European Journal of Epidemiology **15**: 15–22, 1999.

Spatial autocorrelation of cancer in Western Europe

Michael S. Rosenberg[1], Robert R. Sokal[1], Neal L. Oden[2] & Donna DiGiovanni[3]
[1]*Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, New York;* [2]*The EMMES Corporation, Potomac, Maryland;* [3]*Department of Neurology, State University of New York at Stony Brook, Stony Brook, New York, USA*

Accepted in revised form 17 August 1998

Abstract. We applied the techniques of spatial autocorrelation (SA) analysis to 40 cancer mortality distributions in Western Europe. One of the aims of these methods is to describe the scale over which spatial patterns of mortalities occur, which may provide suggestions concerning the agents bringing about the patterns. We analyzed 355 registration areas, applying one- and two-dimensional SA as well as local SA techniques. We find that cancer mortalities are unusually strongly spatially structured, implying similar spatial structuring of the responsible agents. The small number of spatial patterns (4 or 5) in the 40 cancer mortalities suggests there are fewer spatially patterned agents than the number of cancers studied. SA present in variables will bias the results of conventional statistical tests applied to them. After correcting for such bias, some pairwise correlations of cancer mortality distributions remain significant, suggesting inherent, epidemiologically meaningful correlations. Local SA is a useful technique for exploring epidemiological maps. It found homogeneous high overall cancer mortalities in Denmark and homogeneous low mortalities in southern Italy, as well as a very heterogeneous pattern for ovarian cancer in Ireland.

Key words: Cancer mortalities, Europe, Spatial autocorrelation

Introduction

Spatial autocorrelation (SA) is the dependence of the values of a variable at specified geographic locations on the values of the variable at neighboring locations. Spatially autocorrelated data violate the assumption of independence required for most standard statistical tests, calling for special tests designed to remove the dependence of the variable on geography. Although the recognition and analysis of SA is often associated with removing the internal dependence of variables on the underlying spatial structure during hypothesis testing, the analysis of spatial autocorrelation in and of itself can lead to important discoveries about the scale over which spatial patterns occur, which in turn may suggest underlying factors with similar patterns. Spatial autocorrelation analysis has been used to study a variety of phenomena, such as the genetic structure of plant and animal (including human) populations [1–6], mortality patterns [7], and morphological patterns [8–11]. Numerous studies have compiled data for atlases of comparative cancer rates for countries, counties, and regions [12–15], yet these data have not so far been analyzed by the techniques of spatial autocorrelation analysis. Epidemiology can be defined as the search for the determinants of diseases through the study of their occurrence and distribution [14]. SA analysis should provide important insights into the geographic scale over which cancers are occurring, narrowing the list of putative causes of these diseases to those with similar spatial pattern. The purpose of this study was to examine the distribution of cancer mortalities in western Europe using spatial autocorrelation analysis. This will quantify the spatial patterns, enabling us to search for similarities in distribution among different cancers.

Materials and methods

Cancer mortality figures were taken from the Atlas of Cancer Mortality in the European Economic Community [14]. We used cancer mortality rates, rather than incidence or prevalence data, because they represented the most complete and comprehensive data set available for Europe at a fine geographic scale. Mortality rates were presented as the gender-specific, age-standardized (world standard) rate per 100,000 per annum. The atlas contains the mortality rates for 40 cancers for each of 355 registration areas in Western Europe, comprising the 9 countries that were members of the European Economic Community in the 1970s: Belgium, Denmark, Eire, France, Great Britain, Italy, Luxembourg, the Netherlands, and West Germany. The mortality rates represented the years 1971 through 1980, the

Table 1. List of all cancers for which mortality data were available

Male and female	Bladder	Brain	Colon/Rectum	Gall bladder
	Hodgkin's disease	Larynx	Leukemia	Lung
	Lymphoma	Malig. melanoma	Mult. myeloma	Oesophagus
	Oral	Pancreas	Stomach	Thyroid
	Urinary tract			
Female only	Breast	Cervix	Ovary	Uterus
Male only	Prostate	Testis		

years of record varying slightly by country. The 40 cancers are listed in Table 1.

Each registration area was given latitude and longitude coordinates (obtained from a variety of atlases), representing either the approximate center of the area or a town or city near the center. These coordinates were used to calculate the geographic distance between each pair of areas as a great-circle distance in kilometers.

The mortality values over the set of areas for a specific cancer are referred to as a cancer mortality surface. Simple product-moment correlations were calculated for each pair of mortality surfaces and their significances were evaluated using a procedure developed by Clifford et al. [16] that corrects for autocorrelation of the data. This test provides a measure of the similarity of cancer patterns, although it does not describe what those patterns are.

To describe these spatial patterns, two common spatial autocorrelation coefficients were calculated for each cancer surface, Moran's I, a product-moment coefficient, and Geary's c, a standardized squared-distance measure [17, 18]. Given the large number of registration areas, autocorrelation coefficients were calculated for fifteen distance classes, each containing an approximately equal number of distances. The upper limits of the distance classes were 222, 342, 438, 525, 613, 702, 787, 872, 959, 1061, 1176, 1308, 1485, 1747, and 2865 km. Our experience is that studies with such a large sample size are not sensitive to the number of chosen distance classes. When plotted against distance, these coefficients produce a one-dimensional spatial correlogram [17, 19–21]. Moran's I and Geary's c often lead to similar conclusions, and were fairly compatible in the present study. We will report only the results for Moran's I, referring to those of Geary's c only when they differ substantially. Values of Moran's I range approximately from 1 (positive SA) to -1 (negative SA), with an expected value of $-1/(n-1)$, where n is the number of localities. The autocorrelation coefficients were calculated using the SAAP program [22]. The overall significance of each correlogram was evaluated using the Bonferroni test suggested by Oden [23].

We used k-means cluster analysis [24] to group the cancers by their correlograms. This analysis starts by randomly assigning each correlogram to one of k clusters, then shuffles the correlograms between clusters and, by a stepwise procedure, tries to minimize the sum of squared distances from each cluster member to the cluster centroid. One hundred separate attempts were made to optimize the k-clusters for $k = 2$ to $k = 10$. We chose a value of k above which the sum of squares within clusters did not decrease appreciably.

Two dimensional correlograms [2, 25] take into account compass bearing as well as distance. We were able to construct distance/direction classes for 5 annuli, whose outer limits are 150, 600, 1350, 2400, and 3750 km. We calculated such correlograms for each surface using the CIRCLE program written by one of us (NLO). The two-dimensional correlograms were clustered using the k-means clustering scheme described above.

The agreements between the partitions induced by I and c and by one-dimensional and two-dimensional correlograms, were estimated in the form of Rand indexes [26]. We employed expression (5) of these authors which corrects the index for chance and normalizes it, so that perfect agreement is indicated by an index value of 1. We tested its significance by randomly partitioning the members of one of the marginal classifications into the same number of classes as the observed data, with the same number of correlograms per class. All randomizations were carried out 9999 times.

The past few years have witnessed the development of methods for local spatial autocorrelation [27–29]. This is performed by calculating a local SA coefficient for each individual sampling locality i, rather than for an entire distance class. Three different such coefficients were calculated for each locality. I_i and c_i, which are local variants of Moran's I and Geary's c due to Anselin [28], and G_i^*, which is a measure of the amount of clustering between the individual locality i and the neighboring sampling locations developed by Getis and Ord [27]. Both high and low values of G_i^* indicate the presence of strong positive local SA; but high values of G_i^* signify clusters of relatively high values of the variable, while low values indicate clusters of low values. Although the significance of local SA cannot currently be estimated when there is significant overall (global) SA [30–31], we have found relative significance among the coefficients to be a good indicator of local spatial patterns. For each cancer, the probability associated with each coeffi-

cient was calculated by a conditional permutation test [27, 28, 30] for I_i and c_i, and by a total permutation test (ibid.) for G_i^*. These probabilities were then ranked (I_i, c_i, and G_i^* are all ranked separately), and the localities matching the top and bottom 10% (e.g., for I_i–coefficients the nominally most significant positive and negative values) were plotted on a map of Europe. This allows one to identify areas of relatively high or low local spatial autocorrelation. To estimate the average local SA pattern across all cancers, the ranks for a given local statistic (e.g., I_i) for each locality were averaged across all cancers. These average ranks were themselves ranked and plotted as above.

Results

The one-dimensional Moran's I correlograms of the 40 cancers are shown in Figure 1. All 40 correlograms are significant by the Bonferroni criterion [23] at $p \leqslant 0.05$, and 486 out of 600 individual autocorrelation coefficients (81%) are nominally significant at $p \leqslant 0.05$. All 40 cancers show significant positive autocorrelation through the second distance class (342 km). At the largest distance class (distances greater than 1747 km), 75% of the cancers show significant negative autocorrelation. This is an extremely high level of spatial autocorrelation for observational data, with most of the cancers showing some sort of a gradient, often referred to as a clinal pattern. The results of the k-means clustering of the one-dimensional correlograms revealed that these tend to cluster into four groups (Table 2). The mean correlograms for the four groups are show in Figure 2. Two of the groups show a clinal pattern, with positive autocorrelation at short distances and negative autocorrelation at long distances, differing only by the slope of the cline. The third group shows positive autocorrelation at short distances, tapering off to no autocorrelation at longer distances. This type of pattern may indicate a patchy distribution of the variable in question, with the average patch size approximated by the X-intercept. The final group shows strong positive autocorrelation at short distances and negative autocorrelation at intermediate distances, returning to moderate positive autocorrelation at long distances. Such a pattern would be generated when the coordinates of a bowl- or mound-shaped surface are subjected to SA analysis. Because only some countries in Western Europe were members of the European Economic Community in the 1970s, the bowl shape is distorted and positive SA is only moderate at the highest distances. But for the farthest distances (1747 to 2865 km; e.g., Denmark to Sicily or Shetland Islands to Sicily) the mortalities of the 6 cancers are similar. Over distances up to 1500 km, these mortalities effectively form gradients. We shall call this pattern imperfect bowl, for lack of a better term.

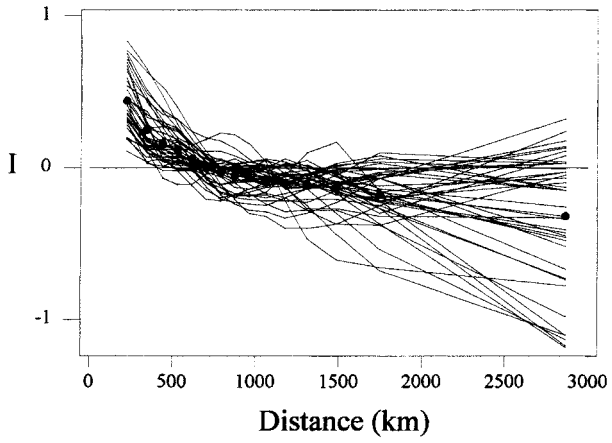

Figure 1. One-dimensional correlograms of Moran's I for mortalities of 40 cancers. Filled circles represent the average correlogram.

In general, the results for Geary's c tend to be less significant for these data than those of Moran's I. When applying the Bonferroni criterion [23] to the Geary's c correlograms, larynx F and thyroid M mortalities lack significant SA. The k-means clustering of Geary's c yielded approximately the same 4 groups and patterns as did the analysis of Moran's I. We tested the partitions induced by I and c and found

Table 2. The groups resulting from the k-means clustering of the one-dimensional correlograms of Moran's I

Strong cline	Breast Lung F Pancreas M,F	Cervix Oesophagus F	Colon/Rectum F Ovary
Weak cline	Bladder Malignant melanoma F Testis	Colon/Rectum M Multiple myeloma M,F Urinary tract F	Lymphoma M,F Oral F
Patchy	Bladder M Larynx F Malignant melanoma M Urinary tract M	Brain M,F Leukemia M,F Stomach M,F Uterus	Hodgkin's disease M,F Lung M Thyroid M,F
Imperfect bowl	Gall bladder M,F Oral M	Larynx M Prostate	Oesophagus

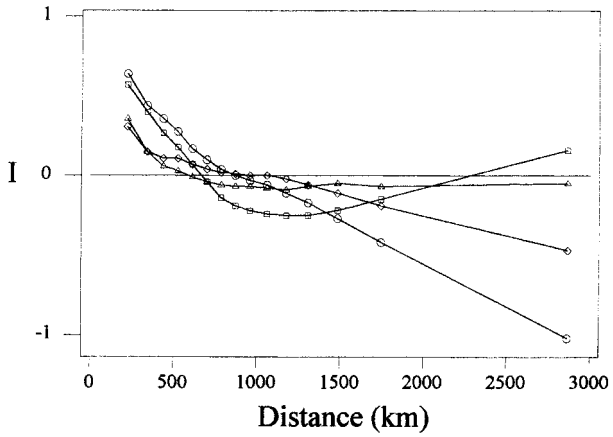

Figure 2. Average one-dimensional correlograms of the 4 *k*-means clusters of the Moran's *I* correlograms. Circles: Group 1, strong cline; diamonds: Group 2, moderate cline; triangles: Group 3, patchy; squares: Group 4, imperfect bowl.

a highly significant ($p = 0.0001$) Rand index of 0.46067.

Two-dimensional correlograms account for direction as well as distance in their calculations, providing additional information on the spatial distribution of values. Because of the additional complexity, the two-dimensional correlograms for individual cancers (not shown) are much more variable than the one-dimensional correlograms. All are Bonferroni significant at $p \leqslant 0.05$. The results from the *k*-means clustering of the two-dimensional correlograms (using Moran's *I*) are shown in Table 3, with the mean correlograms for each cluster pictured in Figure 3. The first group (Figure 3a) shows positive SA out to at least 600 km, with negative SA at greater distances, oriented with a slight NE-SW cline. The second group (Figure 3b) shows significant SA, but is nondirectional. The third group (Figure 3c) shows a complex pattern equivalent to the one-dimensional imperfect bowl, with the bowl oriented in a NW-SE direction. The fourth and fifth groups (Figures 3d and 3e) both show clear, strong N-S clines, differing by the strength of the cline. We compared the

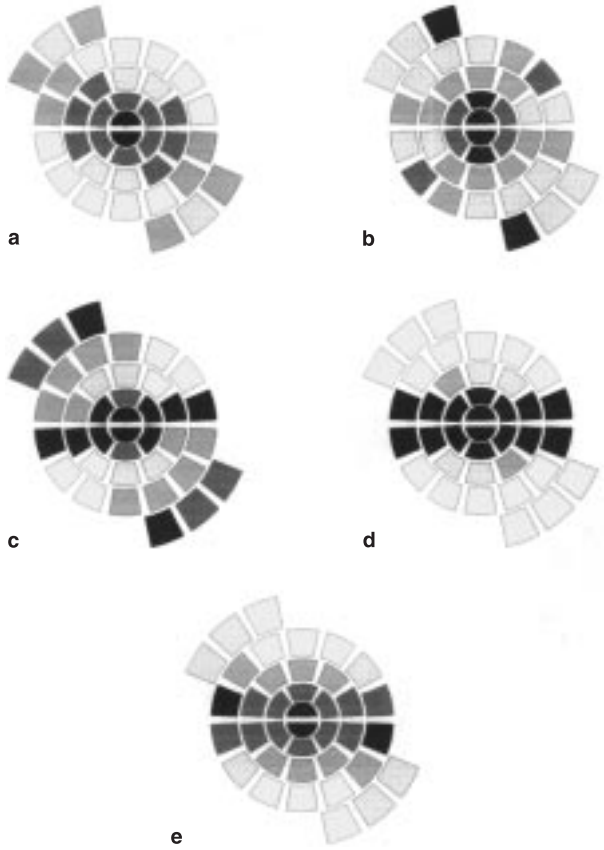

Figure 3. Average two-dimensional correlograms based on 5 clusters. Values of *I*: white: −2.44 to −0.32; pale gray: −0.29 to −0.13; gray: −0.12 to 0.02; dark gray: 0.03 to 0.18; black: 0.19 to 1.21. Upper limits of distance class annuli: 150, 600, 1350, 2400, 3750 km. a, Group 1; b, Group 2; c, Group 3; d, Group 4; e, Group 5.

partition by two-dimensional *I*-correlograms to that induced by the partition of one-dimensional I-correlograms. We found a moderate corrected Rand coefficient of 0.13657, significant at $p = 0.0075$.

Maps of local SA coefficients can reveal whether there are areas of particularly high or low heterogeneity. The map of I_i for ovarian cancer (Figure 4) reveals that Denmark, Southern Italy, and Sicily are areas of high positive local SA, while Ireland and

Table 3. The groups resulting from the *k*-means clustering of the two-dimensional correlograms of Moran's *I*

Group 1	Bladder M,F	Cervix	Malignant melanoma M
	Urinary tract M,F	Uterus	
Group 2	Colon/Rectum M	Gall bladder M,F	Hodgkin's disease M,F
	Larnyx F	Leukemia M,F	Lung M,F
	Multiple myeloma F	Oesophagus F	Oral F
	Prostate	Stomach M,F	Thyroid M,F
Group 3	Brain M,F	Larynx M	Malignant melanoma F
	Oesophagus M	Oral M	
Group 4	Breast	Colon/Rectum F	Multiple myeloma M
	Ovary	Pancreas M,F	
Group 5	Lymphoma M,F	Testis	

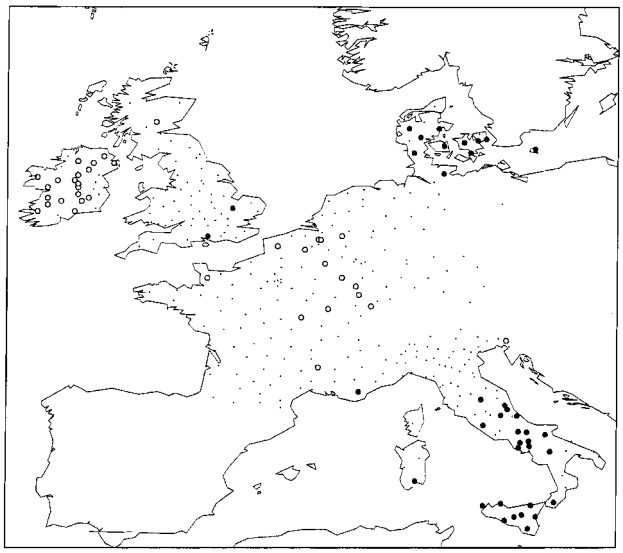

Figure 4. Map of the top (positive local SA; filled circles) and bottom (negative local SA; hollow circles) of the ranked conditional permutational probabilities of I_i for ovarian cancer mortality. We aimed at 10% for both ends of the distribution, but were unable to achieve that exactly. There are more positive areas, because of tied values of I_i and fewer negative areas, because the number of local coefficients below their expected value was less than 10%. The remaining localities are shown as dots.

scattered areas in the low countries, Germany and France are areas of negative local SA. Why mortality from ovarian cancer should be so heterogeneous in Ireland is an interesting question. The plots of G_i^* for individual cancers indicate cohesive areas of relatively high or low cancer rates. The plot of G_i^* for ovarian cancer (Figure 5) indicates that clusters with the highest rates are found in England and Denmark, while those with the lowest rates are found in southern Italy and nearby islands, an observation which can be confirmed from an examination of the cancer mortality maps provided with the atlas [14].

An examination of Figures 4 and 5 reveals the general relationship between I_i and G_i^*. I_i reveals the areas of greatest local homogeneity: Denmark and southern Italy; G_i^* indicates which of these areas contain high mortalities (Denmark) and which contain low mortalities (Southern Italy). There is no correspondence between low values of I_i, and G_i^* on these maps because low values of I_i will necessarily lead to average values of G_i^* which are not displayed.

In addition to plots of individual cancers, maps of the ranked averages across all cancers are also of interest. Figures 6 and 7 are plots of average ranks for I_i and G_i^*, respectively. The maps indicate the top and bottom 10% of the ranked average of the ranked probabilities of the local SA coefficient for each cancer.

Figure 6 shows that on average, Southern Italy, Denmark and Southern England show the most positive local SA; Northern Ireland, the low countries, and Northern Italy show the most negative SA. Figure 7 reveals that the highest cancer mortality rates are found in Denmark and in Northern Italy, while the lowest cancer mortality rates are found in Southern France, Southern Italy, Corsica, and Sardinia. There is an interesting pattern of these local coefficients in Northern Italy; I_i shows negative SA along the fringes of Northern Italy, while G_i^* reveals two small tight clusters of high values. Careful inspection reveals that (with a single exception) these are not the same localities. The hollow circles in

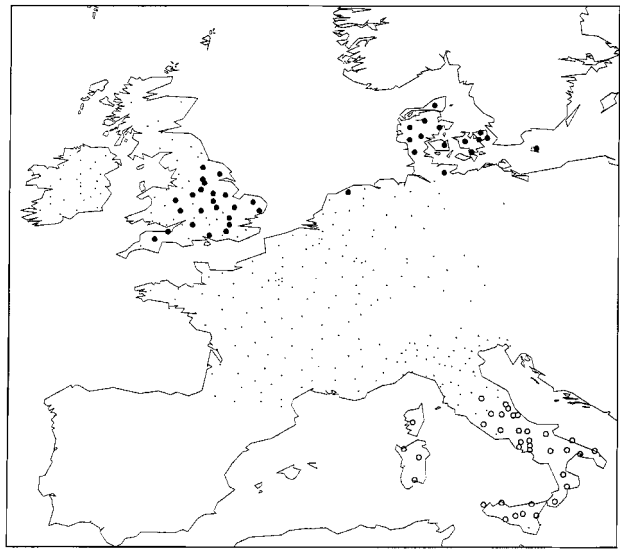

Figure 5. Map of the top (spatial clusters of high mortalities; filled circles) and bottom (spatial clusters of low mortalities; hollow circles) 10% of the ranked total permutational probabilities of G_i* for ovarian cancer mortality. The remaining localities are shown as dots.

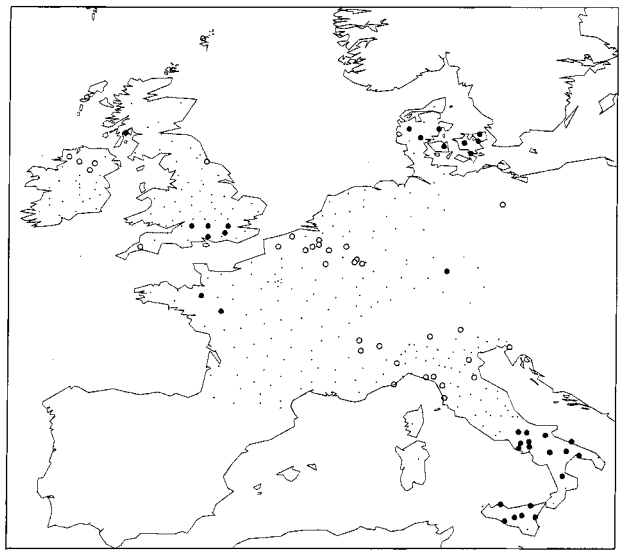

Figure 6. Map of the top (positive local SA; filled circles) and bottom (negative local SA; hollow circles) 10% of the average ranks of conditional permutational probabilities for I_i statistics for 40 cancer mortalities. The remaining localities are shown as dots.

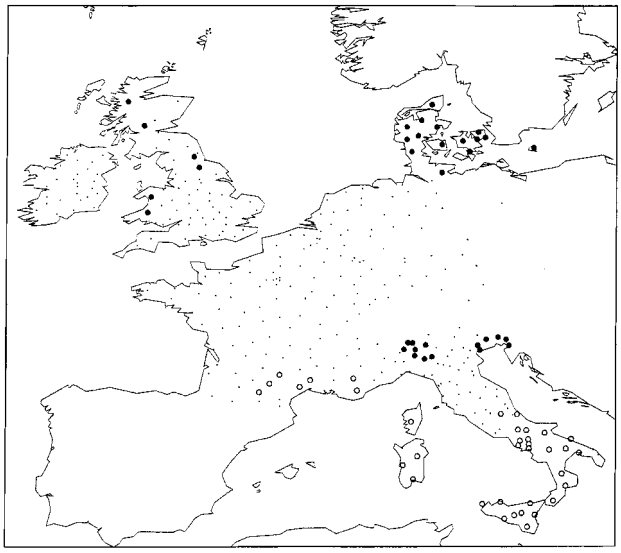

Figure 7. Map of the top (spatial clusters of high mortalities; filled circles) and bottom (spatial clusters of low mortalities; hollow circles) 10% of the average ranks of total permutational probabilities for G_i^* statistics for 40 cancer mortalities. The remaining localities are shown as dots.

Northern Italy in Figure 6, which represent extreme average negative local SA, may result from the juxtaposition of homogeneous high cancer (the filled black circles in Figure 7) with nearby dissimilar locations.

Discussion

The strong spatial autocorrelation of the cancer mortalities indicates clearly that these variables are not spatially random. This would imply that the responsible factors are themselves SA, but does not directly point to the identity of these factors. These could be endogenous (examples: ethnic genetic differences; ethnic or regional cultural differences, such as in food, drink, sexual practices; etc.), or exogenous (e.g., patterning of areas of industrial pollution, intensity and duration of sunlight, regional differences in recording and reporting the vital statistics, etc.). To discover which of these factors might be important requires further investigation.

The patterns are relatively few in number and there is significant agreement of cluster membership between the one- and two-dimensional *I*-correlograms, and also between I and c. This would indicate that there are fewer spatially patterned causal agents than the number of cancers studied.

The cancer mortality data are among the most strongly spatially autocorrelated observational data we have encountered in over 20 years of research into SA. It is well known (see, for example, [16]) that SA in a pair of variables spuriously increases the nominal significance of the correlation between the two variables, leading to inferences about common causation that may be unwarranted. Many of the cancers are strongly correlated with each other. For example, female oesophageal and lung cancers are correlated at $r = 0.775$, which is significant even after the degrees of freedom are reduced to remove the effect of SA (by the procedure of [16]) from 353 all the way down to 7. Other high correlations include cancers with separate male and female rates (e.g., cancers of the stomach, brain, and gall bladder) and cancers that are functionally correlated (e.g., female breast, ovarian, and cervical cancers or male oral, oesophageal, and larynx cancers). In these cases the correlation may well be epidemiologically meaningful.

Cancers whose mortality surfaces are strongly correlated must have similar spatial correlograms. The reverse is not true; it is quite possible to have similar spatial patterns in uncorrelated data, especially with regard to patch size. An example in point is shown by the three mortality surfaces for female oesophageal cancer, ovarian cancer, and cervical cancer. The average correlation of female oesophageal cancer with ovarian and cervical cancer is 0.426. Yet ovarian and cervical cancer correlate at 0.801. These two have very similar patterns of high mortalities (in Denmark and scattered in the British Isles), and low mortalities in much of Italy. By contrast, female oesophageal cancer mortalities are high in Ireland and Britain, and low in central and south Italy. The mean one-dimensional correlograms for each cluster (Figure 2) indicate 2 major patch sizes (approximately 650 and 872 km). Patch size could signify a variety of phenomena [20]. It may indicate the grain of carcinogenic external factors, or of ethnic mixtures, or of cultural practices. Alternatively, patch size could mark distances between political boundaries that mark different standards for mortality reporting. This latter is potentially a major problem for any study which uses data collected from a wide variety of political entities. Potential differences in reporting were a primary concern of the authors who compiled the cancer atlas [14].

Detailed examination of the members of the clusters from the one-dimensional correlograms showed that the strong clines have an average correlation of 0.622 between all pairs of their mortality surfaces. This suggests that most of the patterns will be similar, and closer inspection revealed that a north-south cline with highs in the north and lows in the south is the common theme of the strong clines. As we have seen, there are variations on this theme: cervix and ovary have high mortalities in both Denmark and the British Isles, whereas those of female oesophageal cancer are high only in Ireland and Britain. The weak clines are not nearly as highly correlated. Average correlation between all pairs of mortality surfaces grouped as weak clines is 0.322. The patchy patterns show very little correlation. The average of the ab-

solute correlations is 0.161. In the strong and weak clines we did not compute absolute correlations since their *r*-values are all positive.

When we examine the clustering result of the two-dimensional correlograms, we find that all but one of the group 4 surfaces corresponds to the strong clines of the one-dimensional correlograms. Group 4 exhibits an average correlation of 0.625. Group 1 has a lower average correlation of 0.448. It is a mixture of patchy and clinal mortality surfaces. The other three clusters describe mostly patchy surfaces. Remember that the locations of patches of high and low mortalities need not coincide to yield similar correlograms. The surfaces in the cluster need only to agree in patch size. The three clusters yield average absolute correlations ranging from 0.169 to 0.499. Although a common patch size hints at the possibility that the causal agents of two cancers are the same, this need not be true. Surfaces with different appearing patterns can yield similar correlograms.

Two potential sources of bias in other studies are not of serious concern in this one. Distant noncontiguous areas, such as distant island areas, will force the upper distance classes of the correlogram to reflect mostly the difference, with respect to the variable studied, between the mainland and the outlying areas. The noncontiguous areas in this study, Britain, Ireland, Corsica, Sardinia, Sicily, and the Danish islands, are all close enough to the mainland or to each other so that their links to the main area fall into the two lowest distance classes. The second source of bias is related to the first. If interlocality distances in a region are quite disparate, then the structure shown in the finely sampled areas would have no counterpart in the coarsely sampled areas.

Although there is some disparity in size of the registration areas, (the largest areas are in West Germany, the smallest in Great Britain and Ireland), the sampling points are fairly evenly distributed, as can be seen in any of Figures 4 to 7. Although small-scale (local) positive SA is often easy to explain, small-scale negative SA may be more informative or interesting. For example, in Figure 4 the high positive local SA in Denmark and southern Italy could be due to large country (or regional) differences in reporting; the negative local SA in Ireland cannot be due to this type of large scale regional effect. It could indicate extreme differences in local (county) reporting, but minimally indicates a phenomenon that should be investigated.

We have carried out a similar study of cancer mortalities in central Europe. These are based on an atlas [15] prepared in a comparable manner to the present data set. The records represent the years 1983 to 1987, almost a decade later than the western European mortalities. The two studies agree in showing strong SA in both one- and two-dimensional correlograms. However, the two studies do not agree in the one-dimensional correlogram cluster membership for various cancers, and only partially so for two-dimensional correlogram clusters. This lack of agreement suggests that different or identical but differently patterned causal agents are at work in the two regions. Of the putative factors bringing about differences in mortality, ethnic, genetic, or cultural differences would clearly be cases in point. In a cognate study [32], we have shown that ethnohistorical affinities play a significant role in determining patterns of differences in cancer mortality in Europe. We could also obtain the observed results if the two regions differ markedly in their spatial patterns of reporting vital statistics.

Acknowledgements

This paper is contribution No. 1018 in Ecology and Evolution from the State University of New York at Stony Brook. This research was supported by grants DEB9220538 and SBR9419349 from the National Science Foundation to Robert R. Sokal. We thank Barbara A. Thomson for computational advice and assistance, and two anonymous reviewers for their suggestions and comments.

References

1. Sokal RR, Smouse PE, Neel JV. The genetic structure of a tribal population, the Yanomama Indians XV. Patterns inferred by autocorrelation analysis. Genetics 1986; 114: 259–287.
2. Sokal RR, Oden NL, Barker JSF. Spatial structure in *Drosophila buzzatii* populations: Simple and directional spatial autocorrelation. Am Nat 1987; 129: 122–142.
3. Epperson BK. Spatial patterns of genetic variation within plant populations. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds), Population genetics and germ plasma resources in crop development. Sunderland, MA: Sinauer Associates, 1990: 229–253.
4. Epperson BK. Spatial Structure of genetic variation within populations of forest trees. New Forests 1992; 6: 257–278.
5. Barbujani G, Sokal RR. Genetic population structure of Italy. I. Geographic patterns of gene frequencies. Hum Biol 1991; 63: 253–272.
6. Falsetti AB, Sokal RR. Genetic structure of human populations of the British Isles. Ann Hum Biol 1993; 20: 215–229.
7. Setzer RW. Spatio-temporal patterns of mortality in *Pemphigus populicaulis* and *P.populitransversus* on cottonwoods. Oecologia 1985; 67: 310–321.
8. Jones, JS, Selander RK, Schnell GD. Patterns of morphological and molecular polymorphism in the land snail *Cepaea nemoralis*. Biol J Linn Soc 1980; 14: 359–387.
9. Caugant D, Jones JS, Selander RK. Morphological and molecular polymorphism in *Cepaea nemoralis* in the Spanish Pyrenees. Genetica 1982; 57: 177–191.
10. Epperson BK, Clegg MT. Spatial-autocorrelation analysis of flower color polymorphisms within sub-

structured populations of morning glory (*Ipomoea purpurea*). Am Nat 1986; 128: 840–858.
11. Sokal RR, Uytterschaut H. Cranial variation in European populations: A spatial autocorrelation study at three time periods. Am J Phys Anthro 1987; 74: 21–38.
12. Higginson J, Muir CS, Muñoz N. Human cancer: Epidemiology and environmental causes. Cambridge University Press, 1992.
13. Mehnert WH, Smans M, Muir CS, Möhner M, Schön D. Atlas of Cancer Incidence in the Former German Democratic Republic 1978–1982. Lyon: IARC Scientific Publications No. 106, 1992.
14. Smans M, Muir CS, Boyle P. Atlas of Cancer Mortality in the European Economic Community. Lyon: IARC Scientific Publications No. 107, 1992.
15. Zatonski W, Smans M, Tycynski J, Boyle P. Atlas of Cancer Mortality in Central Europe. Lyon: IARC Scientific Publications No. 134, 1996.
16. Clifford P, Richardson S, Hémon D. Assessing the significance of the correlation between two spatial processes. Biometrics 1989; 45: 123–134.
17. Sokal RR, Oden NL. Spatial autocorrelation in biology I: Methodology. Biol J Linn Soc 1978a; 10: 199–208.
18. Cliff AD, Ord JK. Spatial Processes. London: Pion, 1981.
19. Sokal RR, Oden NL. Spatial autocorrelation in biology II: Some biological implications and four applications of evolutionary and ecological interest. Biol J Linn Soc 1978b; 10: 229–249.
20. Sokal RR. Ecological parameters inferred from spatial correlograms. In: Patil GP, Rosenzweig ML (eds), Contemporary quantitative ecology and related ecometrics. Fairland, MD: International Co-operative Publishing House, 1979: 167–196.
21. Upton GJG, Fingleton B. Spatial data analysis by example. Vol 1 : Point pattern and quantitative data. New York: John Wiley, 1985.
22. Wartenberg DE. SAAP – A spatial autocorrelation analysis program. Setauket, New York: Exeter Software 1989.
23. Oden NL. Assessing the significance of a special correlogram. Geog Anal 1984; 16: 1–16.
24. Späth H. Cluster-Formation und -Analyse. Munich: R. Oldenbourg Verlag, 1980.
25. Oden NL, Sokal RR. Directional autocorrelation: An extension of spatial correlograms in two dimensions. Sys Zool 1986; 35: 608–617.
26. Hubert L, Arabie P. Comparing partitions. J Class 1985; 2: 193–218.
27. Getis A, Ord JK. The analysis of spatial association by use of distance statistics. Geog Anal 1992; 24: 189–206.
28. Anselin L. Local indicators of spatial association – LISA. Geog Anal 1995; 27: 93–115.
29. Tiefelsdorf M, Boots B. A note on the extremities of local Moran's I_is and their impact on global Moran's I. Geog Anal 1997; 29: 248–257.
30. Sokal RR, Oden NL, Thomson BA. Local spatial autocorrelation in a biological model. Geog Anal 1998a; 30: 331–354.
31. Sokal RR, Oden NL, Thomson BA. Local spatial autocorrelation in biological variables. Bio J Linn Soc 1998b; 65: 41–62.
32. Sokal RR, Oden NL, Rosenberg MS, DiGiovanni D. Ethnohistory, genetics and cancer mortality in Europeans. Proc Nat Acad Sci USA 1997; 94: 12728–12731.

Address for correspondence: Robert R. Sokal, Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, New York 11794-5245, USA
Fax: +1 516 632-7626
E-mail: sokal@life.bio.sunysb.edu