DUAL TECHNIQUES FOR THE DETERMINATION OF SPATIAL CLUSTERING OF MORTALITY IN MONTREAL, QUEBEC - 1972

William G. Nisen and Anne M. Hunt
Harvard Laboratory for Computer Graphics
and Spatial Analysis
520 Gund Hall, 48 Quincy Street
Cambridge, Massachusetts 02138

I. Introduction

The determination of spatial clustering in health related studies is crucial for optimizing the design and implementation of prophylactic measures. In addition, by noting where high morbidity and mortality rates predominate, new hypotheses may be generated which would help unravel some of the more perplexing contemporary etiological puzzles.

As an outgrowth of a larger study (Nisen, 1976), standardized mortality rates for neoplasms of all types, excluding lung and breast (ICD, eighth revision, nos. 140-163, 170-173, 180-199, 200-239), for the males in the city of Montreal, 1972, two techniques to determine spatial clustering were employed. The first involved the use of Moran's "I" statistic to statistically determine spatial autocorrelation (Moran, 1950). The second approach utilized computer assisted techniques to produce maps of the entire city for visual analysis, and of selected sets of contiguous tracts as a means of identifying homogeneous areas of high mortality rates.

II. Data Collection and Standardization

The examination of mortality due to neoplasms was

facilitated by the existence of a mortality data base maintained by the Government of Quebec and the City of Montreal. Death certificates are coded in Montreal then geocoded to the census tract of last residence and stored on the data processing system of the Quebec provincial data processing system under the auspices of the Centre Informatique.

Although the mortality information was computerized to facilitate storage and retrieval, considerable time delay was experienced between the request for information and actual examination. This delay was due to problems associated with making operational a new data base management system.

Upon receipt of the mortality tabulations, manual extraction and compilation by census tract produced a file of raw mortality data. This data was divided into deaths by sex, and then age standardized using Liddell's method (Liddell, 1965).

III. Statistical Determination of Positive Spatial Autocorrelation

Spatial autocorrelation may be visualized in the following manner. Given a finite area which is divided into n areal units, if for every pair of areal units i and j, samples yield variates, x_i and x_j, which are uncorrelated, then there is no spatial autocorrelation in the areal system on x. If the samples are not all pairwise uncorrelated, then spatial autocorrelation is said to exist (Cliff and Ord, 1973, 2).

A set of tests have been developed (Moran, 1950, Geary, 1954, Dacey, 1965, and Cliff and Ord, 1969) to detect spatial autocorrelation for nominal, ordinal, and interval data. There are three main test statistics available for use with interval scale data: Geary's "c" statistic, Moran's "I" statistic and the Cliff and Ord statistic.

Both Moran and Geary utilized a binary weighting mechanism in their spatial autocorrelation formulae. Cliff and Ord further extended upon the results of Moran and Geary to incorporate generalized weights (Cliff and Ord, 1975, 151). Since the data are recorded by discrete areal units (census tracts), no

attempt is made to extend beyond binary weighting. Furthermore, the "I" statistic has been judged superior to the "c" statistic; hence, our analysis will be limited to this index (Cliff and Ord, 1975). The last remaining question to be addressed is whether the results obtained from either test for spatial autocorrelation are statistically significant.

Cliff and Ord (1975) state that for the "I" statistic it is necessary to consider two distinct forms for the null hypothesis. These are:

1. Normality assumption: that the values for the set of variates x are the results of n independent drawings from a normal population (or identical populations); and
2. Randomization assumption: whatever the underlying distribution of the random variables, we consider the observed value of I or c within the set of all n! possible values which I or c could take on if the $[x_i]$ were randomly permuted around the areal system (Cliff and Ord, 1975, 152-153).

By reviewing the nature and form of the standardized mortality data as distributions in histogram form, it is apparent that all spatial autocorrelation analysis must be performed assuming the second form (randomization assumption) of the null hypothesis.

The calculated value for Moran's "I" statistic for neoplasms of all types in males, excluding lung and breast, was 0.19277 which is not significant at the 0.05 level.

Even if there was a significant amount of spatial autocorrelation, this type of analysis is limited by its inability to specify the location of the spatial clustering. This deficiency is remedied by use of a more traditional medical geographic approach, visual map analysis, and with recent developments in computer cartography, this technique is more powerful than ever before.

IV. Visual Map Analysis: Problems and a Potential Solution

Visual map analysis has been used extensively in

medical geographic research. The work of Snow (1885) during the cholera outbreak of 1849 in London serves as an example of the types of beneficial findings that can result from an "enlightened" inspection of a mortality or morbidity map. However, in the present study the data was collected on an areal basis, thus necessitating a choropleth display as opposed to individual occurrences which can be investigated via point pattern analysis.

Although associations in certain instances can be ascertained from choropleth maps through visual analysis, there still exist perceptual problems from the interpreter's perspective (Monmonier, 1977). The most important facet of the problem of perceiving areal associations on a mapped surface is that of map complexity. The problem of complexity has been reviewed and investigated by many individuals (Julesz, 1975; Monmonier, 1974; Muller, 1975, Olson, 1975; and Bouchard, 1976).

To evaluate the map complexity problem, two types of cartographic outputs were utilized. A standard two-dimensional shaded map was produced by the POLYPS program developed at Harvard. As evidence in Figure 1, this map is confusing and difficult to interpret. Figure 2 represents the same data, but it makes the perception of high mortality rates much easier to discern. This output was produced by the PRISM program also developed at Harvard. An additional useful characteristic of the PRISM program is the ability to individually represent all the census tracts.

Due to its inherent capabilities to communicate the standardized mortality rates more effectively, the PRISM program was selected for use in an attempt to depict sets of contiguous tracts with relatively high mortality rates and which were homogeneous with respect to these rates, i.e., low variance.

The first step in determining the clusters of high mortality rates was to rank all contiguous sets of tracts by their mean value. Table 1 is an abbreviated representation of this table which illustrates not only the mean values, but the variance as well. The asterisks indicate the contiguous sets of tracts that were selected for mapping.

Figure 3. is a composite cartographic representation of the six sets of tracts where high mortality rates predominate. The small cluster of tracts along Montreal's waterfront is interesting and is deserving of a closer examination.

V. Summary and Conclusions

The increase in chronic disease mortality is one of the most serious health problems faced by the developed nations. This problem becomes acute in an intra-urban context. As a means of elucidating the spatial structure of the Standard Mortality Ratios for neoplasms of all types, excluding lung and breast, (a representative chronic disease) for the male population of the City of Montreal, 1972, two types of analysis were employed. The use of a statistical determination of spatial autocorrelation, specifically Moran's "I" statistic, indicated that no significant amount of spatial clustering of high rates was present in the City of Montreal on a census tract level. Even if a significant amount of spatial autocorrelation was present, this type of analysis has a major deficiency: it cannot specify the location of the high mortality rate clusters. To attempt to solve this problem, a visual analysis of two types of cartographic output were undertaken. To minimize the problems associated with interpreting a two-dimensional shaded map, a program that produces pseudo three-dimensional representations was used. It was found that the combination of interactive computer cartographic programs coupled with rudimentary manipulative and query capabilities result in a tool that can be of tremendous value to the medical geographer and epidemiologist alike as a means of discovering health problem areas for further examination and analysis.

References

Bouchard, Diane. "Spatial Autocorrelation and the Analysis of Health Data." Analysis of Non-Vectored Diseases in Montreal, Report 3. Montreal: Department of Geography, McGill university, 1976.

Cliff, A.D. and Ord, J.K. "The Problem of Spatial Autocorrelation." London Papers in Regional Science, Vol. 1. Edited by A.J. Scott. London: Pion Ltd., 1969.

Cliff, A.D. and Ord, J.K. Spatial Autocorrelation. London: Pion Limited, 1973.

Cliff, A.D., Haggett, P., Ord, J.K., Bassett, K. and Davies, R. Elements of Spatial Structure. London: Cambridge University Press, 1975.

Dacey, M.F. "A Review on Measures of Contiguity for Two and K-Colour Maps." Spatial Diffusion Study, Technical Report No. 2. Evanston, Ill." Department of Geography, Northwestern University, 1965.

Geary, R.C. "The Contiguity Ratio and Statistical Mapping." The Incorporated Statistician, Vol. 5 (1954), 115-45.

Julesz, Bela. "Experiments in the Visual Perception of Texture." Scientific American. Vol 232 (April, 1975), 34-43.

Liddell, F.D.K. "The Measurement of Occupational Mortality." British Journal of Industrial Medicine, Vol. 17 (1965), 228-233.

Monmonier, Mark S. "Measures of Pattern Complexity for Choroplethic Maps." The American Cartographer, Vol. 1 (1974), 159-169.

Monmonier, Mark S. "Maps, Distortion, & Meaning." Commission on College Geography Resource Paper, No. 75-4. Washington, D.C.: The Association of American Geographers, 1977.

Moran, P.A. "Notes on a Continuous Stochastic Phenomenon." Biometrika, Vol. 37 (1950), 17-23.

Muller, J.C. "Definition, Measurement, and Comparison of Map Attributes in Choroplethic Mapping. "Proceedings of the Association of American Geographers, Vol. 7 (1975), 160-164.

Nisen, W.G. "An Examination of Intra-Urban Mortality Patterns in Montreal: A Spatial Analytical Approach." Unpublished thesis, Department of Geography, McGill University, 1976.

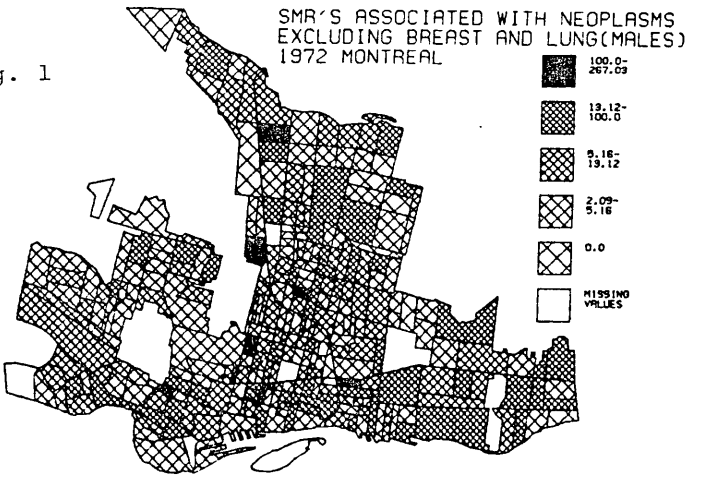
Olsen, Judy M. "Autocorrelation and Visual Map Complexity." Annals of the Association of American Geographers,

Vol. 65 (1975), 189-204.

TABLE 1
Subjective Identification of Sets of Contiguous Tracts with High Mean Mortality Values and Low Variance

TRACT ID	VARIANCE	MEAN	TRACT ID	VARIANCE	MEAN
65	11844.513	109.023	205	155.548	27.260
49	2268.855	59.325	10	303.731	26.753
63	1040.736	58.516	5	680.954	26.665
68	2436.258	55.842	* 142	49.756	26.657
44	457.618	49.630	127	627.882	26.571
138	2045.321	49.418	210	179.876	25.878
67	9333.379	48.402	17	4155.976	24.786
136	4050.622	46.288	* 249	46.479	24.754
275	971.437	43.453	193	426.987	24.648
130	1384.416	39.408	13	988.614	24.634
131	4371.019	38.820	215	395.678	24.072
129	1227.196	38.124	* 29	68.844	23.838
137	873.216	37.144	171	479.685	23.338
51	5853.689	36.311	9	511.927	22.902
134	1301.911	36.082	263	205.731	22.890
164	781.445	36.010	139	225.386	22.656
15	873.921	35.652	219	374.253	22.497
203	607.175	33.100	143	293.428	21.018
273	6749.891	32.792	2	257.022	20.424
32	9247.781	32.304	191	322.475	20.190
211	1324.706	31.812	* 21	91.149	20.010
172	644.649	31.718	128	111.297	19.691
207	5333.825	31.076	214	132.479	19.089
216	4770.407	30.606	7	5152.435	18.384
187	478.480	30.545	* 26	54.215	18.178
30	213.566	29.259	141	255.605	18.094
217	441.686	28.788	16	3188.376	17.670
132	425.854	28.626	11	196.096	17.433
8	185.529	28.480	46	468.706	17.106
3	701.707	27.945	* 64	41.801	16.648

Fig. 1



SMR'S ASSOCIATED WITH NEOPLASMS
EXCLUDING BREAST AND LUNG(MALES)
1972 MONTREAL

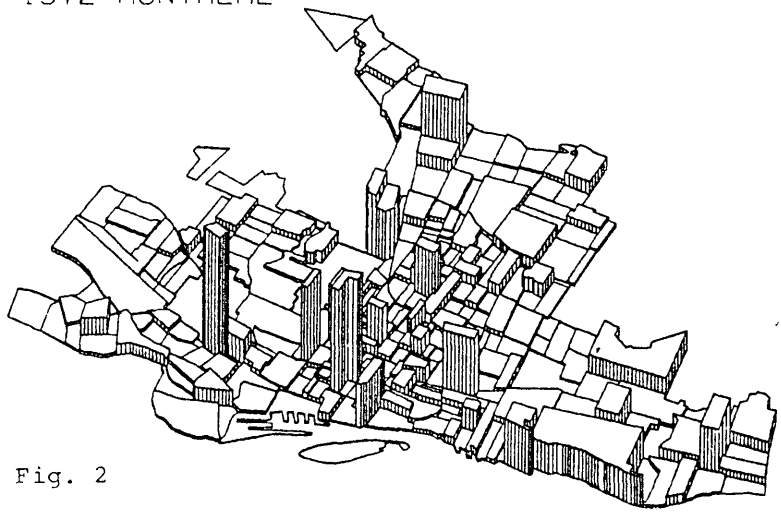


Fig. 2

SMR'S FOR CONTIGUOUS TRACT GROUPS
(VARIANCE LESS THAN 100)

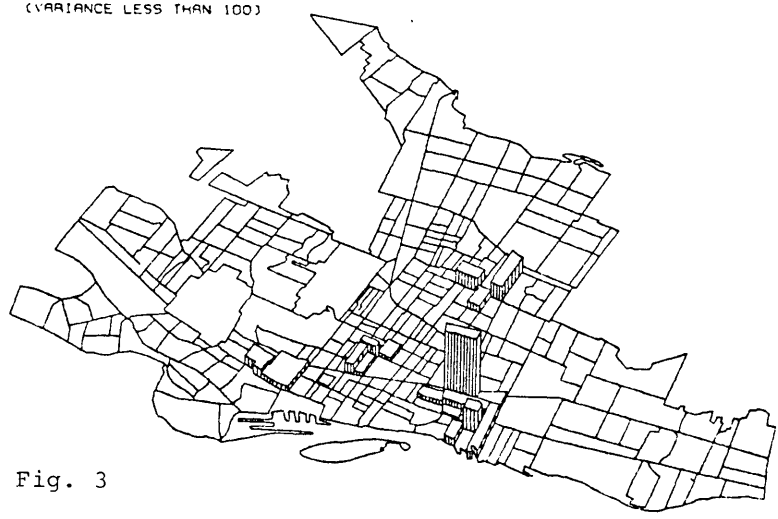


Fig. 3