2003

A Method for Testing Low-Value Spatial Clustering for Rare Diseases

Ge Lin
glin@wvu.edu

Tonglin Zhang

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Acta Tropica 91 (2004) 279–289
ACTA TROPICA
www.elsevier.com/locate/actatropica

A method for testing low-value spatial clustering for rare diseases

Ge Lin [a,*], Tonglin Zhang [b]

[a] West Virginia University Department of Geology & Geography, 425 White Hall, P.O. Box 6300, Morgantown, WV 26506, USA
[b] Department of Statistics, 150 North University Street, 512 West Lafayette, IN 47907-2067, USA

Available online 15 June 2004

Abstract

This paper proposes a method that tests for the existence of low-value spatial clustering while accounting for the influence of high-value clustering. Although the method was developed in reference to the Tango test, it can be extended to other testing methods. The simulation results showed that the proposed method is able to effectively detect low-value clustering with substantially lower rates of type I errors than those of the Tango test, while maintaining comparable statistical power. Applying the method in a case study of leukemia in Minnesota demonstrated an overall tendency toward low-value clustering of leukemia mortality for males but provided inconclusive results for females.

Keywords: Bias; Low-value clustering; Relative risk; Trimmed mean

1. Introduction

Detecting general spatial clustering, explicitly determining the types of spatial cluster(s), and eventually locating them have been widely applied in epidemiology (Lawson and Denison, 2000). According to Marshall, spatial clusters are foci of particularly high incidence or hot spots that are unlikely to happen by chance (Marshall, 1991). Marshall also pointed out that low-value foci or cool spots should be included in this definition. In spatial epidemiology, when a cluster exists, there is an overall tendency toward clustering, which is normally a prerequisite for further study (Cuzick and Edwards, 1990). Many spatial epidemiologists have stressed the importance of identifying and quantifying spatial clustering of elevated risks, as they often provide the basis for the allocation of medical and health resources and etiological clues for disease treatment and prevention (Elliot et al., 2000). Nevertheless, there are several reasons why low-value clustering should also be identified and these issues examined from the opposite perspective.

First, low-value clustering or cool spots are indicative of healthy communities. After detecting low-value clustering, we can further investigate what makes people in a particular clustered area less likely to have a certain disease. The lesser likelihood of disease might result from environmental factors in the geographic area that enhance immunity to a specific disease or from genetically endowed resistance among people in the community. Many genetic studies indicate that some ethnic groups are genetically marked as being more resistant to certain diseases, such as malaria (Hill et al., 1991; Aitman et al., 2000). In many parts of

* Corresponding author. Tel.: +1 304 293 8540; fax: +1 304 293 6699.
E-mail address: glin@wvu.edu (G. Lin).

0001-706X/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.actatropica.2004.05.006

world, ethnicity or ancestral origins can be identified geographically, and geographic surveillance of disease cool spots along the dimension of genetic susceptibility is likely to unpack genetic risk factors for various diseases.

Second, properly detecting low-value clustering can also reveal that a particular prevention program may be at work in a set of communities. For example, regional disease-incidence rates often change in response to health-prevention, environmental, and health-care factors. Normal regions may become high- or low-value clustered regions, and high- or low-value clusters may disappear or exchange positions between each other. Effectively monitoring these potential changes is crucial for the geographic evaluation of intervention and prevention programs (Elliot et al., 2000; Lawson and Kulldorff, 2000), which may provide lessons for other communities seeking to eliminate geographic inequality in health.

Finally, investigating potential cool spots may reveal environmental and cultural processes that operate at a regional or a community level. At the regional level, when geophysical environments are similar, the surveillance of kwashiorkor cool spots, for example, may provide clues for good dietary practices (WHO, 1999). At the community level, population mixing, which describes places with varying degrees of contact between infected and uninfected persons, has been positively associated with the hypothesis of an infectious origin of leukemia. If high rates of leukemia are found in places where many immigrants mix with longtime residents, low-rates of leukemia would be expected in places with little or no population mixing. There are many diseases for which environmental and cultural causes are suspected (e.g., familial Mediterranean fever), and isolating factors that contribute to low-value disease clustering will help to verify or nullify etiological agents.

Even though testing for the existence of low-value clustering seems a straightforward application of general clustering tests, few spatial epidemiologists have undertaken it, and there are several conceptual and testing issues that need to be addressed up front. First, some spatial events, such as infectious diseases, may not necessarily have low-value clustering (Elliot et al., 2000). The West Nile virus, for instance, was originally concentrated along the coastal regions of the mid-Atlantic states in the US in 1999 and 2000 (Marfin et al., 2001); we cannot presume, however, that at that time all other parts of the US were all cool spots with low-value clustering. In this case, low-value clustering does not inherently exist, at least at the outbreak of the virus. Second, some clustering methods have been expressly designed for identifying high-value clustering, and, consequently, may not always be suitable for testing low-value clustering. The scan test of Kulldorff (1997), which tests the existence of one (high-value) cluster within a study area, is an example. Third, some general spatial clustering and autocorrelation tests, such as Moran's I and Tango's C_G, cannot distinguish the existence of low-value clustering from that of high-value clustering (Lin, in press). In this case, a one-sided clustering test would be needed to supplement these test statistics (Lin, 2003).

In this paper, we propose a one-sided testing method for detecting low-value spatial clustering. Following the general approach of spatial epidemiology (Elliot et al., 2000; Lawson, 2001), we use the concept of relative risk, which is defined by the ratio of the region-specific mean and expected count. We first discuss some aspects of general low-value clustering and then provide a testing method based on the Tango test for rare diseases (Tango, 1995). In Section 3, we use simulated data to compare type I errors and the statistical powers between the proposed method and the Tango test. We then provide a case study of leukemia mortality in Minnesota in Section 4, which is followed by concluding remarks in the final section.

2. General tests for high and low-value clustering

Unlike high-value clustering, which has no upper (value) limit, the lower limit of the disease rate is 0 or no occurrence. This difference has several implications. First, it is generally unnecessary to consider low-value outliers because there is no negative value involved. If there is a statistically significant cluster with a low disease rate, it is unlikely to be attributed to outliers, especially when the disease is rare. Second, to avoid the situation in which there is no inherent low-value clustering, rare infectious diseases should not used for testing, because they will not likely occur in most study areas where people are not infected.

Third, the detection of low-value disease clustering tends to be more sensitive to the influence of spatial outliers and high-value clustering.

Consider a study area found to have a general tendency toward low-value clustering. Upon closer examination, however, a hot spot with a relative risk of about 0.15% is found to have a significant leverage on the mean or the average disease risk of 0.05%. If the hot spot, which certainly has not occurred by chance, is deleted or replaced with values around the mean, the tendency of low-value clustering would be reduced to the point of statistical insignificance. This is essentially the problem of comparing means for overlapping groups, a problem that also arises in a high-value clustering test. However, the potential impact of a low-value cluster on the detection of high-value clustering is generally less severe than that of a high-value cluster on the detection of low-value clustering. To develop a robust low-value clustering test, it is necessary to first reduce the potential bias resulting from the influence of high-value clustering.

Extending Marshall's definition, the existence of geographically cool spots can be viewed as abnormally low-values clustered somewhere within the study area. If there is no cluster of an abnormally high value, or a hot spot, the null hypothesis of no low-value clustering (or a constant disease risk across the study area) would be appropriate for a low-value clustering test. However, hot and cool spots often co-exist within a single study area, which may lead to an upward bias of the relative risk from the null hypothesis of a spatially constant mean. Ord and Getis (2001) noticed this problem, and they followed a common method for partitioning means into nonoverlapping groups (Calinski and Corsten, 1985; Looney and Jones, 2003). In this case, the partitions are made within a distance range and the rest of the study area, but the influence of hot spots remains a potential problem. The unresolved issue is how to properly define the null hypothesis such that it incorporates both spatially constant risk and the potential existence of hot spots. Rather than partitioning means, we propose to "partition" the null hypothesis that distinguishes the null hypothesis of (a) no spatial clustering (or a spatially constant risk) from the null hypothesis of (b) the existence of a hot spot without a cool spot.

If we simply use the spatially constant mean, or a, as the null hypothesis, the existing test methods, such as the Whittemore's W (Whittemore et al., 1987), the Getis G (Getis and Ord, 1992), and the Tango C_G, may not be appropriate for detecting low-value clustering, because the mean events in a particular region may not be proportional to the population in the region. By making this distinction, the null hypothesis for testing cool spots becomes $a \cup b$. When the null hypothesis of $a \cup b$ is rejected, for instance, the null distributions under $a \cup b$ should be adjusted for potential upward bias. If there are only a few outliers, it may be easier to delete them. If there are hot spots, we may not know about them until a cluster test is implemented, and we usually cannot delete hot spots from testing anyway. Herein, we propose a conditional replacement method for reducing the potential impact of hot spots or outliers. Although the proposed method is generally applicable to several testing methods, our discussion is based on the Tango C_G, because it is a general test that encompasses several testing methods (e.g., Whittemore W, Oden I, and Rogerson R) (Rogerson, 1999).

Tango's spatial clustering test is an extension of his one-dimensional time-series clustering test (Tango, 1995). Given a population size of ξ_i and a disease count of n_i at region i for a study area that has m regions, let $r = (r_1, \ldots, r_m)$ and $p = (p_1, \ldots, p_m)$, where r_i is the proportion of the ith regional count to the total count and p_i is the proportion of the ith regional population to the total population. Under Tango's null hypothesis that all the relative risks are equal, the Tango statistic

$$C_G = (r - p)^t A (r - p) \tag{1}$$

approximately follows a gamma-distribution, where $A = (a_{ij})$, $a_{ij} = \mathrm{e}^{-d_{ij}/\tau}$, d_{ij} is the distance between region i and region j, and τ is a constant. The P-value of the Tango statistic can be approximately computed by

$$P\{C_G > c\} \approx 1 - I\left(\frac{\nu + T_G\sqrt{2\nu}}{2}, \frac{\nu}{2}\right) \tag{2}$$

where the incomplete gamma $I(x, \phi)$ is defined by

$$I(x, \phi) = \int_0^x \frac{t^{\phi-1}}{\Gamma(\phi)} \mathrm{e}^{-t} dt$$

and Tango T_G (the standardized C_G) is defined by

$$T_G = \frac{C_G - E(C_G)}{\sqrt{\mathrm{var}(C_G)}}$$

with

$$E(C_G) = \frac{1}{n}\mathrm{tr}(AV_p), \qquad \mathrm{var}(C_G) = \frac{2}{n}\mathrm{tr}(AV_p)^2, \quad \text{and}$$
$$v = \frac{[\mathrm{tr}(AV_p)^2]^3}{[\mathrm{tr}(AV_p)^3]^2}$$

where tr(M) is the trace of a squared matrix M and V_p is the matrix with the ith diagonal entry $p_i(1-p_i)$ and the (i,j)th entry $-p_ip_j$ for $i \neq j$.

Tango's C_G, like a Chi-square test or a test similar to Moran's I, is a two-sided test. It can be used to detect both high- and low-value clustering, but there is no way to determine if a detected clustering tendency is attributable to a hot spot, a cool spot, or both (Lin, in press). To make the C_G a one-sided test, we need to reduce any potential effect from the other side (i.e., potential hot spots). Assuming that the relative risks of any normal regions are λ_0, then the relative risk in a hot spot is greater than λ_0, and the expected value within the hot spot is greater than its population times λ_0. Since the null hypothesis for low-value clustering does not exclude the presence of hot spots, it is necessary to eliminate the influence of hot spots when testing for the existence of cool spots.

As mentioned earlier, when the observed count n_i at region i is greater than the expected count $\lambda_0\xi_i$, it could result from either a random high within normal regions or the clustered high within a hot spot. Since the clustered high is not normal, its effect should be reduced before testing for low-value clustering. It is difficult, however, to determine beforehand if high-value regions in a study area are clustered or not. One way to deal with this uncertainty is to use the known λ to generate random numbers to replace high-value regions. Theoretically, we can always generate a disease distribution that resembles the actual disease pattern by using this λ. It turns out that if randomly high-values are replaced with a set of randomly high numbers, the overall effect in normal regions will remain the same; if the numbers in the clustered high regions are replaced with a set of randomly high numbers, the bias caused by high-value clustering can be reduced. It is, therefore, reasonable to blindly replace all the values above the known risk λ_0 with random numbers greater than λ_0. This strategy can be implemented by replacing all the regions with $n_i > \lambda_0\xi_i$ with a Poisson random variable P_i (with parameter $\lambda_0\xi_i$) that is also greater than $\lambda_0\xi_i$, where λ_0 is the real disease rate for the normal regions. With this replacement scheme, potential bias due to spatial outliers or hot spots can be reduced while retaining randomly low values in other regions. A less-biased estimator can, therefore, be obtained with the risks for all regions becoming closer to λ_0. If λ_0 is unknown, a trimmed mean can be used to estimate λ_0. For example, a 10% of trimmed mean can be obtained by cutting off the top and the bottom 10% observed relative risks and then calculating the mean (see Devore, 1999, p. 32).

3. Simulation

Since Tango's spatial clustering test originated from Tango's one-dimensional clustering test, we thought it would be worthwhile to use a one-dimensional simulation to demonstrate the importance of distinguishing between the null hypotheses of no clustering and no low-value clustering. For the purposes of the simulation, we fixed the disease rate λ_0 for normal regions at 0.0001, and independently generated the population at region i ($i = 1, \ldots, 100$) from the closest integer of the $\Gamma(10^4, 0.1)$ distribution. Hence, the mean of the population of any region was 10^5 with a standard error of 10^3. The distance between regions i and j was based on a straight line, or $d_{ij} = |i-j|$. Without loss of generality for the test results, let $\tau = 1$ or the average distance between two adjacent area unit, thus $a_{ij} = e^{-|i-j|}$. We began our simulation by examining type I errors of the Tango and the low-value clustering tests against the null hypothesis of no low-value clustering in the presence of hot spots only. We then evaluated the statistical power of the two tests when there was at least one cool spot. Finally, we compared the powers of these tests by combining the spatial structures from the previous simulations.

The upper panel of Fig. 1 displays the values of relative risks in the simulations with one, two, or three hot spots generated over a range of δ values from 0 to 1 in increments of 0.01. When $\delta = 0$, the highest relative risk was identical to λ_0; when $\delta = 1$, the highest relative risk was twice as much as λ_0. In this way, type I errors for the low-value clustering test were based on the presence of one, two, or three hot spots, and the relative risk of the center point within a hot spot increased gradually from 100 to 200% of the relative risk $R = 1$. In the case of one hot spot, we inserted a

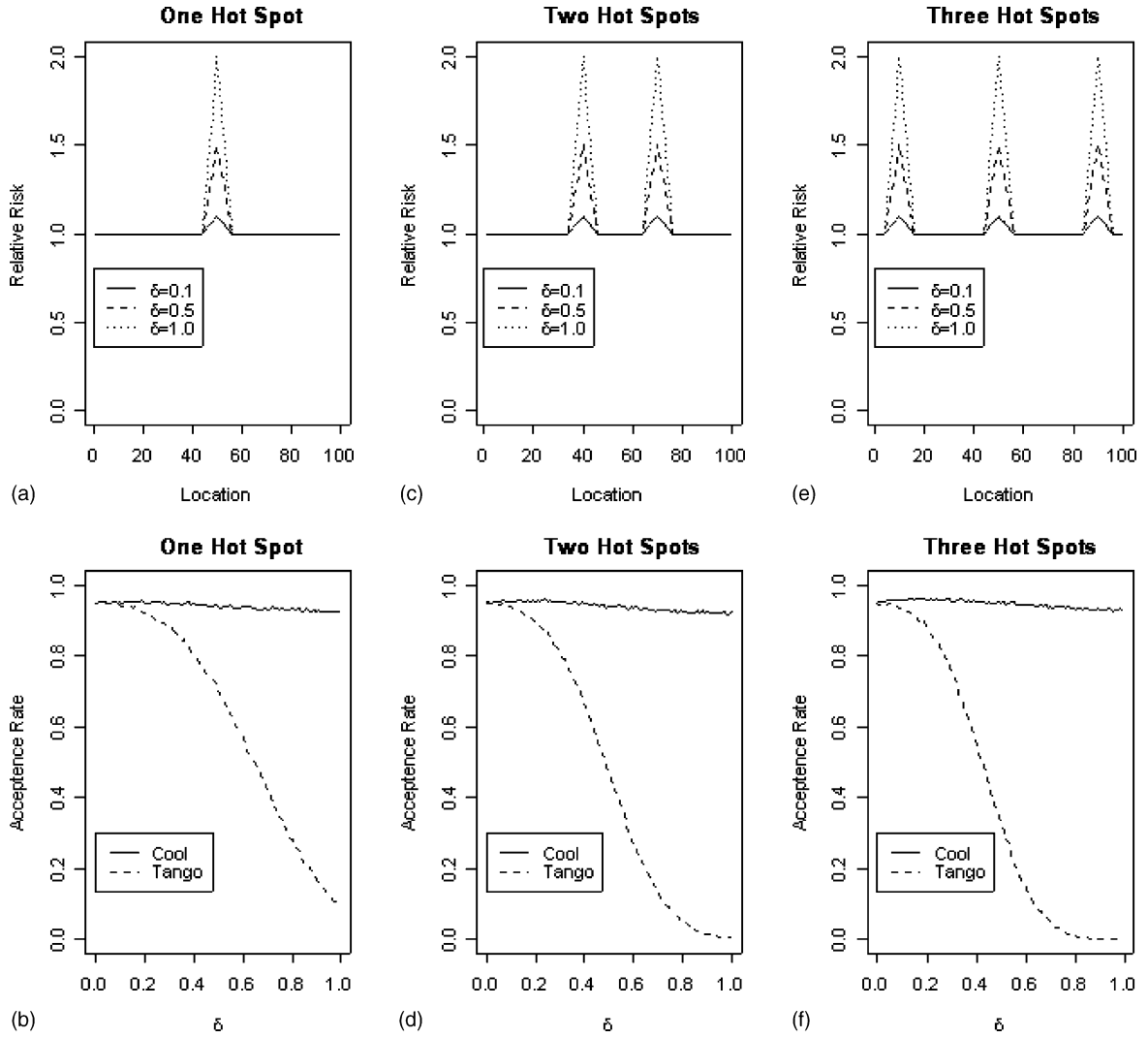

Fig. 1. Acceptance rates for low-value clustering in the presence of hot spots.

hot spot at the midpoint as shown in Fig. 1a (i.e., if $|i - 50| < 6$, $R_i = 1 + \delta|i - 50|/6$; otherwise $R_i = 1$, where $\delta \in [0, 1]$ defines the strength of the cool spot, and 50 is the midpoint). For each δ selected in our simulation, we repeated 10,000 runs, and the results at the 5% level were computed based on the *P*-values of the related statistics for each run. The lower panel of Fig. 1 displays the corresponding results.

In the presence of a hot spot, the test of low-value clustering consistently accepted the null hypothesis for all δs, with the acceptance rate being around 0.95. The acceptance rates of Tango's C_G were at the acceptable level of 0.95 only when there was no hot spot or the strength of the hot spot was very small (δ was close to 0). As the strength of the high-value clustering increased, the acceptance rate for C_G decreased rapidly; nearly reaching 0 when δ was close to 1. There was little difference in type I errors in the presence of additional (i.e., two or three) hot spots. These results suggest a clear inverse relationship between the strength of a hot spot and the type I error rate of C_G: when hot spots existed, the rejection rate of Tango C_G is much greater than the test level if it is used to detect low-value clustering.

We also evaluated the powers of these tests and their effectiveness in detecting low-value clustering (Fig. 2). The results showed that all of the clustering tests were very effective, having an acceptance rate of 0.95 when there was no clustering or $\delta = 0$; the acceptance rates of the corresponding null hypotheses were all close to 0 when there was low-value clustering or δ was close to 1. The statistical powers increased slightly when two or three cool spots were inserted. In all cases, the statistical powers of C_G and the low-value clustering test were almost identical. However, since the results from C_G was unable to distinguish high- from low-value clustering, the results from the low-value clustering test represented an important information gain, which unambiguously rejected the null hypothesis of no low-value clustering and, thus, concluded its existence. The proposed low-value clustering test were indeed supplementary to C_G. Like C_G, the low-value clustering test was likely to be significant in the presence of a cool spot; unlike C_G, it was rarely significant if there was a hot spot only.

Finally, we simulated situations in which both cool and hot spots existed (Fig. 3). The results from the Tango test consistently registered a greater statistical power than did the low-value clustering test. This difference in statistical power could be explained in two different ways. First, the greater power of Tango's test is to be expected, because this test considers the existence of either cool or hot spots as clustering whereas the low-value clustering test only considers the existence of cool spots. The lower statistical power of

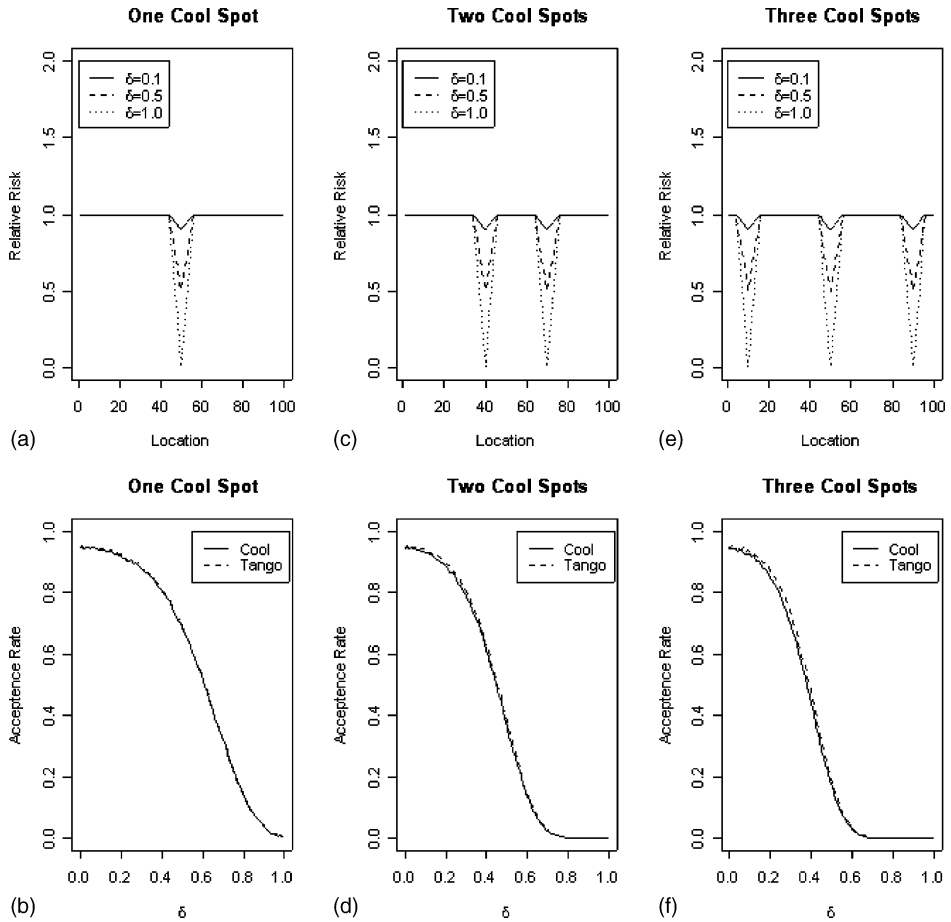

Fig. 2. Acceptance rates for low-value clustering in the presence of cool spots.

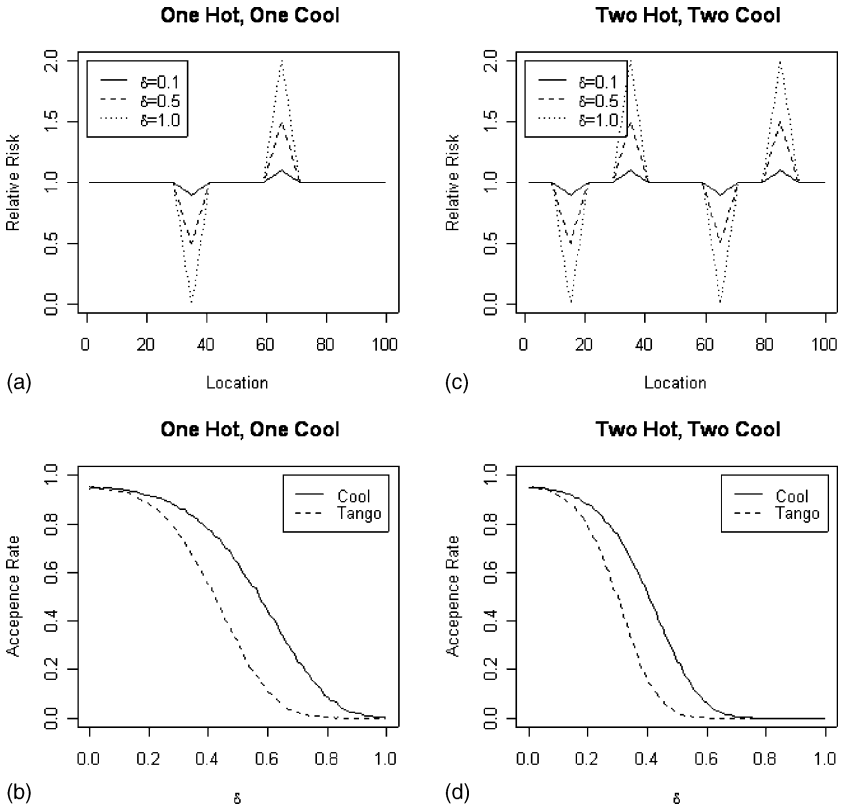

Fig. 3. Acceptance rates for low-value clustering in the presence of both hot and cool spots.

the low-value clustering test is compensated for by greater information gain or less ambiguity. Second, the power gap between the two tests can be evaluated using different δ values, and the area in which a false rejection of the null hypothesis may fall can be determined. In the case of one cool and one hot spot, the rejection rate of no clustering from C_G was greater than the conventional level of 95% when δ was greater than 0.72, but the low-value clustering test did not reach this level until δ was greater than 0.91. This discrepancy indicates that between a weaker cool spot range (i.e., between 1.72 and 1.91 times of the relative risk), C_G may signify significant clustering that may be due to the effects of high-value clustering. Taken these two explanations together, the low-value clustering test is an effective one-sided test that could be used as an alternative to C_G when there is a need to reduce abnormal effects caused by hot spots and outliers.

4. Minnesota leukemia case study

We chose leukemia to test our proposed method, because its etiological causes are largely unknown and because epidemiologists can learn ecological risk factors from both cool and hot spots. Extensive studies have related environmental factors and agrochemicals to leukemia incidence and mortality, but no conclusive geo-environmental leukaemogens have been reported (Boyle et al., 1996; Wartenberg, 1998). For our case study, we selected the 5-year (1992–1996) county-level leukemia-mortality data from the Minnesota Cancer Surveillance System, which records the number of deaths due to leukemia separately for males and females. According to the US National Cancer Institute, the 5-year (1990–1994) leukemia mortality rate in Minnesota for white males was the second highest in the US, or about 11% higher than the national average, and that for white females was ranked 23rd,

or just slightly higher than the national average. We used the county-level populations for the state from the 1990 US Census to analyze leukemia mortality rates for both males and females. We compared the 1990 Census populations with the 1994 county estimates (the mid-year of 1992–1996) and found very small changes in population for most of the counties. For this reason, we decided to use the Census data rather than the estimates. We used Euclidean distance to measure geographic proximity and set $\tau = 35$ miles, which is about the average distance between any two centroids of adjacent counties. We also experimented with τ values between 20 and 50, but the results were not very sensitive within this distance range.

Figs. 4 and 5 displayed the geographic distributions of the 5-year male and female leukemia-mortality rates. The mortality rates were grouped into seven percentiles, with roughly an equal number of counties within each percentile-category. The average mortality rates for males and females per 1000 were 0.80 and 0.55, respectively, with the total number of deaths during the 5-year period being 1718 for males and 1226 for females. Among the 87 counties, two reported a zero death rate for females, but none reported a zero death rate for males. The highest death rate for males was 1.87 per 1000, more than double the corresponding mean; the highest death rate for females was 0.89 per 1000, or 1.62 times the corresponding mean. For males, there appeared to be clusters of both high and low values. The high-value cluster seemed to be located around the central west, and the low-value cluster seemed to be located along the lower Minnesota River basin in the central south. The picture for females was less evident, having no apparent geographical pattern.

In this case study, λ_0 was unknown, making it necessary for us to estimate λ_0 as the true value in the proposed low-value clustering test. Given the emphasis of our study, we used a 10% trimmed mean (Devore, 1999) to derive a less biased mean as the estimated value of λ_0 (denoted by $\hat{\lambda}$). Because this method could still be biased from the true λ_0, we assessed the sensitivity of the *P*-values based on a range of λ values from $0.9\hat{\lambda}$ to $\hat{\lambda}$. In other words, a range of downward λ values was used to correct any potential upward bias resulting from the existence of high-value clusters. The $\hat{\lambda}$ values for males and females were 0.000796 (or 7.96 deaths per 10,000) and 0.000546, respectively,

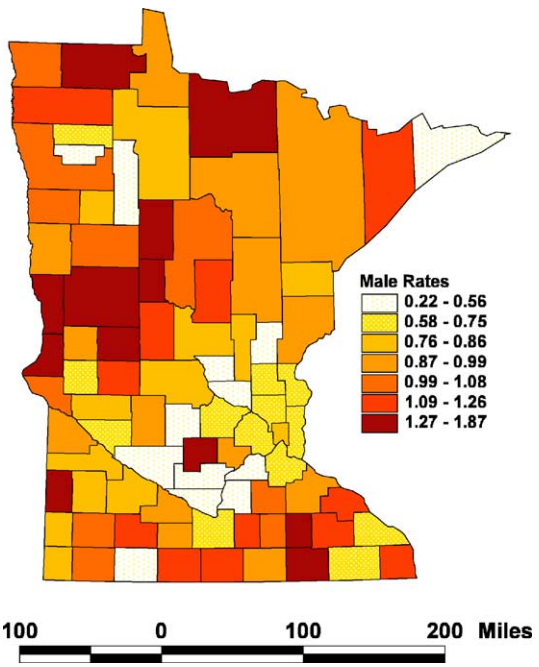

Fig. 4. Male leukemia mortality rates per 1000 in Minnesota.

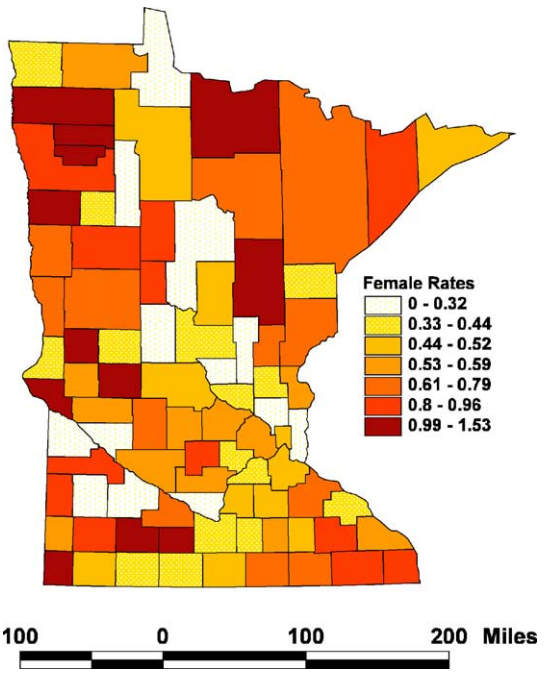

Fig. 5. Female leukemia mortality rates per 1000 in Minnesota.

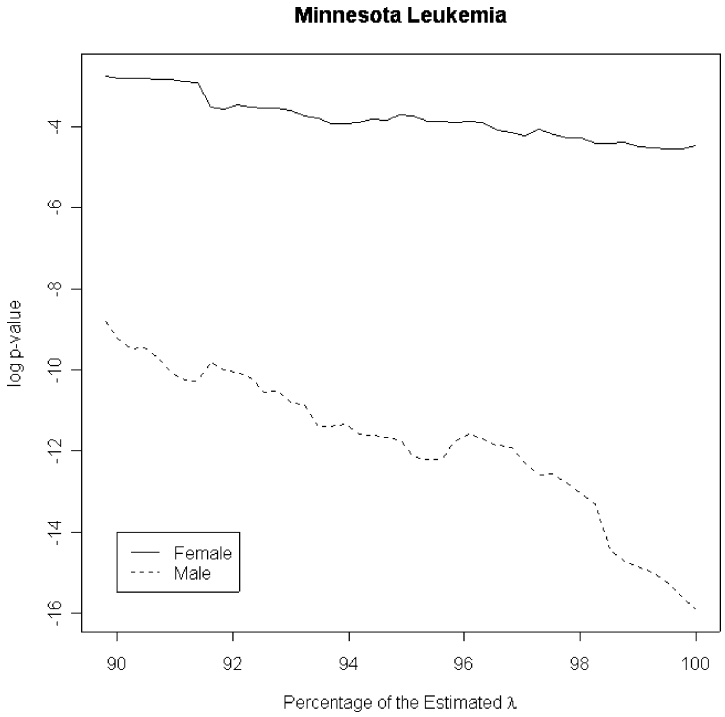

Fig. 6. The *P*-values for testing low-value leukemia clustering in Minnesota.

which were not very different from the average rates of 0.000801 for males and 0.000550 for females.

Taking the male and female $\hat{\lambda}$ values to be the true values, the *P*-values of the low-value clustering test were computed repeatedly 10,000 times. We considered the medians of those repeated *P*-values to be the most trustworthy, and they are displayed in Fig. 6 in the natural log scale. Note that the smaller the *P*-value, that is, the more negative the value along the log scale, the greater is its significance. A log *P*-value greater than 3 would not be significant at the probability level of 0.05, and a log *P*-value greater than 4.6 would not be significant at the probability level of 0.01.

Overall, the *P*-values for males were very small, being less than 0.001 ($e^{-8} \approx 0.003$) anywhere between $0.9\hat{\lambda}$ and $\hat{\lambda}$. As the λ values decreased, the *P*-values of the low-value clustering test increased, and become less significant. Because a potential abnormally hot spot would cause an upward bias from the true λ_0, a correction would lead to a less significant result. In any case, the influence of a potential biased mean in a range of 10% upward, which was evaluated by a range 10% downward of λ, will not lead to a different outcome. We, therefore, concluded that there is an overall tendency of low-value spatial clustering among males who died from leukemia during the study period.

For females, however, the results were not significant at the 0.01 probability level anywhere from the estimated λ to 10% less. In addition, the *P*-values were very sensitive to the shifts in λ values. When the λ fell within 8% of the $\hat{\lambda}$ value, the results were significant at the 0.05 level; when the λ value shifted beyond 8%, the results were no longer significant. Since the sample was fairly large and the estimated mean could easily be biased upward by 6–7%, we used a more conservative confidence level of 0.01. Consequently, the existence of cool spots could not be concluded for females in general.

5. Concluding remarks

In this paper, we have provided a method to test for low-value spatial clustering. When the expected

disease risk was known, the proposed conditional replacement method was effective in reducing a potential overestimate of the disease rate due to the presence of regions of structurally high value. Although the simulations and the data example were in reference to the Tango test, the conditional replacement method can be applied to other clustering tests. In the presence of a hot spot, the type I error rate based on the null hypothesis of the low-value clustering test was much lower than that based on the Tango test. The powers of the two tests, however, were almost identical in the presence of a single cool spot only. When both hot and cool spots coexist, the Tango test has a greater statistical power, which also was also accompanied by a loss of information. In this regard, the low-value clustering test can be used not only for testing the tendency for low-value clustering but also for supplementing other general clustering tests to reduce false alarms. This is especially the case when there is a suspicion of which clustering tendency (high or low) contributes more to a significant test result.

When λ was unknown, as in the case of leukemia mortality in Minnesota, we first estimated the true disease rate for normal regions by using a 10% trimmed mean. Because, according to our null hypothesis, hot spots may exist, this estimated disease rate for normal regions may be biased upward. This means that the true λ_0 may be less than the estimated value. For this reason, we evaluated a range of λ values. The existence of low-value clustering was concluded for males but not for females. Based on these results, we speculate that a low level of population mixing in the rural counties along the lower Minnesota River may reduce the chance of infection through contact. It is also possible that the residues of pesticides may spatially vary in terms of intensity and interaction the with drainage system. These are some of the research questions warrant further epidemiological studies by using a focused test and other covariates.

Several methodological issues also warrant further investigation. First, even though the focus of our study was on low-value clustering, our method can be equally applied in high-value clustering tests. Second, a better method is needed to empirically derive an estimator that is sufficiently close to the true λ. Furthermore, when the study area is part of a larger region, the risk at the regional level should not be ignored. In our empirical case, the rate of mortality in males due to leukemia was 11% higher than the national average; if a hot spot with a highly excessive mortality rate was presented in the study area, we might want to account for it first before estimating the true mean. Third, covariates may still be considered by using covariate-related statistics, such as the scan test (Kulldorff, 1997). Finally, it will be important to determine how the critical region is bounded or affected by an estimated λ value in extending our method to other clustering tests.

Acknowledgements

This paper was written while the first author was a fellow at Baylor College of Medicine. The NIH fellowship grant is acknowledged. The authors would also like to thank Pamela Tice for her editorial assistance.

References

Aitman, T.J., Cooper, L.D., Norsworthy, P.J., Wahid, F.N., Gray, J.K., Curtis, B.R., McKeigue, P.M., Kwiatkowski, D., Greenwood, B.M., Snow, R.W., Hill, A.V., Scott, J., 2000. Malaria susceptibility and CD36 mutation. Nature 405, 1015–1016.

Boyle, P., Walker, A.M., Alexander, F.E., 1996. Historical aspects of leukemia clusters. In: Boyle, P., Alexander, F.E. (Eds.), Methods for Investigating Localized Clustering of Disease. IARC Scientific Publications. Lyon, France, vol. 135, pp. 1–20.

Calinski, T., Corsten, L.C., 1985. Clustering means in ANOVA by simultaneous testing. Biometrics 41, 39–48.

Cuzick, J., Edwards, R., 1990. Spatial clustering for inhomogeneous populations with discussion. J. R. Stat. Soc. Ser. B 52, 73–104.

Devore, J., 1999. Probability and Statistics for Engineering and the Sciences, fifth ed. Duxbury, Pacific Grove, USA.

Elliott, P., Wakefield, J., Best, N., Briggs, D., 2000. Spatial Epidemiology Methods and Applications. Oxford University Press, Oxford.

Getis, A., Ord, J., 1992. The analysis of spatial association by use of distance statistics. Geogr. Anal. 24, 189–206.

Hill, A.V., Allsopp, C.E., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A., Bennett, S., Brewster, D., McMichael, A.J., Greenwood, B.M., 1991. Common west African HLA antigens are associated protection from severe malaria. Nature 352, 560–595.

Kulldorff, M., 1997. A spatial scan statistic. Commun. Stat. Theory Methods 26, 1481–1496.

Lawson, A.B., 2001. Statistical Methods in Spatial Epidemiology. Wiley, New York.

Lawson, A., Kulldorff, M., 2000. A review of cluster detection methods. In: Lawson, A., Biggeri, A., Bohning, D., Lesaffre, E., Viel, J., Bertollini, R. (Eds.), Disease Mapping and Risk Assessment for Public Health, chapter 7. New York, Wiley.

Lawson, A.B., Denison, D.G.T., 2000. Spatial Clustering Modeling. CRC Press, New York.

Lin, G., 2003. A spatial Logit Association Model. Geogr. Anal. 35, 329–340.

Lin, G. Comparing spatial clustering tests based on rare to common spatial events. Comput. Environ. Urban Syst., in press.

Looney, S., Jones, P., 2003. A method for comparing two normal means using combined samples of correlated and uncorrelated data. Stat. Med. 22, 1601–1610.

Marfin, A.A., Petersen, L.R., Eidson, M., Miller, J., Hadler, J., Farello, C., Werner, B., Campbell, G.L., Layton, M., Smith, P., Bresnitz, E., Cartter, M., Scaletta, J., Obiri, G., Bunning, M., Craven, R.C., Roehrig, J.T., Julian, K.G., Hinten, S.R., Gubler, D.J., 2001. Widespread west Nile virus activity, eastern United States, 2000. Emerg. Infect. Dis. 74, 730–735.

Marshall, R.J., 1991. A review of methods for the statistical analysis of spatial patterns of disease. J. R. Stat. Soc. Ser. A 154, 421–441.

Ord, K., Getis, A., 2001. Testing for local spatial autocorrelation in the presence of global autocorrelation. J. Reg. Sci. 41, 411–432.

Rogerson, P.A., 1999. The detection of clusters using a spatial version of the Chi-square goodness-of-fit statistics. Geogr. Anal. 31, 130–147.

Tango, T., 1995. A class of test for detecting ‘General’ and ‘Focused’ clustering of rare diseases. Stat. Med. 14, 2323–2334.

Wartenberg, D., 1998. Residential magnetic fields and childhood leukemia A meta-analysis. Am. J. Public Health 88, 1787–1794.

Whittemore, A., Friend, N., Brown, N., Holly, E., 1987. A test to detect clusters of disease. Biometrika 74, 631–635.

World Health Organization, 1999. Management of severe malnutrition: a manual for physicians and other senior health workers. World Health Organization, Geneva.