Annu. Rev. Public Health 2003. 24:43–56
doi: 10.1146/annurev.publhealth.24.012902.140843

First published online as a Review in Advance on December 2, 2002

PUBLIC HEALTH, GIS, AND SPATIAL ANALYTIC TOOLS

Gerard Rushton
Department of Geography, The University of Iowa, 316 Jessup Hall, Iowa City, Iowa, 52242-1316; email: gerard-rushton@uiowa.edu

Key Words spatial analysis, disease maps, exploratory data analysis, geocoding, Monte Carlo simulation

■ **Abstract** We review literature that uses spatial analytic tools in contexts where Geographic Information Systems (GIS) is the organizing system for health data or where the methods discussed will likely be incorporated in GIS-based analyses in the future. We conclude the review with the point of view that this literature is moving toward the development and use of systems of analysis that integrate the information geo-coding and data base functions of GISystems with the geo-information processing functions of GIScience. The rapidity of this projected development will depend on the perceived needs of the public health community for spatial analysis methods to provide decision support. Recent advances in the analysis of disease maps have been influenced by and benefited from the adoption of new practices for georeferencing health data and new ways of linking such data geographically to potential sources of environmental exposures, the locations of health resources and the geodemographic characteristics of populations. This review focuses on these advances.

INTRODUCTION

The analytic capabilities of geographic information systems, available in a few cases as fully integrated systems, but more commonly as loosely coupled software systems, have developed rapidly in recent years. Public health is now presented with the opportunity to examine key relationships between the health characteristics of populations and both human and physical environmental characteristics. Some authors suggest that a new discipline of spatial epidemiology now exists, recently described (22, p. v) as being "concerned with describing, quantifying, and explaining geographical variations in disease, especially with respect to variations in environmental exposures at the small-area scale." Although few implemented examples of such systems currently exist outside of the area of national disease and mortality mapping systems, this is clearly the direction in which spatial analytic tools will be used in the future. Some descriptions of such systems in various stages of implementation can be found in the public health literature (3, 62, 74). This period, too, has seen the beginning of a more critical literature that points out

the limitations of GIS in answering important public health questions that have a spatial dimension (25, 33).

MAPPING DISEASE RATES

With the increased availability of spatial coverages of individual disease incidences in GIS, statements that were commonly made a decade ago to the effect that disease data are available for areas but rarely as point patterns, now appear quite dated. Several authors have pointed out that choropleth maps (where areas are shaded according to a defined category in a map legend) can be seen as smoothed maps of disease data with the areas chosen acting as filters—albeit of different spatial size. Viewed in this light, such maps are seen by many as inferior representations of the basic data. The traditional public health databases in which observations of diseases are columns and areas are rows reflect the practical requirements for databases prior to GIS. GIS bring the possibility not only of storing geographic information differently but of processing it differently too (63).

With the recognition that different patterns exist at different spatial scales and that maps of any disease should appropriately be examined at different scales, interest has grown in using density estimation methods first popularized by Bithell (7) and usefully compared with other methods in a recent comprehensive review of disease-mapping methods (8). This method produces disease maps in which rates vary continuously across the map. Such maps have several advantages over maps of disease rates for discrete areas. Conceptually, they more closely replicate the facts on the ground. We do not expect disease rates to change suddenly along the borders of the statistical areas for which we traditionally collected and stored disease data. If the data have been geographically coded to their latitude and longitude coordinates, or to quite small areas such as zipcodes, census tracts, census block areas, or to other small postal code areas, they can be spatially aggregated in very flexible ways according to the needs of the user. In most approaches to density estimation, disease rates are computed for circular areas around grid points or for the small areas whose centroids are found there. Sometimes the values of the observations are weighted the same for all observations in the circle (box-car filters) and sometimes they are weighted according to some inverse-distance function from the center of the circle. In Figure 1, numerals indicate the number of births from 1996 through 1998 within 0.8 miles of a grid point. The number of infant deaths in the same filter areas (not shown) are also summed. The legend indicates infant mortality rate (deaths per thousand live births) as interpolated from the grid points to the area in general.

Talbot et al. (69) posit that under conditions of large spatial variability in the density of health events, spatial filters should be adjusted in geographical size in relation to the density of observed events. They argue that the purpose of the spatial filter is to remove random noise in the data. Random noise is a function of sample size. Therefore, choosing the same sample size around each grid point produces a map in which the same degree of random noise is removed. Stated differently, the

estimates of the standardized incidence rate (SIR) values mapped have the same variances across the map. Talbot et al. introduce the idea of an adaptive spatial filter and illustrate its use with birth weight data from New York State. Care is needed when interpreting maps made with spatially adaptive filters because the spatial resolution of the data used varies across the map. Talbot et al. (69) illustrate this fact with a map of the state of New York showing that the size of the filter required to capture 250 births varied from less than 1 km to more than 40 kms.

Disease data are now presented for a far greater variety of types of areas than in the past. Driving this trend is an interest in measuring health and disease for special populations and areas. Facilitating the process is the recognition that areas can be defined in GIS based on different principles—spatial buffer areas of any selected length from defined points or lines, areas with given socio-demographic characteristics, areas with given densities of people or resources, and many more. In Figure 2, at each grid point on a 0.4-mile grid, the difference between the infant mortality rate for 1996 through 1998 is subtracted from the infant mortality rate for 1993 through 1995 as computed for the area (filter area) within 0.8 miles of the grid point. The map shows the difference in the rates as estimated at all grid points and interpolated using an inverse distance-weighted function.

Reservations are being expressed in the recent literature that some results of spatial analysis of disease rates may reflect inaccuracies in population counts and in the registration of disease cases in registers (6, 70), whereas in other cases new methods have been developed to estimate demographic data for small areas (21).

SPATIAL SMOOTHING

Density estimation is an example of spatial smoothing first used in public health as a means to provide more statistically reliable directly adjusted disease or death rates by aggregating contiguous counties in national atlases of disease (17, 58, 59). By developing tables identifying the contiguous counties for each county, the task of making smoothed rates for each county and its contiguous counties—a topological spatial buffer function—could be accomplished without the use of GIS. Before the advent of GIS technology, this was an achievement. Such atlases when made today, however, should attach the caveat, "the disease rates on this map have been estimated using spatial filters of drastically different shapes and sizes."

In the GIS literature smoothing data for irregular data units are known as RESEL (resolution element) filtering and are viewed as an undeveloped field of spatial analysis (53). This literature shows that many of the operational definitions used to apply methods of spatial filtering originally developed for raster (image) data to the polygons for which health data are commonly available are arbitrary and perhaps not as sensible as other definitions developed for RESEL-type data.

Kafadar illustrated many of the principles used in spatial smoothing applications (38). Not all spatial smoothers are applied to disaggregated data as Kafadar

and others have shown (35, 37). Spatial smoothing is often treated in the literature as a distinctive approach, whereas Bithell observed that (8, p. 2204) "smoothing parameters are intrinsic to the density estimation problem." Other approaches have modeled spatio-temporal patterns of disease rates as a generalized additive mixed model (51) or have attempted to "provide a balance between smoothness and the maintenance of discontinuity" for analyses aimed at making resource allocation decisions (47). Both papers attempt to model spatial patterns of disease but not in the sense of using relevant predictors of attributes of areas that contribute to increased disease risk. Rather, they generalize features of the spatial pattern such as local correlation effects or discontinuity components. These approaches acknowledge the problem that although smoothing generally results in more reliable statistical rates, it will often remove small areas where rates are truly different (clusters) or disguise linear features where rates change sharply. Therefore, one use of smoothed maps is to identify distinct areas or geographic features (34). Recent years have seen the application of "head-banging algorithms" (38, p. 3173; 54, 58), pioneered by Hansen (32) for this purpose. The key idea behind this approach is that median values of local rates might better represent the true disease rate surface than mean rates. In an analysis by Mungiole et al. (54), the medians along iteratively adjusted and roughly linear transects across each filter area were weighted by the number of observations that contributed to the rate at each observation point. These studies should be seen as the beginning of research in public health to develop mapping methods that are more successful in describing the true disease rate surface than current methods, which are known from simulation studies to often fail to identify areas of significantly high rates and to often spuriously identify areas as having high rates when they do not (26).

In applying a spatial smoothing algorithm to any area, selecting parameters appropriate to the spatial scale of the database and to the expected disease process is crucial. An analysis of the spatial pattern of cholera in a region of Bangladesh (1, p. 1018) used a "7 x 7 pixel moving window to compute average cholera rates." The pixels were 30-meter squares. No information was provided on the numbers of people living within the 210-meter-square areas for which cholera rates were computed, but results of analyses showed very low correlations with known risk factors for cholera, thus confirming the suspicion that the "smoothed" disease rates for such small geographic areas did not adequately capture the true spatial pattern of disease rates. In rural areas such as this, the density of population may not provide a sufficiently large number of observations for valid disease rates to be computed from information formed from the contiguous areas required to make a valid disease rate. If this problem cannot be solved either by aggregating observations over a longer period of time or making larger spatial filters—neither of which was possible in Ali et al. (1)—then smoothing over noncontiguous spatial areas should be considered. A quite different approach to spatial smoothing involves the aggregation over space of individuals who are in defined spatial relationships with respect either to disease risk factors or to resources for screening or treatment. Rushton & West (66) categorized all women in Iowa who had been diagnosed

with localized breast cancer by their distance to the closest place with a radiation treatment facility. They showed a decline in the proportion of women who selected lumpectomy with radiation over radical mastectomy with increasing distance from the closest radiation treatment facility.

METHODS OF ADJUSTING DISEASE RATES FOR COVARIATES

The spatial pattern of any disease is affected by the spatial distribution of demographic factors such as age and sex. It is standard practice to remove these affects by age-sex adjustment procedures (64). When the purpose of such adjustment is to compare disease rates between populations that have different age-sex characteristics, the direct method of age adjustment is recommended, provided data are available to make valid estimates of disease rates for age-sex subgroups of the population for small areas—a requirement that is difficult to meet for most small areas (60). When the purpose is to assess the need for resources to affect changes in the rates, the indirect method of age adjustment in which national or regional age-sex disease rates are applied to the local demographic characteristics of areas should be used. Indirectly adjusted rates can be estimated more accurately for small areas (2, 28).

This distinction has not always been followed in the recent literature, and a number of recent atlases have used the method of direct age-adjustment to create maps ostensibly for the purpose of guiding decisions on resource allocation (17). Goldman & Brender (28), however, in a series of simulated experiments, provided evidence that results of the two adjustment procedures differed by only about 8%, and they concluded that either of the methods can be used in public health data analysis. This result, if accepted, is significant in that the data requirements for computing directly adjusted rates are far more demanding than the requirements for computing indirectly adjusted rates. The literature is in full agreement that disease rates for small geographic areas are far more likely to be robust when computed by the indirect method than by the direct method (3). Before the public health community accepts this latest comparison of the two methods of age adjustment for disease maps, it must still come to terms with the logical, empirical, and persuasive case made by Kleinman in 1977 (40) that the difference between the two methods is large and important in relation to the purpose for which the adjustment is made.

ADJUSTING FOR SOCIAL AND ECONOMIC DEPRIVATION

In addition to the obvious need to adjust observed disease incidences for differences in the age and sex structure of local populations, there is often a need to adjust for other covariates that influence a person's risk of having the disease. In recent years there has been a marked increase in interest in adjusting for measures of social deprivation or different levels of social capital (16, 20, 41, 48, 73).

Siegel et al. (68) found higher age-adjusted rates of pertussis infection were associated with higher proportions of residents below poverty levels. Järup et al. (35) showed that prostate cancer risks were not associated with deprivation in the United Kingdom. The results of spatial analyses of prostate cancer mortality in the United States between 1970 and 1989 (36) generally support the same conclusion, although the significantly higher prostate cancer rates of American Blacks remain unaccounted for Brown et al. (12) have taken such studies one step further. They use information on the relationship between cancer incidence and social and economic deprivation to develop and use a geodemographic approach to the design of a suitable control group to assess the effects of interventions to reduce the incidence and mortality rates of specific cancers. Preventive health resources can be targeted to areas where concentrations of socially and economically deprived individuals live.

A number of recent studies (35, 38, 71) conclude by noting that the long latency period for many diseases, such as cancer, and the effects of migration are likely to reduce the spatial differences in disease rates that would otherwise exist.

ADJUSTING FOR "NOISE" IN DISEASE MAPS

After adjusting disease rates for age and sex, a second problem faced by the public health worker is adjusting the map for "noise." All maps of diseases that portray disease rates for discrete areas (choropleth maps) are likely to have their extreme rates (high and low) in areas with small populations at risk. This is especially the case where the disease is rare and when it is measured over a short period of time. Similarly, smoothed maps, developed separately for each state, as is sometimes done, can be expected to show more extreme rates along border areas where the filter area is curtailed by the state border or water body in the case of littoral states. The decade of the 1990s saw major developments in Empiric Bayes methods (18, 19, 50) as the preferred methodology to shrink observed rates differentially toward the mean of the distribution of rates in proportion to their expected variability based on the number of observations in the small areas. In addressing the public health community of the United Kingdom recently, the faculty of the UK Small Area Health Statistics Unit noted (3, p. 297), "This technique is the standard in disease mapping" and cited Clayton & Kaldor (18) in support. The past few years, however, have seen the development and increased use of full Bayes methods made possible by the availability of computer codes that permit the use of Markov Chain Monte Carlo (MCMC) methods of full posterior inference. These are extensively discussed in References 43, 44, 46, and a number of examples of their use now exist (75, 78).

One method of analysis viewed by many in the 1990s as holding promise for identifying clusters of disease was the Cartogram or density equalizing projection. This approach (8, p. 2207; 52) appears now to have reached a dead end as more people appreciate the nonuniqueness of the transformed space, the difficulties of maintaining true spatial relationships in the new space (direction, contiguity,

distance), and the availability to researchers of more effective methods of accomplishing the purposes for which this methodology was developed.

ADJUSTING FOR SPATIAL AUTOCORRELATION

For some diseases, rates are most similar in contiguous areas and have a tendency to decline with distance from some core area. Such spatially correlated rates support the hypothesis that the process generating them involves a core area and risks for contracting the disease that decline with distance from the core. Becker et al. (5), using geocoded addresses, illustrate such a process for gonorrhea transmission in Baltimore, Maryland.

Many spatial analytic methods are concerned to remove "spatial autocorrelation effects." The well-known tendency for rates in nearby locations to be more alike than in more distant locations is an effect that violates some of the distributional assumptions of the models used. This fact has led some researchers to adjust findings for the spatial autocorrelation effect (24). Griffith (29) discusses this issue with respect to adjusting for spatial autocorrelation in environmental measures—in his case, lead concentrations in soil in three different types of environments.

A useful point of departure in analyzing any spatial distribution of a disease is to assess the degree of spatial autocorrelation of the rates. "Global Moran's I and local Moran's I_i are the most commonly used test statistics for spatial autocorrelation in univariate map patterns or in regression residuals" (72).

GEODEMOGRAPHICS

As interest increases in interventions for cancer prevention and control, geodemographics data and analysis are used, and GIS-based information displays are beginning to be used in cancer control and prevention activities (49, 55).

Testing Hypotheses/Spatial Clustering

A common question is whether the spatial pattern of disease observed is consistent with a given hypothesis about its spatial pattern. An increasingly common method of testing such hypotheses is by generating the spatial pattern consistent with a given hypothesis using Monte Carlo simulation methods. There is a noticeable increase in reliance on computational approaches—rather than analytic methods—for testing such hypotheses. This trend appears to be driven by increasing evidence that analytic approximations for estimating levels of significance are not as accurate as previously thought (10), and by the fact that quite complex computational models can now be implemented on desktop computing systems. A number of recent studies, for example, have generated *n* simulations and then computed, for a sample of locations on the map, the fraction of the simulations in which the simulated rate is less than the observed rate (35, 43, 67). It is becoming common to map this fraction as a *p*-value surface (44, 65). The use of simulations of data to verify critical

capabilities or weaknesses of alternative models has increased in recent work (14; 20; 39; 45, p. 2217; 70). This is especially the case for dealing with the problem of adjusting *p*-values for the multiple testing of hypotheses in the presence of multiple dependence among tests (13). Both Kulldorff (42) and Tango (70) use Monte Carlo simulations to compute reference distributions for determining the correct *P*-values resulting from multiple tests of spatial clustering of disease incidences.

Gelman et al. (27) used this approach to show that when using recently developed methods for identifying the spatial characteristics of disease patterns, spatial features (artifacts) of disease data are more likely to be recognized in the vicinity of edges or corners of study areas. The significance of observed measures of Moran's I are sometimes tested by simulation experiments in which the observed disease rate values are randomized and reference distributions of the values of Moran's I in the simulated patterns are computed, and the rank of the observed value in the ordered set of values from the simulation experiment is used as a measure of the statistical likelihood of finding the observed value if no spatial autocorrelation of values was present.

Discussions of methods for detecting disease clusters have continued, with many problems remaining to be solved. Burr (13) provides an excellent review of cluster methods and, using simulation, provides evidence of the difficulty of correctly estimating p-values by analytical methods. Burr focuses on the situation of interpreting reports of the highest cancer rate value so far found—the maximally selected value.

In addition to studies that search for ad hoc disease clusters with the expectation that clusters discovered will be discussed and assessed a posteriori, some recent studies have formulated spatial processes and then tested for their existence in real data. Tiefelsdorf (71), for example, postulated that the spatial pattern of bladder cancer in Germany was linked to the urban hierarchy there. He tested for clustering in the regression residuals of the same sign as indications that relationships changed across the map. There is increased interest in testing whether the relationships between covariates of disease rates are the same across the map. To date very few studies have tested for such differences—known as non-stationarities—in the processes that generate the patterns we observe.

Although we have noted the increased use of point data rather than areal data in public health applications of GIS, the use of explanatory variables at the individual level to compute an expected disease density surface has not yet proceeded far. This nevertheless is a promising approach (20, 39), albeit not yet widely adopted. The familiarity of ecological modeling approaches appears to be responsible for the very slow adoption of what surely should be regarded as a most promising approach.

Software Systems for Exploratory Spatial Data Analysis

For public health professionals, the goal is frequently to glean some insight from data. Rather than reporting the result of testing a formal statistical hypothesis, as in

a classical scientific approach, they successively apply graphical or other visualization tools, data enhancement methods, and a mix of descriptive statistics and more formal data models. For this purpose, a number of prototype analysis systems have been proposed and constructed (2, 4, 9, 31, 74). Carr et al. (15) developed linked micromap plots that show, in one connected view, relationships between attributes of areas and the spatial pattern of disease rates. Their illustration of spatial patterns of lung cancer rates among white males aged 65–74 in the South-East and poverty rates should be seen by public health specialists as a potential approach for studying and organizing preventive and control programs for many diseases. In Figure 3, the states are shown by their rank order on mortality rate for cervix cancer and their estimated percent pap test rate. Ninety-five percent confidence limits on both variables are also shown. The website characteristics show that the user can select from types of cancer displayed, types of geography and control of color schemes.

Relationships Between Disease and Environmental Factors

One area of increased interest in GIS and the spatial analysis of health data is that of establishing relationships between disease rates and exposures to environmental factors (23). Järup (35) provides a review of geographical studies in this area. GIS is being used to reconstruct from historical records residential exposures to pesticide use (11, 76, 77). A common principle used is that of reconstructing past land-use patterns and assuming that historical rates of pesticide application were applied on particular pieces of land. One study design created "a set of addresses that were distributed throughout the study area in roughly the same density as the age-stratified distribution of subjects in the case-control study." This work includes a review of the use of GIS in exposure assessment. Data on the intensity of pesticide applications at a fine geographic scale are now available to researchers from the California Department of Pesticide Regulation. Researchers from the California Department of Health Services (61) used a geographic information system to calculate the density of pesticide use in pounds per square mile of total land area for all Census Block Groups in California. In a second paper (30), they analyzed population-based rates of childhood cancer incidence throughout California in relation to agricultural pesticide use. They used a geographic information system to assign summary population, exposure, and outcome attributes at the block group level. They generally found no association between density of pesticide use and rates of childhood cancer incidence.

Spatial analysis methods are also becoming embedded in spatial decision support systems for public health. Njemanze et al. (56) used risk analysis methods and geographic information system technologies to evaluate the health effects of types of water sources available to rural populations in Nigeria. "Probabilistic layer analysis"—a computer-based modeling approach—spatially displayed the water source as layers of geology, hydrology, population, environmental pollution, and electricity. The spatial analysis methodology enabled estimation of the probability

of the risk of diarrheal diseases occurring at each layer and proposed solutions to eliminate these risks.

THE FUTURE

There are some forecasts of the future use of GIS in public health that appear to this author to be self-evident. It will, however, take effort and commitment to make them happen, and they are more likely to be seen first in state and local areas rather than at the national level in the United States. They may also be seen in other countries before they are seen in the United States. First will be the development of a more structured and organized system of disease surveillance in which geo-referenced data systems are examined by agent-based search systems designed to alert public health workers to areas and social and economic groups that have substantially higher disease rates. Cancer is one disease where, with the recent development of state cancer registries, the infrastructure for such spatial analysis might first emerge. Second, evidence-based research in clinical practice is likely to broaden into the geographic domain as evidence accumulates that differences in clinical practice frequently occur between geographic areas (57). For such developments to happen, for each of the areas served by a public health institution, an infrastructure of spatial data to which health data attributes can easily and accurately be attached will need to be developed. In addition, public health personnel will be needed who are trained to work in such an environment. There are few signs that such data and analysis systems are being planned at this time. Until such planning is implemented, the potential value of GIS and spatial analysis in public health will remain largely unrealized.

ACKNOWLEDGMENT

Figures 1 and 2 were made by Rudi Banerjee.

LITERATURE CITED

1. Ali, M, Emch M, Donnay JP, Yunus M, Sack RB. 2002. The spatial epidemiology of cholera in an endemic area of Bangladesh. *Soc. Sci. Med.* 55:1015–24
2. Anselin L. 2000. Computing environments for spatial data analysis. *J. Geogr. Syst.* 2:201–20
3. Aylin P, Maheswaran R, Wakefield J, Cockings S, Järup L, et al. 1999. A national facility for small area disease mapping and rapid initial assessment of apparent disease clusters around a point source: the UK Small Area Health Statistics Unit. *J. Public Health Med.* 21:289–98
4. Bao S, Anselin L, Martin D, Stralberg D. 2000. Seamless integration of spatial statistics and GIS: the S-PLUS for ArcView and the S+Grassland links. *J. Geogr. Syst.* 2:287–306
5. Becker KM, Glass GE, Brathwaite W,

Zenilman JM. 1998. Geographic epidemiology of gonorrhea in Baltimore, Maryland, using a geographic information system. *Am. J. Epidemiol.* 147:709–16

6. Best N, Wakefield J. 1999. Accounting for inaccuracies in population counts and case registration in cancer mapping studies. *J. R. Stat. Soc.* A 162:363–82
7. Bithell JF. 1990. An application of density estimation to geographical epidemiology. *Stat. Med.* 9:691–701
8. Bithell JF. 2000. A classification of disease mapping methods. *Stat. Med.* 19:2203–15
9. Bivand R, Gebhardt A. 2000. Implementing functions for spatial statistical analysis using the R language. *J. Geogr. Syst.* 2:307–17
10. Boots B, Tiefelsdorf M. 2000. Global and local spatial autocorrelation in bounded regular tessellations. *J. Geogr. Syst.* 2:319–48
11. Brody JG, Vorhees DJ, Melly SJ, Swedis SR, Drivas PJ, et al. 2002. Using GIS and historical records to reconstruct residential exposure to large-scale pesticide application. *J. Expos. Anal. Environ. Epidemiol.* 12:64–80
12. Brown PJB, McCulloch PG, Williams EMI, Ashurst DC. 2001. Control sample design using a geodemographic discriminator: an application of super profiles. *J. Geogr. Syst.* 3:107–35
13. Burr TL. 2001. Maximally selected measures of evidence of disease clusters. *Stat. Med.* 20:1443–60
14. Byers S, Besag J. 2000. Inference on a collapsed margin in disease mapping. *Stat. Med.* 19:2243–49
15. Carr DB, Wallin JF, Carr DA. 2000. Two new templates for epidemiology applications: linked micromap plots and conditioned choropleth maps. *Stat. Med.* 19:2521–38
16. Carstairs V. 2000. Socio-economic factors at areal level and their relationship with health. In *Spatial Epidemiology: Methods and Applications*, ed. P Elliott, JC Wakefield, NG Best, DJ Briggs, 4:51–67. Oxford: Oxford Univ. Press
17. Casper ML, Barnett E, Halverson JA, Elmes GA, Braham VE, et al. 2000. *Women and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality*. Morgantown, WV: West VA Univ., Off. Soc. Environ. Health Res.
18. Clayton DG, Kaldor J. 1987. Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. *Biometrics* 43:671–81
19. Devine O, Louis T. 1994. A constrained empirical Bayes estimator for incidence rates in areas with small populations. *Stat. Med.* 13:1119–33
20. Diggle P, Morris S, Morton-Jones T. 1999. Case-control isotonic regression for investigation of elevation in risk around a point source. *Stat. Med.* 18:1605–13
21. Dobson JE, Bright EA, Coleman PR, Durfee RC, Worley BA. 1999. A global population database for estimating population at risk. *Photogramm. Eng. Remote Sens.* 16:849–57
22. Elliott P, Wakefield JC, Best NG, Briggs DJ, eds. 2000. *Spatial Epidemiology: Methods and Applications*. Oxford: Oxford Univ. Press
23. English P, Neutra R, Scalf R, Sullivan M, Waller L, Zhu L. 1999. Examining associations between childhood asthma and traffic flow using a geographic information system. *Environ. Health Perspect.* 107(9):761–67
24. Fortin M-J, Jacquez GM. 2000. Randomization tests and spatially auto-correlated data. *Bull. Ecol. Soc. Am.* 81:201–5
25. Gatrell AC, Loytonen M. 1998. *GIS and Health.* London: Taylor & Francis
26. Gelman G, Price PN. 1999. All maps of parameter estimates are misleading. *Stat. Med.* 18:3221–34
27. Gelman G, Price PN, Lin C-Y. 2000. A method for quantifying artifacts in mapping methods illustrated by application to headbanging. *Stat. Med.* 19:2309–20
28. Goldman DA, Brender JD. 2000. Are

standardized mortality ratios valid for public health data analysis? *Stat. Med.* 19:1081–88
29. Griffith DA. 2002. The geographic distribution of soil lead concentration: description and concerns. *URISA J.* 14:5–15
30. Gunier RB, Harnly ME, Reynolds P, Hertz A, VonBehren J. 2001. Agricultural pesticide use in California: pesticide prioritization, use densities, and population distributions for a childhood cancer study. *Environ. Health Perspect.* 109:1071–78
31. Haining R, Wise S, Ma J. 2000. Designing and implementing software for spatial statistical analysis in a GIS environment. *J. Geogr. Syst.* 2:257–86
32. Hansen KM. 1991. Head-banging: robust smoothing in the plane. *IEEE Trans. Geosci. Remote Sens.* 29:369–78
33. Jacquez GM. 2000. Spatial analysis in epidemiology: nascent science or a failure of GIS. *J. Geogr. Syst.* 2:91–97
34. Jacquez GM, Maruca S, Fortin M-J. 2000. From fields to objects: a review of geographic boundary analysis. *J. Geogr. Syst.* 2:221–41
35. Järup L, Best N, Toledano MB, Wakefield J, Elliott P. 2002. Geographical epidemiology of prostate cancer in Great Britain. *Int. J. Cancer* 97:695–99
36. Jemal A, Kulldorff M, DeVesa SS, Hayes RB, Fraumeni JF Jr. 2002. A geographic analysis of prostate cancer mortality in the United States, 1970–89. *Int. J. Cancer* 101:168–74
37. Kafadar K. 1997. Geographic trends in prostate cancer mortality: an application of spatial smoothers and the need for adjustment. *Ann. Epidemiol.* 7:35–45
38. Kafadar K. 1999. Simultaneous smoothing and adjusting mortality rates in U.S. counties: melanoma in white females and white males. *Stat. Med.* 18:3167–88
39. Kelsall JE, Diggle PJ. 1998. Spatial variation in risk of disease: a nonparametric binary regression approach. *Appl. Stat.* 47:559–73
40. Kleinman JC. 1977. Age-adjusted mortality indexes for small areas: applications to health planning. *Am. J. Public Health* 67:834–40
41. Krieger N, Williams DR, Moss NE. 1997. Measuring social class in U.S. public health research. *Annu. Rev. Public Health* 18:342–78
42. Kulldorff M. 1997. A spatial scan statistic. *Commun. Stat. Theor. Methods* 26:1481–96
43. Lawson AB. 2001. Tutorial in biostatistics: disease map reconstruction. *Stat. Med.* 20:2183–204
44. Lawson AB. 2001. *Statistical Methods in Spatial Epidemiology*. New York: Wiley
45. Lawson AB, Biggeri AB, Boehning D, Lesaffre E, Viel J-F, et al. 2000. Disease mapping models: an empirical evaluation. *Stat. Med.* 19:2217–41
46. Lawson AB, Bohning D, Lessafre E, Biggeri A, Viel J-F, Bertollini R, eds. 1999. *Disease Mapping and Risk Assessment for Public Health.* New York: Wiley
47. Lawson AB, Clark A. 2002. Spatial mixture relative risk models applied to disease mapping. *Stat. Med.* 21:359–70
48. Leyland AH, Langford IH, Rasbash J, Goldstein H. 2000. Multivariate spatial models for event data. *Stat. Med.* 19:2469–78
49. Lubenow A, Tolson K. 2001. GIS technology helps pinpoint patients Health Management Technology. http://www.medstat.com/pdf/medstat.pdf
50. Marshall R. 1991. Mapping disease and mortality rates using empirical Bayes estimators. *Appl. Stat.* 40:283–94
51. McNab YC, Dean CB. 2002. Spatiotemporal modeling of rates for the construction of disease maps *Stat. Med.* 21:347–58
52. Merrill DW, Selvin S, Close ER, Holmes HH. 1996. Use of density equalizing map projections (DEMP) in the analysis of childhood cancer in four California counties. *Stat. Med.* 15:1837–48
53. Moore K. 2000. RESEL filtering to aid visualization within an exploratory data

analysis system. *J. Geogr. Syst.* 2:375–98
54. Mungiole M, Pickle LW, Simonson KH. 1999. Application of a weighted head-banging algorithm to mortality data maps. *Stat. Med.* 18:3201–9
55. Natl. Cancer Inst. 2002. Consumer health profiles. Cancer Control Popul. Sci., Appl. Cancer Screen. Res. Branch. http://cancercontrol.cancer.gov/ACSRB/consumer.html
56. Njemanze PC, Anozie J, Ihenacho JO, Russell MJ, Uwaeziozi AB. 1999. Application of risk analysis and geographic information system technologies to the prevention of diarrheal diseases in Nigeria. *Am. J. Trop. Med. Hyg.* 61:356–60
57. Parchman ML, Ferrer RL, Blanchard KS. 2002. Geography and geographic information systems in family medicine research. *Fam. Med.* 34(2):132–37
58. Pickle LW, Su Y. 2002. Within-state geographic patterns of health insurance coverage and health risk factors in the United States. *Am. J. Prev. Med.* 22:75–83
59. Pickle LW, Mungiole M, Jones GK, White, AA. 1999. Exploring spatial patterns of mortality: the new Atlas of United States Mortality. *Stat. Med.* 18:3211–20
60. Pickle L, White AA. 1995. Effects of the choice of age-adjustment method on maps of death rates. *Stat. Med.* 14:615–27
61. Reynolds P, Von Behren J, Gunier RB, Goldberg DE, Hertz A, Harnly ME. 2002. Childhood cancer and agricultural pesticide use: an ecologic study in California. *Environ. Health Perspect.* 110:319–24
62. Richards TB, Croner CM, Rushton G, Brown CK, Fowler L. 1999. Geographic information systems and public health: mapping the future. *Public Health Rep.* 114:359–73
63. Rigby JE, Gatrell AC. 2000. Spatial patterns in breast cancer incidence in northwest Lancashire. *Area* 32:71–78
64. Rothman KJ. 1986. *Modern Epidemiology.* Chapter 5. Boston/Toronto: Little, Brown
65. Rushton G, Lolonis P. 1996. Exploratory spatial analysis of birth defect rates in an urban population. *Stat. Med.* 15:717–26
66. Rushton G, West M. 1999. Women with localized breast cancer selecting mastectomy treatment, Iowa, 1991–1996. *Public Health Rep.* 114(4):370–71
67. Semenciw RM, Le ND, Marrett LD, Robson DL, Turner D, Walter SD. 2000. Methodological issues in the development of the Canadian cancer incidence atlas. *Stat. Med.* 19:2437–49
68. Siegel C, Davidson A, Kafadar K, Norris JM, Todd J, Steiner J. 1997. Geographic analysis of pertussis infection in an urban area: a tool for health services planning. *Am. J. Public Health* 87:2022–26
69. Talbot TO, Kulldorff M, Forand SP, Haley VB. 2000. Evaluation of spatial filters to create smoothed maps of health data. *Stat. Med.* 19:2399–408
70. Tango T. 2000. A test for spatial disease clustering adjusted for multiple testing. *Stat. Med.* 19:191–204
71. Tiefelsdorf M. 1998. Some practical applications of Moran's I's exact conditional distribution. *Pap. Reg. Sci.* 77(2):101–29
72. Tiefelsdorf M. 2002. The saddlepoint approximation of Moran's I's and local Moran's I_i's reference distributions and their numerical evaluation. *Geogr. Anal.* 34(3):187–206
73. Wakefield J, Elliott P. 1999. Issues in the statistical analysis of small area health data. *Stat. Med.* 18:2377–99
74. Wall PA, Devine OJ. 2000. Interactive analysis of the spatial distribution of disease using a geographic information system. *J. Geogr. Syst.* 2:243–56
75. Waller L, Carlin B, Xia H, Gelfand A. 1997. Hierarchical spatio-temporal mapping of disease rates. *J. Am. Stat. Assoc.* 92:607–17
75a. Wang X, Chen JX, Carr DB, Bell BS, Pickle LW. Geographic statistics visualization: web-based linked micromap plots. *Comput. Sci. Eng.* 4(3):90–94
76. Ward MH, Nuckols JR, Weigel SJ,

Maxwell SK, Cantor KP, et al. 2000. Identifying populations potentially exposed to agricultural pesticides using remote sensing and a geographic information system. *Environ. Health Perspect.* 108:5–12

77. Xiang H, Nuckols JR, Stallones L. 2000. A geographic information assessment of birth weight and crop production patterns around mother's residence. *Environ. Res.* (Sect. A) 82:160–67
78. Zia H, Carlin BP, Waller LA. 1997. Hierarchical models for mapping Ohio lung cancer rates. *Environmentrics* 8:107–20

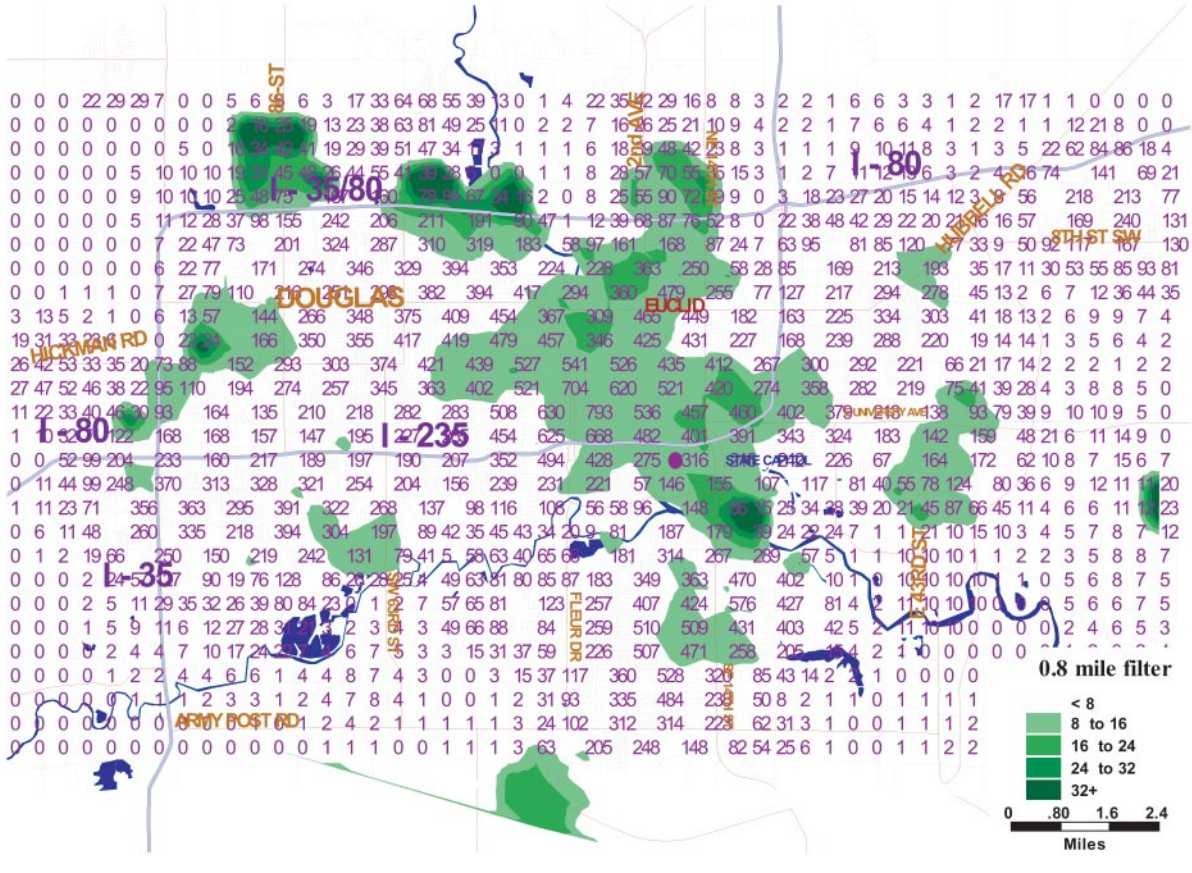

Figure 1 Des Moines, Iowa: births and infant mortality rates (all races), 1996–1998.

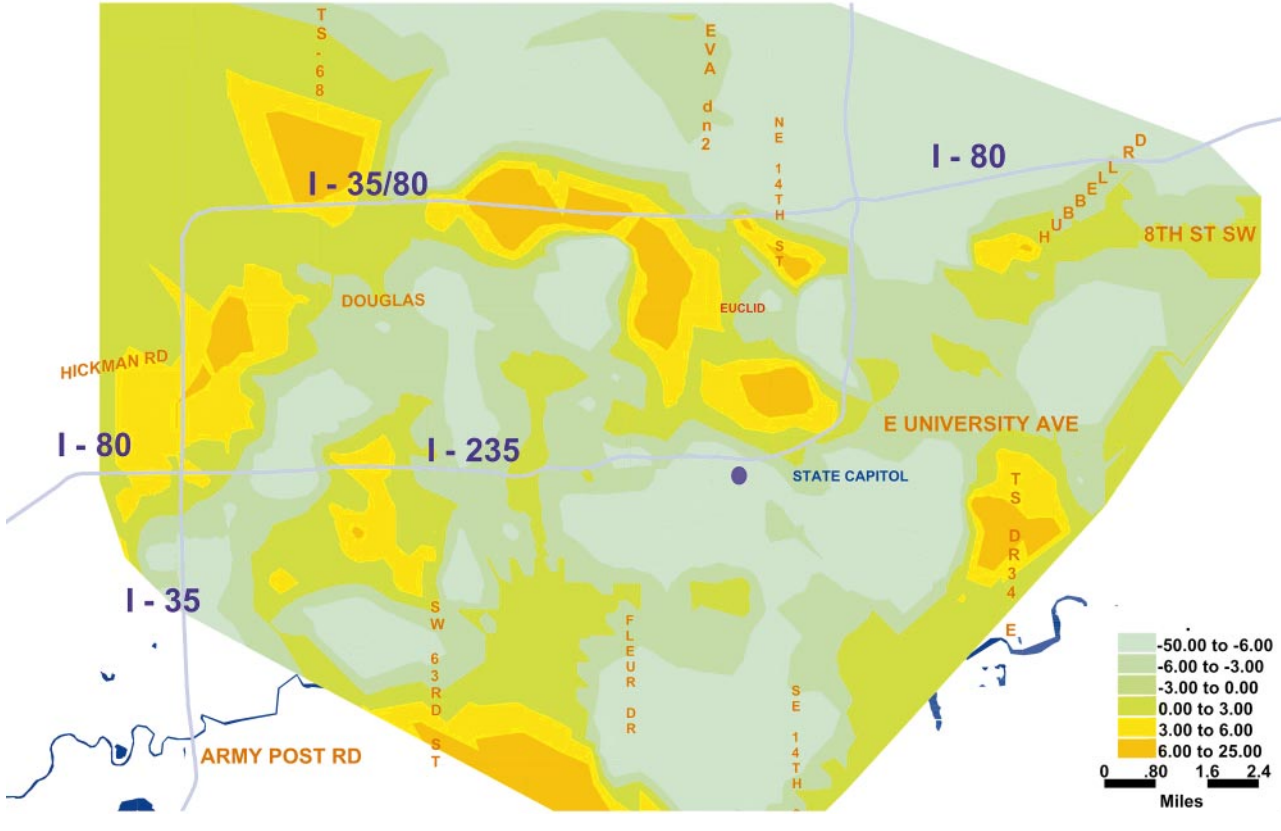

Figure 2 Des Moines, Iowa: change in infant mortality, 1993–1998.

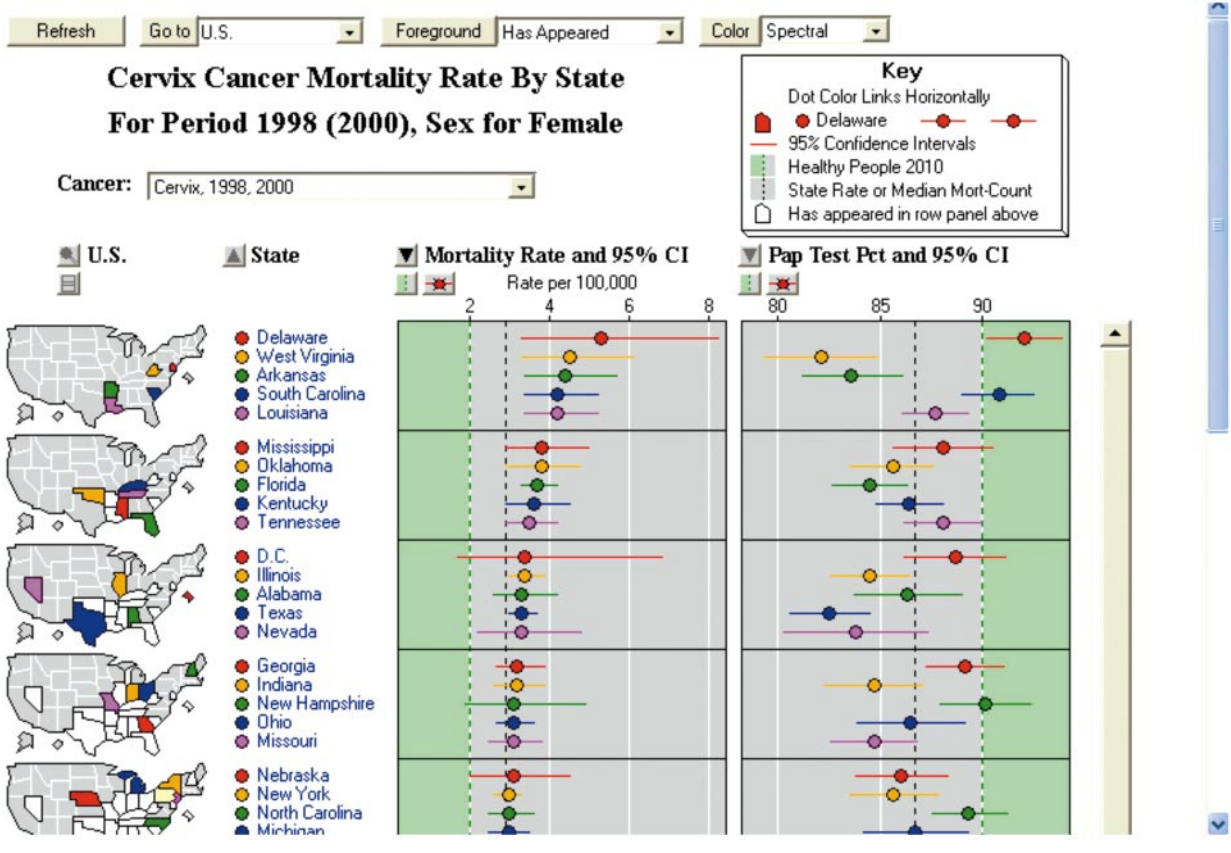

Figure 3 A snapshot of linked micromap plots of U.S. cancer statistics. Source: X Wang, JX Chen, DB Carr, BS Bell, LW Pickle, "Geographic statistics visualization: web-based linked micromap plots," *Comput. Sci. Eng.* 4(3):90–94 May/June 2002, p. 91 (75a).

Annual Review of Public Health
Volume 24, 2003

CONTENTS