{O_{i,j}} \left(\frac{(O_t - O_{i,j})/(n_t - n_{i,j})}{O_t/n_t}\right)^{O_t - O_{i,j}}$$

where $O_{i,j}$ ($n_{i,j}$) is the case (respectively, population) count inside cluster i,j and O_t (n_t) is the total case (respectively, population) count in the study region.

For a general test for localized clusters, Kulldorff [13] suggested using the maximum likelihood ratio over all potential clusters, $\text{LR}_{\max} = \max_{i,j} \text{LR}_{i,j}$, as a test statistic. The distribution of this statistic under the null hypothesis of no clustering can be found through Monte Carlo simulation. If the null hypothesis is rejected, the cluster associated with $\text{LR}_{\max}$ (effectively, the maximum likelihood estimate of the true cluster) may be viewed as the estimated cluster location. This MLE cluster provides only an approximate cluster location, because many clusters overlapping the MLE cluster will also be associated with large likelihood ratios.

The previous observation suggests that the $\text{LR}_{\max}$ test is not using all the information about localized clustering available in the data and hence may be inefficient. As an alternative, we propose using the weighted average of the likelihood ratios, $\text{LR}_{\text{wgt}} = \sum_{i=1}^{N} \sum_{j=1}^{m_i} w_{i,j} \text{LR}_{i,j}$, where $w_{i,j} \geqslant 0$ is a known weight associated with cluster i,j. Discussion of specific choices for the weights will be postponed; the present discussion is quite general and does not depend on the chosen weights. As with the $\text{LR}_{\max}$ test, in our applications the null distribution of the LR_{wgt} test is found by Monte Carlo simulation.

The LR_{wgt} statistic has an appealing Bayesian interpretation as approximating the marginal likelihood ratio of a one cluster model to the no clustering model. To see this, note that the Bayesian information criterion (BIC) approximation [14] to the marginal likelihood for cluster model i,j is proportional to the maximized likelihood of cluster model i,j, assuming

flat priors for the disease rates. Thus, if we view the $w_{i,j}$ as a prior distribution on the clusters, the LR_{wgt} statistic is proportional to the approximate marginal likelihood of the one cluster model. In this sense, LR_{wgt} is a reasonable measure of the evidence for a single cluster model in the data.

We more fully exploit our Bayesian interpretation of the weights to interpret the data if the null hypothesis is rejected. Specifically, after rejection, we assume that the one cluster model is correct and formally treat the $w_{i,j}$ as a prior distribution on the clusters. Conditional on the one cluster model, the posterior probability of the true cluster being cluster i,j is approximately $w_{i,j}LR_{i,j}/LR_{wgt}$. Using these probabilities, we can average over the single cluster models to obtain estimates of the probability that a cell belongs to the cluster and the cell-specific disease rates. Such statistics can pinpoint the location of a single cluster or suggest the presence of multiple clusters. If there are indications of multiple clusters, more sophisticated modelling techniques may be employed [15].

To this point, we have taken the weights $w_{i,j}$ as given. We now discuss two possible sets of weights. Both sets of weights are motivated by schemes for selecting a cluster at random. For both schemes, we first select a cluster centre and then, conditional on that centre, select a cluster radius. The first selection scheme consists of selecting a point from a uniform distribution over the study area and making the centroid of the cell to which the point belongs the cluster centre. The radius is then selected at random from a uniform distribution on $[0, r_{max}]$. Then the weight $w_{i,j}$ is defined to be the probability that cluster i,j is selected by this procedure, and we denote the average likelihood ratio test statistic using these weights as $LR_{wgt,1}$. We note that

$$w_{i,j} = \frac{a_i}{A}\frac{r_{i,j+1} - r_{i,j}}{r_{max}}$$

where a_i is the area of cell i, A is the area of the study region, and $r_{i,m_i+1} = r_{max}$. Similar schemes could be developed for other sets of potential clusters, including the population-based circles used by Kulldorff and Nagarwalla [12].

The second set of weights is based on a crude version of the above scheme that requires less information about the cells. First, a cell centroid is chosen at random, with probability $1/N$, to be the cluster centre. Second, a radius $r_{i,j}$ is chosen at random, with probability $1/m_i$, from $\{r_{i,1}, r_{i,2}, \ldots, r_{i,m_i}\}$. Thus, the probability of selecting cluster i,j under this scheme is $1/(Nm_i)$; the average likelihood ratio test using these probabilities as weights is denoted $LR_{wgt,2}$.

3. SIMULATION RESULTS: POWER AND BIAS

To assess the merits of the tests based on $LR_{wgt,1}$, $LR_{wgt,2}$ and LR_{max}, we consider two criteria: power and unbiased cluster selection. Since all three testing procedures maintain exact control of the type I error rate, comparisons based on the error rates under the null hypothesis are uninteresting. None the less, there is an important property of the tests under the null hypothesis that can be compared. Associated with each test is an 'estimated cluster' (for LR_{max}, the most likely cluster; for the weighted average LR tests, the posterior over the clusters). If, under the null hypothesis, each cell in the study region has an equal chance of belonging to the 'estimated cluster', we say the test for localized clustering is unbiased. This

is an intuitive notion of a fair test, similar to the idea of unbiased split selection for tree-based models [16].

To evaluate the unbiasedness of these tests, we utilized the structure of the New York leukaemia data. The New York leukaemia data set consists of data on leukaemia incidence between 1978 and 1982 in eight counties in upstate New York: Broome; Cayuga; Chenango; Cortland; Madison; Onondaga; Tioga, and Tompkins. The two largest cities in the study region are Syracuse in Onondaga County and Binghamton in Broome County. The eight-county region is divided into 790 cells, either census blocks or census tracts. For each cell, the population at risk, count of leukaemia cases and geographic centroid are available. In our analysis, cell areas are imputed using the Dirichlet tessellation of the centroids [15]. Additional background information on the New York leukaemia data is available in the paper by Waller *et al.* [17].

We simulated 1000 data sets under the null hypothesis, using a rate of 5 cases per 10 000 persons. This rate is comparable to the overall rate in the New York data. For these analyses, we considered all circles centred at the cell centroids with radii less than or equal to 20 km as potential clusters. For each simulated data set, we found the most likely cluster and the posterior distributions associated with the two weighted averages. For each cell, we then found the proportion of the simulations in which the cell belonged to the most likely cluster and the average posterior probability that the cell belonged to a cluster under the two weighting schemes.

Figure 1 displays these estimates of cluster membership probabilities on a map of the New York data. For an unbiased test, this graphic will show a uniform dark colour. Greater colour contrasts between light and dark colours indicates greater bias in the test, with dark areas being relatively likely to belong to the estimated cluster and light areas being relatively unlikely to belong. The results are quite striking. $LR_{wgt,1}$ shows minimal bias with estimates ranging from 0.014 to 0.037. Both LR_{max} and $LR_{wgt,2}$ show evidence of bias with $LR_{wgt,2}$ showing more severe bias; estimates range from 0.001 to 0.077 for LR_{max} and from 0.003 to 0.306 for $LR_{wgt,2}$. (Applying LR_{max} with a fixed maximum population size of clusters instead of a fixed geographic size of clusters does not substantially alter the bias; estimates range from 0.001 to 0.068.) Both of these tests appear to favour clusters in the Syracuse area over clusters in rural areas. The likely cause of this bias is the large number of geographically small cells in these urban areas. These cells produce a lot of potential clusters, many of which significantly overlap. Neither LR_{max} nor $LR_{wgt,2}$ downweights these clusters to account for the overlap, and hence there is substantial bias favouring an urban cluster. We note that, on a regular grid of cells, the bias would likely be minimal; however, such cases are unlikely to occur in practice. All methods, including LR_{max} with a fixed maximum population size of clusters, show evidence of edge effects in that cells at the edge of the study region are less likely to belong to a cluster than interior cells. Again, $LR_{wgt,1}$ appears to minimize this bias although it does not eliminate it.

We next examined the power of the three tests to detect clusters in simulated data sets. For each of 15 sets of simulations, we introduced a single circular cluster centred at a cell centroid into the data set. The rate outside the cluster was 5 cases per 10 000 population. A rate ratio (RR) was assigned to the cluster commensurate with the population inside the cluster, that is, clusters with large (respectively, small) populations had relatively low (respectively, high) RRs. Since the power of the likelihood ratio test depends primarily on the expected number of cases, that is, on the product of the population and the rate, this method of assigning

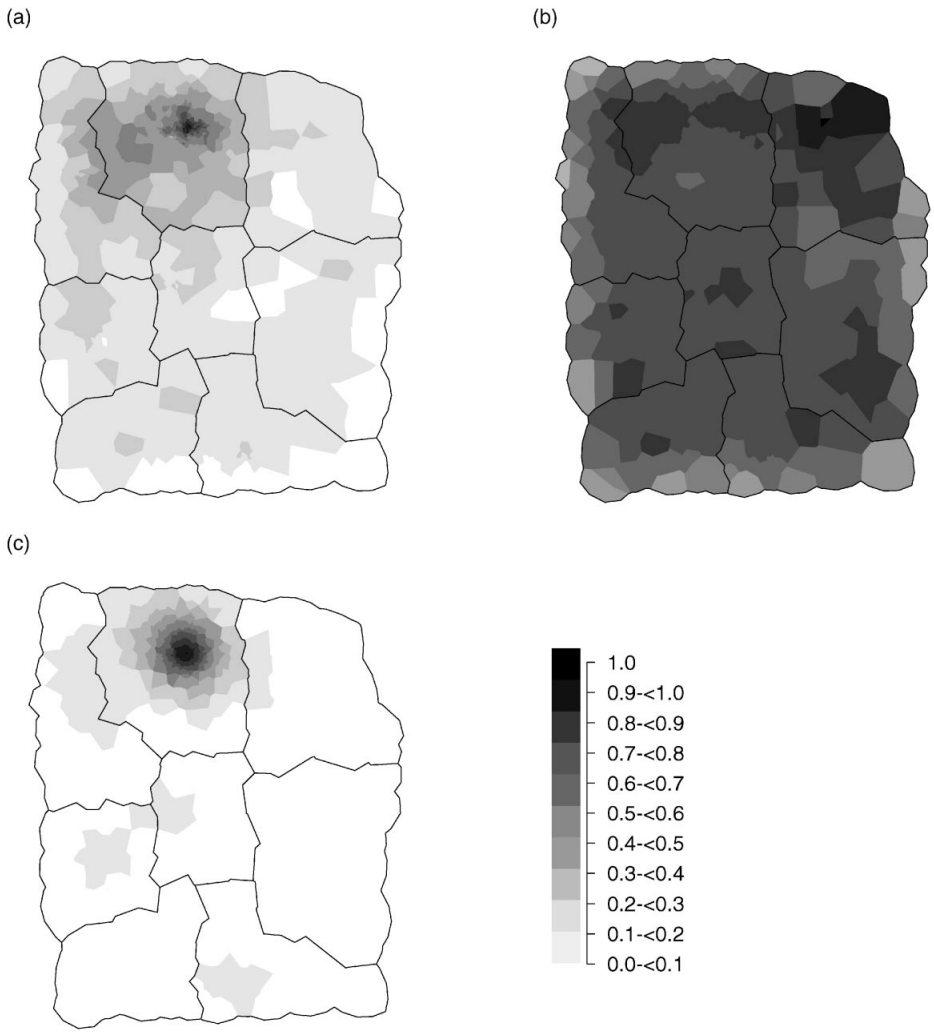

Figure 1. Estimated bias of proposed test statistics based on 1000 simulations under the null hypothesis. (*a*) Proportion of simulations in which cell belonged to the most likely cluster divided by the maximum such proportion (0.077). (*b*) Average estimated probability that cell belonged to a cluster using $LR_{wgt,1}$ divided by the maximum such probability (0.037). (*c*) Average estimated probability that cell belonged to a cluster using $LR_{wgt,2}$ divided by the maximum such probability (0.306).

RRs was designed to ensure reasonable (not exceptionally high or low) power to detect each cluster. The specific rate ratios used (1.4, 1.5, 2, 3, 4 and 6) were chosen arbitrarily and achieved our desired goal of reasonable power. The 15 clusters used for this power study are displayed on a map of the New York data in Figure 2. Estimated rejection rates based on 1000 simulations are provided in Table I. Note that, for each simulated data set, the 5 per cent critical value for each test statistic was found using 100 simulations under the conditional null distribution.

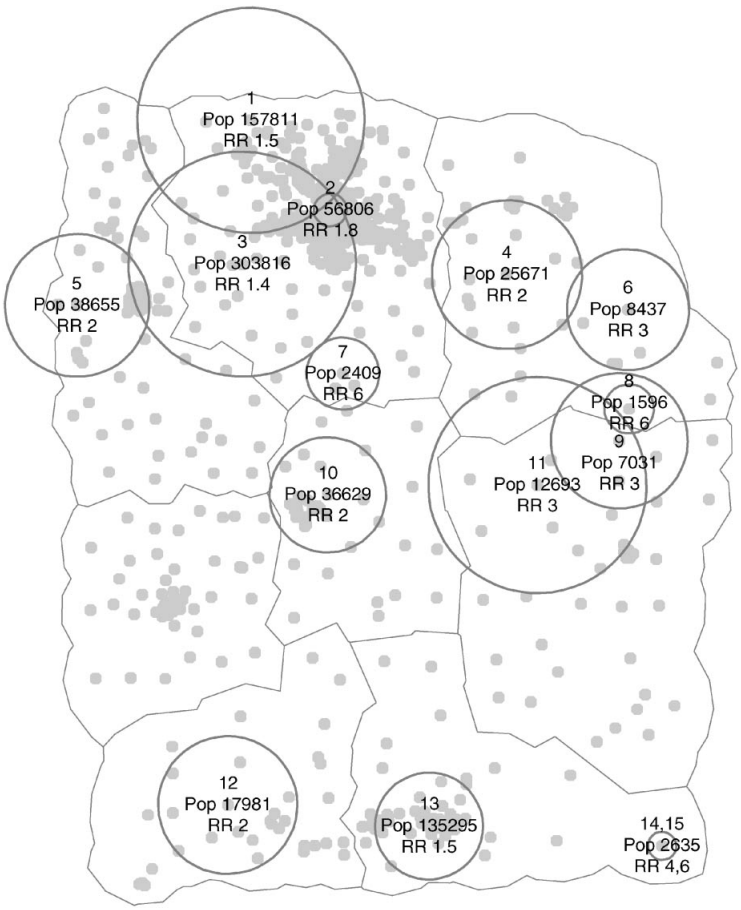

Figure 2. Map of the centroids of cells in New York data. Grey dots indicate cell centroids. Circles indicate clusters used for power simulations. Text indicates cluster number, population and rate ratio (RR).

To summarize the results, the $LR_{wgt,1}$ test provides competitive power in most settings and often the best power of the three test statistics considered here. Major exceptions are clusters 1–3, which overlap portions of Syracuse. For cluster 2, the power of $LR_{wgt,1}$ is only 29 per cent versus 66 per cent for $LR_{wgt,2}$ and 59 per cent for LR_{max}. These findings are not especially surprising in light of the inherent bias of the latter tests towards clusters in areas with many cells, which are typically urban areas. In addition, we note that, with the one exception given above, both weighted average likelihood ratio tests are always at least as good as the maximum likelihood ratio test. Thus, for general usage, we would recommend the $LR_{wgt,1}$ test. If clustering in dense, urban areas were of special concern, we would suggest using either the $LR_{wgt,2}$ test or, preferably, a tailored set of weights designed for the specific application.

Table I. Estimated power of the $LR_{wgt,1}$, $LR_{wgt,2}$ and LR_{max} tests to detect clusters based on 1000 simulations for each cluster. Cluster information provided includes the co-ordinates of cluster centre (km from centroid of region), its radius and population, and rate ratio used for simulations. Bold text indicates the test with largest power for each cluster.

Number	Cluster information					Power of tests		
	X	*Y*	Radius	Population	RR	$LR_{wgt,1}$	$LR_{wgt,2}$	LR_{max}
1	−27.17	56.19	19.83	157,811	1.5	53.3	**66.8**	53.7
2	−13.32	40.41	2.89	56,806	1.8	29.2	**65.8**	59.0
3	−28.69	30.85	19.84	303,816	1.4	54.3	**69.9**	56.7
4	17.48	29.01	13.11	25,671	2.0	**33.2**	25.8	23.6
5	−57.47	23.60	12.57	38,655	2.0	61.9	**63.0**	49.5
6	38.58	22.83	10.65	8,437	3.0	**48.4**	36.0	30.5
7	−11.13	11.58	6.37	2,409	6.0	**51.6**	37.6	40.1
8	38.79	5.35	4.34	1,596	6.0	**32.6**	19.4	19.3
9	37.09	−0.26	11.93	7,031	3.0	**37.2**	22.3	20.9
10	22.79	−8.07	19.04	12,693	3.0	**60.3**	52.8	50.8
11	−13.77	−9.79	10.16	36,629	2.0	**66.8**	64.0	49.6
12	−31.19	−64.46	12.13	17,981	2.0	**27.2**	21.0	18.5
13	3.89	−68.17	9.41	135,295	1.5	55.0	**60.0**	44.0
14	44.54	−71.66	2.49	2,635	4.0	**31.0**	13.7	13.3
15	44.54	−71.66	2.49	2,635	6.0	**65.5**	37.9	35.4

Given the two-sided nature of our testing procedure, a concern might arise that some rejections of the null found in this study could reflect detection of false negative clusters instead of the true positive clusters. An additional set of 1000 simulations using model 11 found that, regardless of statistical significance, the true cluster was identified as the estimated cluster in all but one of those simulations.

4. EXAMPLE: NEW YORK LEUKAEMIA DATA

To assess the evidence for localized clustering in the New York leukaemia data, we considered all circles centred at the cell centroids with radii less than or equal to 20 km as potential clusters. We tested the null hypothesis with the three test statistics discussed earlier – LR_{max}, $LR_{wgt,1}$ and $LR_{wgt,2}$. We conducted each test at the 1 per cent significance level, rejecting the null hypothesis if the observed test statistic was larger than the 99th percentile of 1000 values simulated under the null.

The observed values of the test statistics, on the log scale, were 13.06 for LR_{max}, 4.96 for $LR_{wgt,1}$, and 5.81 for $LR_{wgt,2}$. All three test statistics were clearly significant at the 1 per cent level. (The simulated 99th percentiles of the test statistics were 11.55, 3.59 and 2.73, respectively.) The estimated clusters using each test statistic are displayed in Figure 3 along with a greyscale map of the observed five-year leukaemia rates.

Our analysis using LR_{max} indicates a significant high rate cluster in Broome County, as did the analysis of Kulldorff and Nagarwalla [12]. Our analysis differs from theirs in two aspects: first, we considered a slightly different class of clusters, and second, we tested a two-sided alternative (high or low cluster rates) instead of the one-sided alternative (high cluster

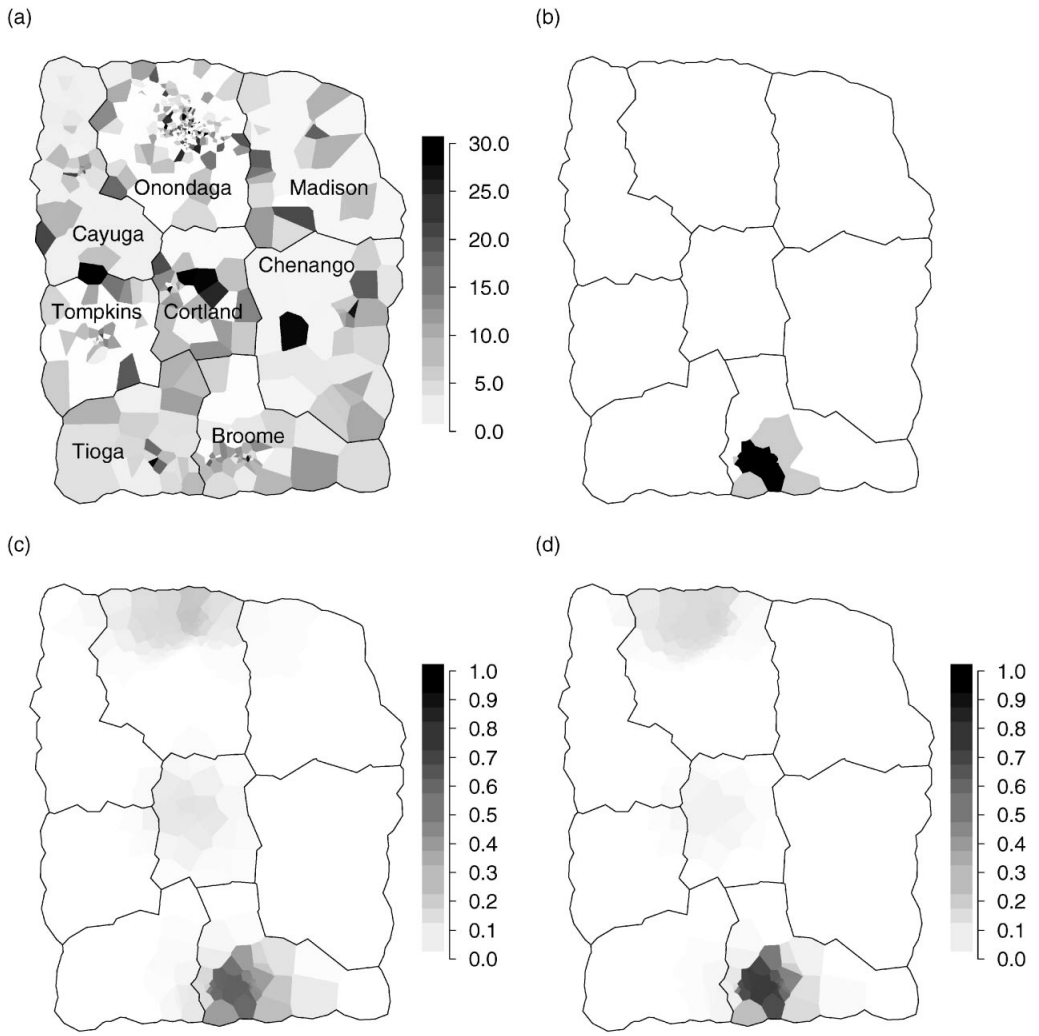

Figure 3. (*a*) Greyscale map of the observed five-year leukaemia rates per 10000 persons for the New York leukaemia data. Rates above 30 per 10000 displayed in uniform black. (*b*) Cells belonging to most likely cluster, that is, the cluster associated with LR_{max}, (drawn in black) and cells belonging to secondary significant clusters (drawn in grey). (*c*) Posterior probabilities of cells belonging to a single cluster using $LR_{wgt,1}$ weights. (*d*) Posterior probabilities of cells belonging to a single cluster using $LR_{wgt,2}$ weights.

rates). On the map, cells belonging to the circle corresponding to the most likely cluster are drawn in black; cells belonging to circles corresponding to other clusters whose likelihood ratios are 'significant', that is, greater than 11.55, are drawn in grey. With this analysis, we found statistically significant evidence for a single cluster in Broome County. If we used a significance level of 0.10 (comparable to the 0.05 used by Kulldorff and Nagarwalla [12] for

the one-sided alternative), we would also find a second 'statistically significant' cluster in Cortland County.

The Bayesian estimates of cluster membership probabilities associated with the $LR_{wgt,1}$ and $LR_{wgt,2}$ statistics are quite similar. Both indicate much evidence for a cluster in Broome County; the 'posterior probability' of that cluster is 62 per cent using $LR_{wgt,1}$ and 77 per cent using $LR_{wgt,2}$. There is additional evidence for two clusters – a cluster in Onondaga County (probabilities 23 per cent and 16 per cent, respectively) and a cluster in Cortland County (probabilities 11 per cent and 5 per cent, respectively). The clusters in Broome and Cortland counties appear to have higher rates than the background, while the cluster in Onondaga County appears to have a lower rate than the background. The differences in the probability estimates result from the differences in weights for urban and rural clusters under the two schemes, because the area of clustering in Broome County is relatively urban (it overlaps the city of Binghamton) and the areas of clustering in Cortland and Onondaga counties are relatively rural. The posterior probabilities must be interpreted in the context of the assumed single cluster model. The probabilities assess the relative likelihood of a single cluster being in Broome County or being in Cortland County. They do not assess the likelihood of there being two clusters, one in Broome County and one in Cortland County. The posterior probabilities do suggest the possibility of multiple clusters, although they cannot provide a formal assessment of that possibility. Follow-up using methods suitable for detecting multiple clusters would be the next step in a complete analysis of these data.

5. CONCLUSIONS

For a global test for localized clusters, the proposed weighted average likelihood ratio test provides an attractive alternative to the maximum likelihood ratio test. The proposed test statistic has a natural interpretation as the marginal likelihood under a simple Bayesian model. It also appears to have superior power and, with the proper choice of weights, can be relatively unbiased. Estimated cluster membership probabilities (found using the Bayesian interpretation of the test statistic) quantify the possible extent of the cluster in ways that a single 'approximate cluster location' cannot.

The test statistics discussed here, both the maximum likelihood ratio test and the weighted average likelihood ratio test, should be principally viewed as surveillance tools. A significant test result will require additional epidemiologic investigation. When there are indications of multiple clusters, we would also suggest further analysis using estimation methods suitable for multiple clusters, such as the Bayesian approach of Gangnon and Clayton [15].

REFERENCES

1. Besag J, Newell J. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society*, *Series A* 1991; **154**:143–155.
2. Kulldorff M. Statistical methods for spatial epidemiology: tests for randomness. In *GIS and Health for Europe*, Gatrell A, Løytønen M (eds). Taylor & Francis: London 1998; 49–62.
3. Whittemore A, Friend N, Brown BW, Holly EA. A test to detect clusters of disease. *Biometrika* 1987; **74**:31–35.
4. Ross A, Davis S. Point pattern analysis of the spatial proximity of residences prior to diagnosis of persons with Hodgkin's disease. *American Journal of Epidemiology* 1990; **132**:S53–S62.
5. Selvin S, Schulman J, Merrill D. Distance and risk measures for the analysis of spatial data: a study of childhood cancers. *Social Science and Medicine* 1992; **34**:769–777.

6. Cuzick J, Edwards R. Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society*, *Series B* 1990; **52**:73–104.
7. Stone RA. Investigations of excess environmental risks around putative sources: statistical problems and a proposed test. *Statistics in Medicine* 1988; **7**:649–660.
8. Diggle PJ. A point process modelling approach to raised incidence of a rare phenomenon in the vicinity of a prespecified point. *Journal of the Royal Statistical Society*, *Series A* 1990; **153**:349–362.
9. 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.
10. Openshaw S, Craft AW, Charlton M, Birch JM. Investigation of leukaemia clusters by use of a geographical analysis machine. *Lancet* 1988; **1**:272–273.
11. Turnbull BW, Iwano EJ, Burnett WS, Howe HL, Clark LC. Monitoring for clusters of disease: application to leukemia incidence in upstate New York. *American Journal of Epidemiology* 1990; **132**:S136–S143.
12. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Statistics in Medicine* 1995; **14**:799–810.
13. Kulldorff M. A spatial scan statistic. *Communications in Statistics – Theory and Methods* 1997; **26**:1487–1496.
14. Schwarz G. Estimating the dimension of a model. *Annals of Statistics* 1978; **6**:461–464.
15. Gangnon RE, Clayton MK. Bayesian detection and modeling of spatial disease clustering. *Biometrics* 2000; **56**:922–935.
16. Loh W-Y, Shih Y-S. Split selection methods for classification trees. *Statistica Sinica* 1997; **7**:815–840.
17. Waller LA, Turnbull BW, Clark LC, Nasca P. Spatial pattern analyses to detect rare disease clusters. In *Case Studies in Biometry*, Lange N, Ryan L, Billard L (eds). Wiley: New York, 1994; 3–22.

