Considerations for Improving Geographic Information System Research in Public Health

Gerard Rushton, Gregory Elmes, Robert McMaster

Abstract: *Many activities to promote better health and to reduce disease are directed at the changing environments in which people live. In response to the University Consortium for Geographic Information Science (UCGIS) 1999 Summer Assembly application challenge, this paper describes the uses of geographic information systems in the field of public health care. With the geographic information system, observations regarding the social, economic, political, and physical environments can be referenced to a common geospatial data framework. This permits varying organizations to share spatial data regarding these phenomena. Geographic information science has the potential to create rich information databases, linked to methods of spatial analysis, to determine relationships between geographical patterns of disease distribution and social and physical environmental conditions. As the core of a decision-support system, geographic information science also has the potential to change the way that allocations of resources are made to facilitate preventive health services and to control the burden of disease.*

Introduction

The social, economic, political and physical environments in which people live are changing, as activities to encourage better health and reduce disease are being promoted. Using the geographic information system (GIS), observations regarding any of these environments can be referenced to a common geospatial data framework. Organizations can then share spatial data regarding these phenomena. Geographic information science (GIScience) has the potential to bring rich information databases, linked to methods of spatial analysis, to determine relationships between geographical patterns of disease distributions and social and physical environmental conditions. As the core of a decision support system, GIScience also has the potential to change the way in which geographical resource allocations are made to facilitate the establishment of preventive health services and to control the burden of disease in patients.

Traditionally, administrative areas or other spatial units such as census-defined areas have been the geographic units where health status and health outcomes are measured and where health resources allocated. However, if the regions are, for example, small counties, the movement of people from one county to another has made this spatial accounting scheme inaccurate for measuring the relationships between health status, health resources, and health outcomes, and, therefore, inappropriate to be used as decision-making units. For example, the federal government has struggled for decades to devise appropriate spatial accounting methods by which local communities deemed to be disadvantaged with respect to health resources could be assisted without spending scarce resources on communities not economically disadvantaged (see Lee 1991, General Accounting Office 1997, Bureau of Primary Health Care 1998). If the areas used to report disease data and resource allocations are larger than counties, however, differences within the areas may be great and they are inappropriate for analysis and decision-making. Whether in the U.S. or other countries, counties differ greatly in size and population.

It is clear that diverse health resources affect the levels of health and disease over local areas of different size. Health systems operate at a multitude of spatial scales that are constantly changing with the reorganization of health resources and the behavior of health-seeking populations. As many public and private health organizations have discovered, to understand and make decisions about this complex system of inputs and outcomes requires operational access to information at very localized geographic scales. In this respect, health systems are not different from many other systems for which people receive services from dispersed facilities. As with such systems, the early model of a spatial accounting framework of counties or fixed service areas is now being replaced by a GIS-based model in which the geographic scale that information is analyzed changes according to the kind of question being addressed. Using GIS, data from small, localized areas can be flexibly aggregated to larger areas that are meaningful for the questions asked and the decisions to be made.

Background

Federal, state, and local institutions have recently developed a strong interest in GIS as it relates to health. The Centers for Disease Control (CDC) have supported the development of software for mapping diseases (Dean 1999), and the National Cancer Institute (NCI) has supported the development of software for disease cluster identification (Kulldorff et al. 1998). The CDC and the Agency for Toxic Substances and Disease Registry

(ATSDR), a branch of the Environmental Protection Agency (EPA), have cooperated in sponsoring and organizing four annual meetings on GIS and public health, the most recent held in San Diego in August 1998. The CDC also collaborate with the National Center for Health Statistics and the U.S. Department of Health and Human Services on an annual conference on health statistics that increasingly has geographic information and Internet technology components. In 1998, the NCI in collaboration with the National Institute of Environmental Health Sciences (NIEHS), under Public Law 103-43 requested proposals to develop a health-related GIS for Long Island. "The prototype health-related GIS will provide researchers a new tool to investigate relationships between breast cancer and the environment on Long Island, and to estimate exposures to environmental contamination." (NCI 1999a: 3).

Despite this, considerable skepticism exists in many quarters about the role of mapping and spatial analysis in the analysis of disease patterns and resource allocation. For example, Dennis Whalen, Executive Deputy Commissioner, New York State Department of Health, in testimony before a Committee of the New York State Assembly on March 8, 1999 noted in a discussion of "cancer mapping challenges" that:

> Some experts say mapping itself, is ineffective—that maps will provide little additional information about cancer patterns, so resources should be directed to more promising research. They argue that maps presuppose a geographic link to cancer cases that cannot be proven, and in fact, may be completely irrelevant. For example, one would have to question the validity of overlaying a map of current environmental exposure data on a 1991 through 1996 cancer incidence map knowing that a particular type of cancer may have a latency period of 10 to 20 years, and that many of those diagnosed with a common cancer may not have lived in the area long enough for their cancers to have a common cause. Yet many mapping supporters expect that maps will definitively identify "hot spots". They expect that maps will demonstrate a cause and effect relationship between cancer cases and a particular risk factor or factors.

Whalen went on to argue that mapping might be a useful tool on which to focus further research and for generating hypotheses. Furthermore, maps can help target efforts such as increased physician education on available treatments. He concluded, "We believe that cancer mapping is the next logical step to address the call by New Yorkers for more information about cancer cases in their communities." The New York State Health Department has formed an advisory committee on cancer surveillance and much of its research plan relates to GIS use.

There have been many reviews on the use of GIS in public health or in the provision or planning of health services (Ricketts et al. 1994, Briggs and Elliott 1995, Clarke et al. 1996, Croner et al. 1996, Waller 1996, Rushton et al. 1997, Vine et al. 1997, Croner and Stroup 1999, Richards et al. 1999, Rushton 1999, Yasnoff and Sondik 1999). These reviews, however, have focused on the potential use of GIS as currently conceived.

In this work, we are not reviewing attempts to use GIS in epidemiology and health services research. Instead, we focus on the educational and research needs to fulfill the potential of improving health with GIS. It is our contention that, although a great deal can be accomplished to improve health with GIS, significant challenges exist that only further research in GIS and health can solve. This paper is one of several organized by the University Consortium for Geographic Information Science (UCGIS) that address applications of GIScience. It follows, and is patterned after, the in-depth reviews of fundamental GIScience research and education issues previously conducted by UCGIS.

The UCGIS Approach: What Are the Key Issues?

Research in the area of GIS and public health can be associated with the 10 "research challenges" defined by the UCGIS 1999 Summer Assembly, which now constitutes its "research agenda." Based on the UCGIS assemblies of 1996 and 1998, the challenges that emerged include:

- Spatial data acquisition
- Scale
- Cognition of geographic information
- Extensions to geographic information
- Uncertainty
- Interoperability
- Distributed and mobile computing
- The spatial information infrastructure
- Spatial analysis and modeling
- GIS and Society

Although we will not attempt to link each of the challenges to GIS and public health, each appears within the structure that we use below to explore this subject.

Educational

Much of the education on GIS and public health will fall under the heading of "Professional GIS education." Unfortunately, professionals seeking education/training in GIS rarely find specific courses or workshops on the application of GIS in public health. Both degree and certificate-granting graduate programs in GIS, such as those found at San Diego State University (certificate), Clark University (M.A.), and the University of Minnesota (M.GIS.), allow students to become educated and trained in the areas of GIS/GIScience. Specific coursework in GIS and public health, however, is not usually available. Some UCGIS member universities have schools of public health such as those found at the Universities of Iowa, Minnesota, Washington, and North Carolina, and at Ohio State University. Perhaps these institutions will develop programs to address this need. Although most of the teachers, researchers, and leaders of public health organizations hold advanced degrees (Masters of Public Health (M.P.H.) for professional work and Ph.D. for advanced instruction and research), the curricula for these degrees and the research experiences do not normally include GIS-related subjects.

Perhaps the bridge between these two areas—GIS and public health—will be made by the biostatisticians, as they begin to use GISs to test their methods. Increasingly, they are investigating relationships between disease and environmental factors where, for instance, in the case of major polluted areas (Superfund sites), they often inherit spatially referenced data systems from legacy information systems. In such situations, there is a tendency to confuse GIS with GIScience. Wanting to develop support systems for researchers working in this area, they often hire staff for their institutes without stipulating in their personnel search that the person to be hired know the concepts and terminology of GIS. They believe that experience and knowledge of GISs, meaning the particular software with which they intend to work, are what is needed. Only later do they discover that their new personnel do not know the basic data models and the conversion methods between data models. These may be raster or vector, Triangular Irregular Network (TIN), or network models. Their staff should know how to use recognized georeferencing and coordinate systems, including relative georeferencing and map projections. They should know the language and concepts of geometric and attribute accuracy. They should know about buffer zones for points, lines, and areas as well as relational and hierarchical database systems and object-oriented systems. Also important is knowledge about positional data accuracy, hash functions, quad-trees, and spatial logic operations. It is important to be familiar with the principles of aerial photo interpretation as well as supervised and unsupervised classifications. They should know some principles of surveying and remote sensing, street centerline systems and address matching, digitizing, and scanning of spatial data. These topics are not covered within the typical curricula of departments of statistics, biostatistics, public health, or computer science. From where will practicing public health workers or academic scientists who study the relationships between environments, disease, and health find people with this knowledge? (Bernhardsen 1999)

There are serious educational needs for both researchers and practitioners in the health fields who are using GIS in their work and who are struggling to find the educational resources to meet their needs. The CDC and ATSDR are currently developing distance learning modules on GIS and public health and they expect to broadcast these soon using satellite-based, video, broadcasting systems. Rushton and colleagues, with a grant from the Department of Education, organized five three-day workshops for health professionals between 1993 and 1997. They also developed a Web presence and CD-ROM on the subject of GIS and public health (Rushton et al. 1997). Getis and Goodchild presented a workshop at the GIS and public health conference in San Diego in 1998. A widely held view is that more needs to be done to educate health professionals on the use of GIS in public health activities.

In the health sciences, one common approach to educating advanced professionals in areas outside their area of traditional education is through focused, post-doctoral training programs. The National Institutes of Health (NIH) frequently support such programs through its ongoing aid of focused research institutes. The Basic Science Research Program for Superfund sites of the NIEHS, for example, offers aid for research and education units that include GIS among the core support areas of several of the research institutes they aid. Such a program in GIS might contribute to the twin goal of preparing new teachers and researchers in GIS and public health and in advancing critical research areas.

Research

As an application area that only recently recognized the potential contributions of GIS-based research, the health disciplines have not yet formulated a plan for research on GIS use in public health. The NCI prepared, for its Long Island Breast Cancer Project, a glossary of GIS-based methods that have been used in investigations of the spatial distributions of disease and possible relationships with environmental factors. It is reproduced as Appendix A. For each application area, at least one citation to published research was provided. The list is an expansive view of the relationship of GIS to more traditional analysis methods. It recognizes that GIS is far more than a mapping tool. It is seen here as an analytic system that brings together in a geographic context information on disease incidence, health services availability and access, demographic characteristics, and environmental factors. Each of these information systems has its own geographic referencing system and the functional relationships between them operate at different geographic scales.

Many current applications of GIS in health are extremely wasteful of resources in that their ad hoc nature requires costly GIS resources to be developed to support single project plans. The recently developed and NCI-supported Long Island Breast Cancer Project attempts to address this problem by supporting, under contract, the development of a GIS utility for this region. The system, currently being developed by AverStar Inc., will develop selected spatial coverages and will implement selected spatial analysis methods prioritized by the NCI from the taxonomy of methods in Appendix A.

There are other areas of research (e.g., the location of health facilities), for which focused reviews of work completed do not yet exist or problems are not yet addressed (Mohan 1983, Cromley and Shannon 1986, Hirschfield et al. 1993, 1995). Recent developments in the organization of health care through the development of managed care systems have strong geographical information and analysis components. Little research exists on this subject (Perkins 1999).

In the case of infectious diseases, the environmental conditions conducive to an outbreak can be identified and steps taken to control or eliminate disease transmission. GIS containing analytical modules can quickly identify hot spots and aid in the study of viral transmission vectors. Surveillance data collected in Puerto

Rico (Morrison et al. 1998) and currently in Peru enable researchers to identify the pattern of dengue cases with respect to its disease vector, the mosquito species Aedes aegypti. The unavailability of reliable georeferencing systems stands in the way of this type of work in difficult tropical environments.

In the case of diseases such as most cancers, exposures to agents that might increase the risk of disease often predate by 10 to 20 years the diagnosis of the disease. In such circumstances, the location of diagnosis and the location of probable exposure are unlikely to be the same. With a population that moves its residence so frequently, it is a formidable challenge to estimate the places of likely exposure by people whose location at the time of first diagnosis is known. Mark and Egenhofer (1998) and Mark et al. (1999) have recently begun geographic demographic research in the U.S. on methods to estimate the likelihood that a person whose current residence at the time of first diagnosis is at "x" might have lived in exposure area "y," "t" years ago (see the section "Temporal Aspects of GIS and Health"). Research on possible prior exposure to risks is proceeding in Sweden for environmentally linked leukemia and child-onset diabetes (Kohli et al. 1997).

Several authors have argued the merits of exploratory spatial data analysis for health applications. Haining et al. (1998) illustrated a system of analysis, SAGE, that can undertake exploratory spatial analysis held in the ArcInfo GIS. They illustrate the system with analyses of standardized incidence rates for cancer in Sheffield, England. Their system permits "brushing" of the region (identifying regions) and displaying relationships between variables for the data of the region brushed. Tools for regionalization are also developed in the SAGE system (Wise et al. 1997), as are computations of local statistics (Unwin and Unwin 1998) such as the widely used Getis-Ord (G_i^*-) statistic (Getis and Ord 1992, Ord and Getis 1995). Anselin and Bao (1997) also developed an interactive computational system that links many methods of spatial analysis to ArcView GIS.

Demographic data for small areas are crucial for many research applications of GIS in health, particularly for estimating the values of denominators in computing disease rates in these areas (Elliott et al. 1992, Martin 1996). Quality demographic data for small geographic areas, particularly publicly available data, is often not available, especially during inter-census periods.

Priority Areas and Recommendations for Research

Improving Disease Surveillance Data Systems

There is general agreement that location variables have not been collected well in most current disease surveillance systems. Until recently, for example, the New York State Department of Health's Cancer Registry recorded the current address of people with cancer. When their residence changed, the new location replaced the old. In the Iowa Cancer Registry, the locations of specific treatments are not coded even though they are available in the written record. There is a need for disease surveillance systems to adopt uniform methods for locational coding and to introduce quality assurance and quality testing standards for locations comparable to the standards they use for other data items they code. MacDorman and Gay (1999) recently reviewed state initiatives in geocoding vital statistics data.

The NIH and CDC are not unaware of this problem. In a recent report of the Surveillance Implementation Group of the NCI (1999b), one of 11 research opportunities identified is:

> Research Opportunity 4. Explore the feasibility and utility of employing geographic information systems for geocoding surveillance data and reporting geographic relationships among screening measures, risk factors (including environmental exposures), and improved cancer outcomes. Methods need to be developed for assuring data confidentiality. (The cost of this effort is expected to be moderate; work should be initiated within the next 1-2 years.) Research is needed on the utility of geographic information systems (GIS) as an innovative addition to the cancer surveillance infrastructure.

There is a need to develop methods of spatial analysis that can be routinely used for exploratory analysis of surveillance data (see Getis and Ord 1992, Rushton 1998, Rosenberg et al. 1999). There is also a need for a national dialog on the improvement and standardization of the quality and quantity of spatial information associated with health statistics. This should include an examination of existing national record and database systems (e.g., Health Care Financing Administration, Medicare/Medicaid Parts 1 and 2, and death certificates).

In concert with the need to improve health surveillance systems, calls have risen for better assessment of rural health and the health of minorities (Ricketts 1994, Bureau of Primary Health Care 1998). The Office of Social Environment and Health Research at West Virginia University in cooperation with CDC has published atlases of social environment affecting heart disease in Appalachia, and for women at the national and state scale (Barnett et. al. 1998, Casper et al. 1999). Gender and minority issues have not only been relatively neglected in the epidemiological literature but also raise elevated concern about confidentially.

Geographical data and the results of analysis are, more often than not, presented in the form of a map. Maps are complicated representations that cartographers, even with decades of cognitive research, do not fully understand in terms of reading and interpretation. Whereas the cartographic symbolization for static maps has been standardized for nearly 100 years, today we are quickly developing and implementing new symbolization schemes, many that deal with dynamic—four-dimensional—techniques. One can point to the work of MacEachren and colleagues at Pennsylvania State University (MacEachren and DiBiase 1991) in both the development of the chorodot technique for looking at the autoimmune deficiency syndrome (AIDS) distribution and an animation of AIDS data for the State of Pennsylvania. However, significant research is needed to understand exactly how such representations are read and, more importantly, understood by the users of these new cartographies. In particular, there is an "ethics" or set of concerns of which persons producing maps of public health data need to be aware. The significant question deals with the potential impact of such representational techniques on the public, and on individuals who may misread and misinterpret such maps. Cognitive research on map reading and interpretation is thus needed.

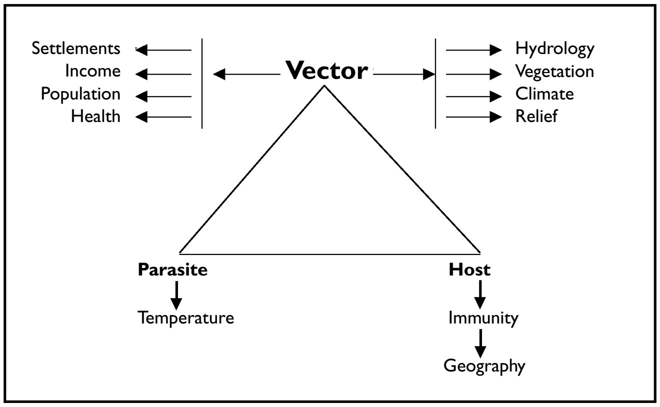
All geographical data have uncertainty, but in the area of public health assessment, this is of particular concern. Critical issues involve the positional and the attributed accuracy of the health information. The slight displacement of an enumeration boundary that is being used to show an affected population is of lesser concern, whereas the uncertainty associated with address-specific data (e.g., cases of tuberculosis in El Paso mapped as addresses) might be more problematic. New visualization techniques are needed to portray to the user the relative "reliability" of the data and how this reliability might vary from point to point. Simple techniques, such as the presentation of a "blanket of error" for the classification on choropleth maps—for better or worse the most commonly used mapping method for health data—should be considered. Another related topic is that of the data lineage. Where did the data come from? What year were they generated? What statistical manipulations have been applied? Such metadata is crucial for the user to ascertain fitness for use.

Risk Factors as Contributors to Disease and Ill Health

Behavioral risk factors are often discovered through national health surveys—for example, the Behavioral Risk Factor Surveillance Survey and the National Health Interview Survey. There is a need to link the findings from such surveys to local socio-demographics to estimate local risk factors based on expected local behavior patterns (see Brown et al. 1991, Braden and Beauregard 1994). Sometimes, attributes that can be observed in local administrative databases can be used as surrogates to estimate disease incidence rates. For example, the density of retail alcohol sites has been linked to local rates of alcohol abuse (Mackinnon et al. 1995). Haining et al. (1994) investigated the relationship between material deprivation and the rates of colorectal cancer.

It does seem clear that, with a few exceptions, theories of spatial diffusion and related spatial models are rarely given serious consideration in CDC and NIH research activities. Gould (1993) described the lack of interest he encountered in NIH groups that discussed the spread of AIDS in the U.S. in the 1980s. Some examples of spatial diffusion and core disease areas as explanations for current patterns of disease can be found in the work of Becker et al. (1998) and Cook et al. (1999).

Figure 1. Relationship between socioeconomic conditions and physical environmental conditions and the spread of infectious diseases. (This model was developed by Jamil Kazmi and E. Lynn Usery in April 1999, at the Department of Geography, University of Georgia, Athens, Georgia.)



Ecological Studies of the Relationship Between Environmental Factors and Disease Transmission

As a World Health Organization (WHO, 1996) Report recently noted, the spread of many infectious diseases is related to the climate, vegetation, and socioeconomic conditions in local areas.

Kasmi and Usery note that, although their model was developed for malaria, it is applicable to other vector-borne diseases as well. They write:

> The basic idea is that malaria is a three-factor disease which develops with the interaction of the vector (mosquito), parasite (plasmodium) and host (man). Absence of any of these three basic factors means the absence of malaria from various parts of the world. Each factor at the individual level has many contributing elements, for example, the vector has many physical and socio-economic elements which may contribute to the transmission and control of malaria. Therefore, the role of remote sensing and geographic information systems (GIS) as modern tools to study vector-borne diseases is to identify and interpret these contributing elements.

Recent research by Spear et al. (1998) illustrates the contribution of GIScience to this area. An illustration of an ongoing project in South Africa that uses GIS in the control of malaria activities can be seen at http://www.malaria.org.za/homested.htm. This project illustrates the sensitivity of malaria control activities to the geographic scale of surveillance activities.

With the possibility of significant climate change in many areas of the world, research is needed to project the likely effects of such changes on human health. The frequency and magnitude of extreme events increase health risks (WHO 1996, Smoyer 1998).

Those epidemiological studies that require environmental data, such as those assessing the dispersion of Lyme disease, must acquire up-to-date, and often classified (i.e., put into categories and coded) land use and land-cover data. Issues of spectral and spatial resolution, image classification, and ground truthing must be addressed for these applications.

Temporal Aspects of GIS and Health

The movements of persons through geographic space are a critical factor in exposures to environmental health hazards. Computational models that can account for the fact that a person's location in geographic space is dynamic rather than static will greatly enhance the power and potential of data analysis and reasoning methods for examining environmental exposures or discovering past clusters of currently-ill patients. Individuals navigate through space, they stay at locations where they meet other individuals, and they perform regularly recurring tasks that involve variable or fixed locations in geographic space. These movements often expose the individual to environmental factors that can result in health problems at latency periods, ranging from seconds to decades. For example, establishing whether a particular U.S. soldier was exposed to hazardous chemicals during Operation Desert Storm requires not only the space-time distribution of environmental risk, but also a record of the space-time behavior of the soldier. If the former is not known, space-time places of high risk might be inferred by comparing the space-time behaviors of soldiers showing symptoms of ill health with the behaviors of a control group of soldiers not showing symptoms.

For many health conditions, the application of GIS has been hampered by the poor ability of commercial GISs to handle multitemporal geographic information or movement (Langran 1992, Peuquet 1994). This shortcoming severely impedes the utility for GIS to assist in understanding health problems with long latency periods, such as many forms of cancer, since with mobile populations, the location of the patient at the time of diagnosis or mortality may have little relation to the location of the exposure to toxic substances or other environmental risks.

Recently, the NIEHS has supported a research project that focused on the extraction of health-related information from *geospatial lifelines*, which capture an individual's location in geographic space at regular or irregular temporal intervals (Mark and Egenhofer 1998, Mark et al. 1999). The objectives of this project are to develop and test the theory of geospatial lifelines in the environmental health sciences by:

- developing methods to trace locations of individual people (patients, cases, or controls) back through time, to discover spatial clusters in the past or to determine past environmental exposures;
- designing, prototyping, and assessing computational models that can deal with large sets of geospatial lifelines and environmental information; and
- examining the ethical and legal implications of recording an individual's geospatial lifelines in databases and establishing procedures for appropriate restrictions on data analysis and dissemination.

Geospatial lifelines (Mark and Egenhofer 1998) consist of a series of discrete space-time samples over the domain of continuous movements, describing an individual's location in geographic space at regular or irregular temporal intervals. Methods for the analysis of and reasoning about movement in geographic space are based in theory outlined three decades ago by Torsten Hägerstrand (1970). Although Hägerstrand's work has influenced conceptualizations of spatiotemporal constraints on human activities, it has rarely been implemented computationally (however, see Miller 1991). Geospatial lifeline data may be recorded at different resolutions; however, as they apply to environmental health, researchers are mainly concerned with data acquired over a period of days extending to entire lifetimes, with a resolution of hours to years. The research will develop new methods for the analysis of geographically and temporally referenced medical information and new methods for reasoning about environmental exposures and their consequences over space and through time (Mark et al. 1999). The methods will also be applicable to hazardous exposures of shorter time periods with more immediate impacts. The recent interest in moving points within the spatiotemporal database community further suggests that methods based on moving points (Erwig et al. 1998), and particularly on Hägerstrand's time geography model (Fauvet et al. 1998, 1999, Dumas et al. 1999), will become realistic tools in environmental health in the near future. Recently, Forer (1998) articulated time-space primitives, defining timelines and activity volumes that have direct implications for locational histories in health research. Further articulation of time geography concepts in a GIS environment and their evaluation with empirical data are essential.

This research project will also examine statistical approaches for identifying clusters of hot spots of ill health in space-time. Often, it is important to determine whether observations of some phenomenon are clustered in space, time, or both. When attempting to determine the causes of an outbreak or chronic pattern of ill health, analysts frequently plot the distributions of cases on maps. This method has been used at least since Dr. John Snow's now famous map of cholera deaths in London, England, which helped identify a particular public water pump as the source of the epidemic (Snow 1936, McLeod 2000). For infectious diseases with short incubation periods, an analysis of the spatial distribution alone may be sufficient. There are problems, however, with such methods in the study of environmentally induced diseases with long latency periods, such as many forms of cancer, since those stricken with cancer could have changed their residence several times since their exposure to environmental hazards, thus breaking up clusters and obscuring patterns.

Methods of analysis based on reasoning about geospatial lifelines of specific cases may reduce or eliminate the problem of cluster dispersion. If researchers have the data and information-manipulation tools, cases can be rolled back to places of residence or travel when the cases might have been more markedly clustered. Clusters could also be identified directly in three-dimensional space-time. A discussion of how Finnish census data are uniquely organized to permit the tracing of residences to places with known radon level measurements was described by Loytonen (1998).

Integrate the Literature of Spatial Choice in Geography and Econometrics with the Literature of Preventive Care Choices

Many critical choices that affect a person's health are made where the controlling factors are in a spatial context. The decision of when and where to seek health care is known to be affected by the geographical distribution of relevant resources, hence, the importance of geographic accessibility in seeking timely medical care. It is possible, for example, that the development stage of a tumor at the time of first diagnosis might relate to the choice of place and type of treatment (see Fortney et al. 1995, 1998).

Propose Additional Systematic Studies of Access, Health Treatment Choice, and Health Outcomes

The traditional spirit of public health has always been a focus on the health of the public. Consequently, there is concern when particular population groups experience a greater burden of disease (Townsend et al. 1988; see Cohen and Lee 1985, Piette and Moos 1996, Gober 1997, Siegel et al. 1997, McLafferty 1988).

The issues of scale relate to geographic information and analysis. One needs to differentiate between geographical scale (the areal units for which information is processed), and data resolution (the granularity of the data). Other issues include the "cartographic scale" and the "extent" of the analysis (neighborhood, urban, regional, state). A recent paper by Sexton et al. (2000), described many of the issues related to public health and scale. For public health analyses, an important consideration is the Modifiable Areal Unit Problem (MAUP). Increasingly, researchers are relying on governmentally derived public health data for assessing the potential populations at risk and for more robust standardization of census data (Waller and McMaster 1997). The MAUP relates to the geographical concern that results change as the boundaries of units of analysis change. Thus, one result is obtained if a researcher uses census Minor Civil Divisions in a health analysis, while a different conclusion may be obtained if other boundaries of units of the same size are used. In addition, as the geographic scale of analysis changes, the results may change. Researchers must be keenly aware of this, and attempt to tease out the "ideal" resolution or resolutions. Alternatively, researchers completing such multi-scale/multi-resolution analyses have been encouraged to present the results of each and to allow the user to determine why the results are scale-dependent.

Another potential data challenge occurs when using census information for risk assessment and standardization. As we move further away from the census year, the data become increasingly obsolete, with differential spatial and thematic error. For instance, certain neighborhoods in a city will witness remarkable social changes between the censuses, whereas other, more stable neighborhoods (often much wealthier neighborhoods) see little change in their socioeconomic conditions and race/ethnicity structures. Age data, however, change dramatically in all areas by the end of the 10-year period, to the point of being non-usable. One example would be using data from the 1990 census in 1999 for a health survey attempting to measure characteristics of children under 5.

Develop Methods for Targeting Health Resources

There is agreement that one important use of GIS is to target health resources to places where they are most in need (Larimore and Davis 1995, Geronimus et al. 1996, Bureau of Primary Health Care 1998). The health science community is often unaware of the extent of the development of general methods in geography and regional science for this purpose (Ayeni et al. 1987, Malczewski and Ogryczak 1988, McLafferty and Broe 1990, Walsh et al. 1997). The Federal Department of Health and Human Services, for example, assigns resources for reducing the rate of infant mortality in areas of U.S. cities where the rates are double the overall U.S. rate. The program, called "Healthy Start," currently operates in selected areas in at least 50 U.S. cities. Methodologies for selecting these areas involve ad hoc methods of regionalization. Evaluating the geographical effects of policies designed to improve access to health services is another area of application of GIS in health. The problem is well illustrated in the General Accounting Office's (1997) review of the implementation of the Rural Health Care Centers Program of the Department of Health and Human Services.

Using GIS to Improve Ways to Communicate to the Public the Results of Research on Health

Maps indicating the locality of disease in local areas are difficult to interpret. Indeed, they may be misinterpreted by a public not aware that, in small areas, there are naturally marked geographic variations, even when the true disease rates do not vary. Maps are also open to interpretation by groups whose purpose may include the deliberate manipulation of public opinion toward some end. The public is often suspicious that information is being withheld from their view. The challenge for the scientist is to assist in the interpretation of results from GIS-based analyses. Pickle and others (1995, 1997) have reported on the extensive perception research undertaken in conjunction with the *Atlas of United States Mortality* to reduce unintended interpretations. The public health community has had a focus on small-area analysis for some time, but the issue of misinterpretation is not resolved by more efficient algorithms (e.g., those of Carvalho et al 1996). Headlines in the popular press can swiftly present conclusions that are not supported by scientific analysis. When health risks and the environment are the subjects, it is important that information be presented to help the public sort through material that is often conflicting. Monmonier (1997) has written about risk communication in the health area. There have been a number of recent proposals to the NIH for projects to support the development of Web-based mapping that the public can access.

Related to another societal issue of GIS, researchers must also consider the ethics behind certain types of analysis. For instance, when maps indicating morbidity and mortality of certain types of diseases are made public without appropriate explanation, undue stress can result in affected communities. It is clear that, as GISs become commonplace in public health analysis, there are a plethora of societal concerns that must be addressed with caution.

Maintaining the Confidentiality of Health Records

Personal health records are sensitive, confidential pieces of information on individuals, and many laws exist to ensure the privacy of individuals and the protection of information from those who have no right to view it. Releasing health information data for small areas may result in the failure to protect the privacy of individuals. Often, the desire to see health data in its geographic context is in conflict with protecting the confidentiality of individuals. Methods need to be developed for ensuring confidentiality while preserving the capability of geographical analysis. A report of the NCI (1999: 30) recognized this problem in the context of GIS.

However, it is critical that mechanisms for protecting confidentiality be developed to maximize the utility of this technology. Spatial aggregation, which has been the standard method for preserving confidentiality of geographic data, will not suffice for health-related GIS activities. The SIG (special interest group) recommends that research be conducted to develop alternative methods to guard the privacy of health records incorporated in GIS-based geographic analysis.

Only a small volume of literature has addressed this issue (Armstrong et al. 1999). These problems affect spatial data acquisition. Public health researchers encounter a variety of difficulties in the acquisition of appropriate geospatial data. Basic public health data records, as mentioned before, often have strict confidentiality and privacy concerns resulting in the release of records at only crude levels such as counties. Additionally, much of the data, although tagged with addresses, requires significant effort in geocoding if address-specific release is allowable. Although improvements have been made over the past decade, accurate geocoding is still a time-consuming procedure fraught with complications, including inaccurate and out-of-date geographical databases, typographical errors in the addresses, and address-range problems in rural areas.

Policy Implications

One role of the UCGIS is to influence the formulation of policy that contributes to the advance of GIScience and technology. Within the domain of health and public services, such policy is in the hands of the U.S. Congress, the Executive Branch, and in the budgetary priorities of the mainline agencies, such as the Department of Health and Social Security and the Veterans' Administration. Additionally, policy direction and priority are of significance in the CDC, the NIH, the National Science Foundation, and in numerous additional groups who access and use geographic information. Good policy would advance GIScience for public health and human services by establishing "best practice" for spatial analysis within GIS. It would also address the need for standardization of geospatial data used in public health and services applications and in establishing standards for privacy and the ethical use of geographic information. GIS can help target need in reducing the burden of premature mortality and morbidity and providing health care for the elderly and uninsured. The incorporation of GIS should enable higher quality needs assessment and intervention evaluation. Peer review organizations, such as the American Quality of Health Association, will increasingly use GIS for evaluation and assessment of the geographical equity of the use of procedures and interventions, especially in the assessment of the effectiveness of Medicare/Medicaid funds (see Durch et al. 1997).

Recent developments have seen the emergence of a policy directed at national medical readiness in the face of mass casualty events. Massive casualties anywhere in the U.S. due to natural disasters, bioterrorism, or weapons of mass destruction would quickly overwhelm the current medical care establishment. As yet, few places have been able to adopt lessons learned from the Oklahoma bombing in terms of the instantaneous need for large quantities of diverse information coordinated primarily by geography. It could be argued that the nation is fundamentally unprepared for a concerted response to major disasters in densely populated areas and is in urgent need of policy to redress this deficiency.

Perhaps the overriding health information policy issue at the turn of the 21st century is that of privacy. The public senses considerable vulnerability in the face of the technical ability of corporations to combine personal data from multiple sources. Widespread availability and sharing of geographic data exacerbates questions of confidentiality. At this point, the search for answers continues as many groups grapple with the competing claims from those who see that the sharing of personal data has benefits to the individual and the society and those who are concerned with the potential abuses of privacy that such capabilities present.

Authors

Gerard Rushton is a Professor of Geography and Adjunct Professor of Public Health at the University of Iowa. He was awarded a B.A. and M.A. from the University of Wales and a Ph.D. from the University of Iowa. His research interests are in spatial decision support systems, location modeling, and the spatial analysis of disease distributions.

Gregory Elmes is a Professor of Geography at West Virginia University. His current research interests are in GIS and spatial analysis, especially as applied to medical geography, and the role of geographic information in society.

Robert B. McMaster is an Associate Professor of Geography at the University of Minnesota, where he serves as Director of the Master of Geographic Information Science Program. He earned a B.A. from Syracuse University (1978) and a Ph.D. from the University of Kansas (1983). His research interests are in environmental risk assessment and modeling (in particular environmental justice and GIS), cartographic generalization, and the history of academic cartography.

Acknowledgments

The authors wish to thank David Mark and Max Egenhofer for contributing the section "Temporal Aspects of GIS and Health" and Jamil Kazmi and Lynn Usery for the section "Ecological Studies of the Relationship between Environmental Factors and Disease Transmission." Thanks are also given to Art Getis for the many valuable suggestions and two other reviewers for their comments.

References

Anselin, L., 1995, Local Indicators of Spatial Association—LISA. *Geographical Analysis,* 27, 93-115.

Anselin, L. and S. Bao, 1997, Exploratory Spatial Data Analysis Linking SpaceStat and ArcView. In Fischer M. and A. Getis (Eds.), *Recent Developments in Spatial Analysis: Spatial Statistics., Behavioural Modelling and Neuro-computing* (Berlin: Springer-Verlag), 35-59.

Armstrong, M.P., G. Rushton, and D.L., Zimmerman, 1999, Geographically Masking Health Data to Preserve Confidentiality. *Statistics in Medicine,* 18 497-525.

Ayeni, B., G. Rushton, and M. McNulty, 1987, Improving the Geographical Accessibility of Health Care in Rural Areas: A Nigerian Case Study. *Social Science and Medicine,* 25, 1083-1094.

Barnett, E., Elmes G., Braham V., Halverson J., Lee, J., and Loftus., 1998, Heart Disease in Appalachia: An Atlas of County Economic Conditions, Mortality and Medical Care Resources, *Office for Social Environment and Health Research*, West Virginia University, Morgantown WV: 94pp.

Becker, K.M., G.E. Glass, W. Brathwaite, and J.M. Zenilman, 1998, Geographic Epidemiology of Gonorrhea in Baltimore, Maryland, Using a Geographic Information System. *American Journal of Epidemiology,* 147, 709-716.

Bernhardsen, T., 1999, *Geographic Information Systems: An Introduction* (New York: John Wiley & Sons, Inc.).

Braden, J. and K. Beauregard, 1994, *Health Status and Access to Care of Rural and Urban Populations* (AHCPR Pub. No. 94-0031). National Medical Expenditure Survey Research Findings 18, Agency for Health Policy and Research (Rockville, MD: Public Health Service).

Briggs, D.J. and P. Elliott, 1995, The Use of Geographical Information Systems in Studies on Environment and Health. *World Health Statistics Quarterly,* 48, 85-94.

Brown, P.J.B., A. Hirschfield, and P.W.J. Batey, 1991, Applications of Geodemographic Methods in the Analysis of Health Condition Incidence Data. *Papers in Regional Science: Journal of the Regional Science Association International,* 70, 329-344.

Bureau of Primary Health Care, 1998, Procedures for Medically Underserved Areas and Medically Underserved Populations Designation http://www.vdh.state.va.us/primcare/0143.htm

Carvalho, M.S., O.G. Cruz, and F.F. Nobre, 1996, Spatial Partitioning Using Multivariate Cluster Analysis and a Contiguity Algorithm. *Statistics in Medicine,* 15, 1885-1894.

Casper M.L., Barnett E., Halverson J.A., Elmes G.A., Braham V.E., Majeed Z.A., Bloom A.S., Stanley S. 1999. Women and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality. *Office for Social Environment and Health Researc*h, West Virginia University, Morgantown WV:240 pp.

Clarke, K.J.C., S.L. McLafferty, and B.J. Tempalski, 1996, On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Diseases,* 2(2), 85-92.

Cohen, M. and H. Lee, 1985, The Determinants of Spatial Distribution of Hospital Utilization in a Region. *Medical Care,* 23, 27-38.

Cook, R.L., R.A. Royce, J.C. Thomas, and B.H. Hanusa, 1999, What's Driving an Epidemic? The Spread of Syphilis Along an Interstate Highway in Rural North Carolina. *American Journal of Public Health,* 89, 369-373.

Cromley, E.K. and G.W. Shannon, 1986, Locating Ambulatory Medical Care Facilities for the Elderly. *Health Services Research,* 21, 499-513.

Croner, C.M., J. Sperling, and F. Broome, 1996, Geographic Information Systems (GIS): New Perspectives in Understanding Human Health and Environmental Relationships. *Statistics in Medicine,* 15, 1961-1977.

Croner, C.M. and D.F. Stroup, 1999, GIS and EIS: Geographic Information Systems and the Epidemic Intelligence Service. *Epidemic Intelligence Service Bulletin,* Fall/Winter.

Dean, A.G., 1999, Epi Info and Epi Map: Current Status and Plans for Epi Info 2000. *Journal of Public Health Management and Practice,* 5(4), 54-57.

Dumas, M., M.-C. Fauvet, and P.-C. Scholl, 1999, TEMPOS: A Temporal Database Model Seamlessly Extending ODMG, Research-Report 1013-I-LSR-7, LSR-IMAG Laboratory, Grenoble, March.

Durch, J.S., L.A. Bailey, and M.A. Stoto (Eds.), 1997, *Improving Health in the Community: A Role for Performance Monitoring*. (Washington, DC: Institute of Medicine (U.S.), Committee on Using Health Performance Monitoring to Improve Health, National Academy of Sciences).

Elliott, P., J. Cuzick, D. English, and R. Stern (Eds.), 1992, *Geographical and Environmental Epidemiology: Methods for Small Area Studies* (Oxford: Oxford University Press).

Entwisle, B., R.R. Rindfuss, et al., 1997, Geographic Information Systems, Spatial Network Analysis, and Contraceptive Choice. *Demography,* 34, 171-187.

Erwig, M., R.H. Güting, M. Schneider, and M. Vazirgiannis, 1998, Spatio-Temporal Data Types: An Approach to Modeling and Querying Moving Objects in Databases. Fern Universität Hagen, Informatik-Report 224, December 1997. Abstract, Paper.ps.gz (compressed postscript, 52 KB). Revised Version, August.

Fauvet, M.C., S. Chardonnel, M. Dumas, P.C. Scholl, and P. Dumolard, 1998, Analyze de données géographiques: application des bases de données temporelles. *Révue Internationale de Géomatique,* 8(1-2), 149-165.

Fauvet, M.C., S. Chardonnel, M. Dumas, P.C. Scholl, and P. Dumolard, Forthcoming, Applying Temporal Databasesto Geographical Data Analysis. In *Proceedings of the DEXA Workshop on Spatio-Temporal Data Models and Languages, Florence Italy, September, 1999.*

Forer, P., 1998, Geometric Approaches to the Nexus of Time, Space, and Microprocess: Implementing a Practical Model for Mundane Socio-spatial Systems. In Egenhofer, M. and R. Golledge (Eds.), *Spatial and Temporal Reasoning in Geographic Information Systems* (New York: Oxford University Press), 171-190.

Fortney, J., B. Booth, F. Blow, J. Bunn, and C. Cook, 1995, The Effects of Travel Barriers and Age on the Utilization of Alcoholism Treatment Aftercare. *American Journal of Drug and Alcohol Abuse,* 21, 391-406.

Fortney, J., K. Rost, M. Zhang, and G.R. Smith, 1998, A Joint Choice Model of the Decision to Seek Depression Treatment and Choice of Provider Sector. *Medical Care*.

Gatrell, A.C. and T.C. Bailey, 1966, Interactive Spatial Data Analysis in Medical Geography. *Social Science and Medicine,* 42, 843-855.

Gatrell, A.C. and M. Loytonen, 1998, *GIS and Health* (London: Taylor & Francis Ltd).

General Accounting Office, 1997, *Rural Health Clinics: Rising Program Expenditures not Focused on Improving Care in Isolated Areas* (Washington, DC: GAO/HeHS-97-24).

Geronimus, A., J. Bounnd, and L. Neidert, 1996, On the Validity of Using Census Geocode Characteristics to Proxy Individual Socioeconomic Characteristics. *Journal of the American Statistical Association,* 91, 529-537.

Getis, A. and J.K. Ord, 1992, The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24, 189-206.

Glass, G.E., B.S. Schwartz, et al., 1995, Environmental Risk Factors for Lyme Disease Identified with Geographic Information Systems. *American Journal of Public Health,* 85, 944-948.

Gober, P., 1997, The Role of Access in Explaining State Abortion Rates. *Social Science and Medicine*, 44, 1003-1016.

Gould, P., 1993, *The Slow Plague: A Geography of the AIDS Pandemic* (Oxford: Blackwell Scientific).

Hägerstrand, T., 1967, *Innovation Diffusion as a Spatial Process* (Chicago: University of Chicago Press).

Hägerstrand, T., 1970, What About People in Regional Science? *Papers, Regional Science Association,* 24, 1-21.

Haining, R., 1990, *Spatial Data Analysis in the Social and Environmental Sciences* (New York: Cambridge University Press).

Haining, R., S.M. Wise, and M. Blake, 1994, Constructing Regions for Small Area Analysis: Material Deprivation and Colorectal Cancer. *Journal of Public Health Medicine,* 16, 429-438.

Haining, R., S. Wise, and J. Ma, 1998, Exploratory Spatial Data Analysis in a Geographic Information System Environment. *The Statistician,* 47(3), 457-469.

Haslett, J., G. Wills, and A.R. Unwin, 1990, SPIDER—an Interactive Statistical Tool for the Analysis of Spatially Distributed Data. *International Journal Geographical Information Systems,* 4, 285-296.

Hirschfield, A., P. Brown, and P. Bundred, 1993, Doctors, Patients and GIS. *Mapping Awareness,* 7(9)

Hirschfield, A., P. Brown, and P. Bundred, 1995, The Spatial Analysis of Community Health Services on Wirral Using Geographic Information Systems. *Journal of the Operational Research Society,* 46, 147-159.

Jacquez, G.M, 1995, The Map Comparison Problem: Tests for the Overlap of Geographic Boundaries. *Statistics in Medicine,* 14, 2343-2361.

Kerner, J.F., H. Andrews, A. Zauber, and E. Struening, 1988, Geographically Based Cancer Controls: Methods for Targeting and Evaluating the Impact of Screening Interventions on Defined Populations. *Journal of Clinical Epidemiology,* 41, 543-553.

Kohli, S., K. Sahlén, O. Löfman, Å Sivertun, M. Foldevi, E. Trell, and O. Wigertz, 1997, Individuals Living in Areas with High Background Radon: A GIS Method to Identify Populations at Risk. *Computer Methods and Programs in Biomedicine,* 53, 105-112.

Kulldorff, M., et al., 1998, SaTScan V. 2.1, Software for the Spatial and Space-time Scan Statistics (Bethesda, MD: National Cancer Institute).

Kulldorff, M., 1999, Geographic Information Systems (GIS) and Community Health: Some Statistical Issues. *Journal of Public Health Management and Practice,* 5(2*)*, 100-106.

Langran, G., 1992, *Time in Geographic Information Systems* (London: Taylor and Francis).

Larimore, W. and A. Davis, 1995, Relation of Infant Mortality to the Availability of Maternity Care in Rural Florida. *Journal of the American Board of Family Practice,* 8, 392-399.

Lee, R., 1991, Current Approaches to Shortage Area Designation. *The Journal of Rural Health,* 7, 437-450.

Love, D. and P. Lindquist, 1995, The Geographical Accessibility of Hospitals to the Aged: a Geographic Information Systems Analysis Within Illinois. *Health Services Research,* 29, 629-651.

Loytonen, M, 1998, GIS, Time Geography and Health. In Gatrell A.C. and M. Loytonen (Eds.), *GIS and Health* (London: Taylor & Francis Ltd.), 97-110.

MacDorman, M.F. and G.A. Gay, 1999, State Initiatives in Geocoding Vital Statistics Data. *Journal of Public Health Management and Practice,* 5(2), 91-93.

MacEachren, A.M. and D. DiBiase, 1991, Animated Maps of Aggregate Data: Conceptual and Practical Problems. *Cartography and Geographic Information Science,* 18, 221-229.

Mackinnon, D.P., R. Scribner, and K.A. Taft, 1995, Alcohol Availability and Alcohol Problems Database. *Statistics in Medicine,* 14, 591-603.

Malczewski, J. and W. Ogryczak, 1988, A Multi-objective Approach to the Reorganization of Health Services Areas: A Case Study. *Environment and Planning, A,* 20, 1461-1470.

Mark, D.M. and M.J. Egenhofer, 1998, Geospatial Lifelines. In Guenther, O., T. Sellis, and B. Theodoulidis (Eds.), *Integrating Spatial and Temporal Databases*. Dagstuhl Seminar Report No. 228.

Mark, D.M., M. Egenhofer, L. Bian, P.A. Rogerson, and J. Vena, 1999, Spatio-Temporal GIS Analysis for Environmental Health. National Institute of Environmental Health Sciences, National Institutes of Health, Grant No. 1 R 01 ES09816-01.

Martin, D., 1996, *Geographic Information Systems: Socioeconomic Applications* (London: Routledge).

McLafferty, S., 1988, Predicting the Effect of Hospital Closures on Hospital Utilization Patterns. *Social Science and Medicine,* 27, 255-262.

McLafferty, S. and D. Broe, 1990, Patient Outcomes and Regional Planning of Coronary Care Services: A Location-allocation Approach. *Social Science and Medicine,* 30, 297-304.

McLeod, K.S., 2000, Our Sense of Snow: The Myth of John Snow in Medical Geography. *Social Science and Medicine,* 50, 923-935.

Miller, H.J., 1991, Modeling Accessibility Using Space-Time Prism Concepts within a GIS. *International Journal of Geographic Information Systems,* 5(3), 287-301.

Mohan, J., 1983, Location-allocation Models, Social Science and Health Service Planning: An Example from North East England. *Social Science and Medicine,* 17, 493-499.

Monmonier, M., 1997, *Cartographies of Danger: Mapping Hazards in America* (Chicago: The University of Chicago Press), Chapter 13, John Snow's Legacy.

Morrison, A.C., A. Getis, M. Santiago, J.G. Rigau-Perez, and P. Reiter, 1998, Exploratory Space-Time Analysis of Reported Dengue Cases During an Outbreak in Florida, Puerto Rico, 1991-1992, *American Journal of Tropical Medicine and Hygiene*, 58, 287-298.

National Cancer Institute, 1999a, The Long Island Breast Cancer Study Project, http://www-dccps.ims.nci.nig.gov/LIBCSP/index.html

National Cancer Institute, 1999b, Cancer Surveillance Research Implementation Plan. Surveillance Implementation Group, NCI, March, 30.

Ord, J.K. and A. Getis, 1995, Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis,* 27, 286-306.

Perkins, B.B., 1999, Re-forming Medical Delivery Systems: Economic Organization and Dynamics of Regional Planning and Managed Competition. *Social Science and Medicine,* 48, 241-251.

Peuquet, D.J., 1994, It's About Time: a Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers,* 84(3), 441-461.

Pickle, L.W., M. Mungiole, G.J. Jones, and A. White, 1997, *Atlas of United States Mortality* (National Center for Health Statistics, Centers for Disease Control and Prevention).

Pickle, L.W. and A. White, 1995, The Choice of Age-adjustment Methods on Maps of Death Rates. *Statistics in Medicine,* 14, 615-627.

Piette, J.D. and R.H. Moos, 1996, The Influence of Distance on Ambulatory Care Use, Death, and Readmission Following a Myocardial Infarction. *Health Services Research,* 31, 573-591.

Richards, T.B., C.M. Croner, G. Rushton, C.K. Brown, and L. Fowler, 1999, Geographic Information Systems and Public Health: Mapping the Future. *Public Health Reports,* 114, 359-373.

Ricketts, T.C., L.A. Savitz, W.M. Gesler, and D.N. Osbourne, 1994, *Geographic Methods for Health Services Research : A Focus on the Rural-Urban Continuum* (Lanham, MD: University Press of America).

Rosenberg, M.S., R.R. Sokal, N.L. Oden, and D. DiGiovanni, 1999, Spatial Autocorrelation of Cancer in Western Europe. *European Journal of Epidemiology,* 15, 15-22.

Ross, N., M. Rosenber, and D. Pross, 1994, Siting a Women's Health Facility: A Location-allocation Study of Breast Cancer Screening Services in Eastern Ontario. *The Canadian Geographer,* 38, 150-161.

Rushton, G., 1998, Improving the Geographic Basis of Health Surveillance Using GIS. In Loytonen, M. and A. Gatrell (Eds.), *GIS and Health* (London: Taylor & Francis), 63-79.

Rushton, G., 1999, Methods to Evaluate Geographic Access to Health Services. *Journal of Public Health Management and Practice,* 5(2), 93-100.

Rushton, G., M.P. Armstrong, C. Lynch, and J. Rohrer, 1997, Improving Public Health Through Geographical Information Systems: an Instructional Guide to Major Concepts and Their Implementation [CD-ROM] (Iowa City: University of Iowa Department of Geography). Available at URL, http://www.uiowa.edu/~geog/health/

Sexton, K., L.A. Waller, R.B. McMaster, G. Maldonado, and J.L. Adgate, 2000, The Importance of Spatial Scale for Environmental Health Policy Research. In *Human and Ecological Risk Assessment.*

Siegel, C., A. Davidson, K. Kafadar, J.M. Norris, J. Todd, and J. Steiner, 1997, Geographic Analysis of Pertussis Infection in an Urban Area: A Tool for Health Services Planning. *American Journal of Public Health,* 87, 2022-2026.

Spear, R., P. Gong, E. Seto, Y. Zhou, B. Xu, D. Maszle, S. Liang, G. Dagis, and X. Gu, 1998, Remote Sensing for Schistosomiasis Control in Mountainous Areas in Sichuan, China. *Geographic Information Sciences,* 4, 14-22.

Smoyer, K.E., 1998, Putting Risk in its Place: Methodological Considerations for Investigating Extreme Event Health Risk. *Social Science and Medicine,* 47, 1809-1824.

Snow, J., 1936, *Snow on Cholera* (London: Oxford University Press).

Unwin, A. and D. Unwin, 1998, Exploratory Spatial Data Analysis with Local Statistics. *The Statistician,* 47(3), 415-421.

Vine, M.F., D. Degnan, and C. Hanchette, 1997, Geographic Information Systems: Their Use in Environmental Epidemiologic Research. *Environmental Health Perspectives,* 105(6), 598-605.

Waller, L.A., 1996, Geographic Information Systems and Environmental Health. *Health and Environment Digest*, 9, 85-88.

Waller, L.A. and R.B. McMaster, 1997, Incorporating Indirect Standardization in Tests for Disease Clustering in a GIS Environment. *Geographical Systems,* 4, 327-342.

Walsh, S., P. Page, and W. Gesler, 1997, Normative Models and Healthcare Planning: Network-based Simulations Within a Geographic Information System Environment. *Health Services Research,* 32, 243-260.

Wise, S.M., R.P. Haining, and J. Ma, 1997, Regionalization Tools for the Exploratory Spatial Analysis of Health Data. In Fischer M. and A. Getis (Eds.), *Recent Developments in Spatial Analysis: Spatial Statistics, Behavioural Modelling and Neuro-computing* (Berlin: Springer-Verlag), 83-100.

World Health Organization, 1996, *Climate Change and Human Health* (Geneva).

Yasnoff, W.A. and E.J. Sondik, 1999, Geographic Information Systems (GIS) in Public Health Practice in the New Millennium. *Journal of Public Health Management and Practice,* 5(4), ix-xii.

Appendix A

GIS Functions Identified by the National Cancer Institute as Necessary or Desirable Functions for Its GIS for the Long Island Breast Cancer Project

June 1998

Function Name / Citation

Address Matching with TIGER Files
    Broome and Meixler, 1990
    Gribb et al., 1990
    Marx, 1990
    Rizzardi et al., 1993
    Rushton and Lolonis, 1996
    U.S. Bureau of the Census, 1992
Aggregation Effects (see also Geographic Scale Analysis)
    Clark and Avery, 1976
    Waller, 1996
Areal interpolation
    Fisher and Langford, 1995
    Flowerdew and Green, 1989
    Goodchild and Lam, 1980
    Goodchild et al., 1993
Bayes -see Empirical Bayes
Buffering
    Feychting and Ahlbom, 1993
    Wartenberg et al., 1993
Confidence Intervals for Age-Adjusted Rates
    Dobson et al., 1991
Confidentiality Issues (see also Spatial Masks)
    Cox, 1996
    Duncan and Lambert, 1989
    Duncan and Pearson, 1991
    Institute of Medicine, 1994
    Jacquez, 1994, 1996
    Jacquez and Waller, 1996
    U.S. Office of Management and Budget, 1994
Confounding factors
    Kelsey and Horn-Ross, 1993
Constrained Bayes - see Empirical Bayes
Cluster Detection
    Approaches
    Grimson and Oden, 1996
    Neutra et al., 1992
    Smith and Neutra, 1993
    Waller et al., 1994
Density Estimation - see Spatial Filters
Demographics - Small Area Population Estimates
    Byerly, 1990
    Ericksen, 1974
    Isserman, 1977 & 1984
    National Research Council, 1980

- Disease Cluster Investigations (see also Cluster, approaches)
  - Marshall, 1991
  - Neutra, et al. 1992
- Disease Mapping
  - Cliff and Haggett, 1988
  - Haybittle et al., 1995
  - Hoover et al., 1975
  - Marshall, 1991
  - Olson, 1975
  - Pickle et al., 1987, 1990
  - Walter and Birnie, 1991
  - Walter, 1993
- Ecological Analysis
  - Blot et al., 1978
  - Goodman et al., 1989
  - Openshaw, 1984
  - Robinson, 1950
- Empirical Bayes Estimates
  - Bernardinelli and Montomoli, 1992
  - Clayton and Kaldor, 1987
  - Devine and Louis, 1994
  - Devine et al., 1994, 1996
  - Ghosh, 1993
  - Langford, 1994
  - Louis, 1984
  - Manton et al., 1989
  - Mollie and Richardson, 1991
  - Stabilized Rates
  - Chen, 1996 (1994?)
  - Clayton and Kaldor, 1987
  - Kennedy-Kalafatis, 1995
    - Moulton et al., 1994
- Exposure Analysis
  - Louis, 1984
- GIS-H
  - Clarke et al., 1996
- Geographic Scale Analysis
  - Amrhein and Reynolds, 1996
  - Cleek, 1979
  - Holtet al., 1996
  - Moellering and Tobler, 1972
  - Munasinghe and Morris, 1996
  - Rushton and Lolonis, 1996
  - Schneider et al., 1993
  - Waller and Turnbull, 1993
- Interactive Spatial Data Analysis
  - Bailey & Gatrell, 1995
  - Gatrell and Bailey, 1996
- Interpolation - see Areal Interpolation
- Kriging
  - Carrat and Valleron, 1992
  - Webster et al., 1994
- Map Overlay
  - Tomlin, 1990
- Methods
  - Besag and Newell, 1991
  - Cuzick and Edwards, 1990
  - Fotheringham and Zhan, 1996
  - Grimson, 1993
  - Kingham et al., 1995
  - Kulldorf and Nagarwalla, 1995
  - Le, Petkau, and Rosychuk, 1996
  - Mantel et al., 1976
  - Mantel, 1967
  - Marshall, 1991
  - Oden et al., 1996
  - Pompe-Kirn and Ferligoj, 1991
  - Stone, 1988
  - Wojdyla et al., 1996
  - Focused Tests
  - Bithell, 1992
  - Hills and Alexander, 1989
  - Lagakos et al., 1986
  - Lawson, 1993
    - Le et al., 1996
  - Stone, 1988
  - Waller and Lawson, 1995
  - Waller et al., 1992, 1994
  - Hierarchical Clustering
    - Grimson et al., 1981
  - Point-vs. Area-based Measures
  - Oden et al., 1996
  - Point Pattern Analysis (see also Scan Statistic)
    - Boots and Getis, 1988
  - Diggle, 1983
- Modifiable Areal Unit
  - Fotheringham and Wong, 1991
- Point Pattern Analysis (see also Software for Spatial Analysis)
  - Cliff and Ord, 1975
  - Oden, 1995
- Poisson Distributions
  - Aickin et al., 1992
  - Downer, 1996
  - Reynolds et al., 1996
  - Schlattmann et al., 1996
- Poisson Regression Models
  - Frome and Checkoway, 1985
- Power simulations
  - Oden et al., 1996
  - Waller, 1996
  - Wartenberg and Greenberg, 1990
- Proximity Analysis (see also Buffering)
  - Geschwind et al., 1992

- Scan Statistic
    - Cressie, 1977
    - Glaz, 1993
    - Hjalmars et al., 1996
    - Hryhorczuk et al., 1992
    - Nagarwalla, 1996
    - Naus, 1965, 1966
    - Wallenstein and Neff, 1987
    - Wallenstein et al., 1993
    - Weinstock, 1981
- Small-Area Variation
    - Cliff and Haggett, 1988
    - Elliot et al., 1992
    - Ghosh and Rao, 1994
    - Reynolds et al., 1996
    - Weiss and Wegener, 1990a, b
- Smoothing Maps - see Spatial Filters
- Software for Spatial Analysis
    - Clustering
        - Hall et al., 1996
    - Jacquez, 1994
    - General
        - Biomedware, http://www.biomedware.com/
    - McCune and Mefford, 1995
    - Mapping
    - Dean, 1993
    - Schlattman, 1996
        - Strassburg and Williams, 1995
    - Point Pattern Analysis
        - Rowlingson and Diggle, 1993
    - Skelton, 1996
- Spatial Autocorrelation
    - Glick, 1979
    - Jacquez, 1992
    - Moran, 1950
    - Munasinghe and Morris, 1996
    - Sokal et al., 1993
    - Waldhor, 1996
- Spatial Clustering Algorithms
    - Huel et al., 1986
    - Morris and Munasinghe, 1993
- Spatial Data Analysis
    - Cliff and Ord, 1981
    - Haining, 1990
    - King, 1979
    - Ripley, 1981
    - Upton and Fingleton, 1985
- Spatial Filters (see also Scan Statistic)
    - Silverman, 1978, 1986
    - Bithell, 1990
    - Cressie, 1992
    - Cressie and Read, 1989
    - Rushton and Lolonis, 1996
- Spatial Masks
    - Jacquez, 1996
- Space-Time Pattern Analyzer
    - Alexander, 1992
    - Chen et al., 1984
    - Ederer et al., 1964
    - Knox, 1964
    - McAuliffe and Afif, 1984
    - McKnight et al., 1996
    - Moran, 1950
    - Openshaw, 1994
- Spatial Time Series
    - Bennett, 1979
    - Geographical Analysis Machine
        - Openshaw et al., 1987, 1988a, b
    - Klauber and Angulo, 1974
- Standardized Incidence Ratio
    - Maskarinec, 1996
- Statistical Power
    - Walter, 1992a, 1996
    - Wartenberg and Greenberg, 1990
- Temporal Clustering
    - Wallenstein, 1980
- TIGER - see Address Matching with TIGER files
- Visualization
    - Dorling and Openshaw,1992

References

Aickin, M., C.A. Chapin, T.J. Flood, S.J. Englender, and G.C. Caldwell, 1992, Assessment of the Spatial Occurrence of Childhood Leukemia Mortality Using Standardized Rate Ratios with a Simple Linear Poisson Model. *International Journal of Epidemiology,* 21, 649-655.

Alexander, F.E., 1992, Space-Time Clustering of Childhood Acute Lymphoblastic Leukemia: Indirect Evidence for a Transmissible Agent. *British Journal of Cancer*, 65(4), 589-592.

Amrhein, C.G. and H. Reynolds, 1996, Using Spatial Statistics to Assess Aggregation Effects. *Geographical Systems,* 3, 143-158.

Bailey, T.C. and A.C. Gatrell, 1995, *Interactive Spatial Data Analysis* (London: Longman).

Bennett, R.J., 1979, *Spatial Time Series* (Pion: London).

Bernardinelli, L. and C. Montomoli, 1992, Empirical Bayes versus Fully Bayesian Analysis of Geographical Variation in Disease Risk. *Statistics in Medicine,* 11, 983-1007.

Besag, J. and J. Newell, 1991, The Detection of Clusters in Rare Diseases. *Journal of the Royal Statistical Society Series A*, 154, 143-155.

Biomedical Statistical Software, 1983 (Berkeley, CA: University of California Press).

Bithell, J.F., 1990, An Application of Density Estimation to Geographical Epidemiology. *Statistics in Medicine,* 9, 691-701.

Bithell, J.F., 1992, Statistical Methods for Analyzing Point-Source Exposures. In Elliot, P., J. Cuzick, D. English, and R. Stern (Eds.), *Geographical and Environmental Epidemiology: Methods for Small-Area Studies* (Oxford: Oxford University Press).

Blot, W.J., J. Fraumeni, F. Jun, and B.J. Stone, 1978, Geographic Correlates of Pancreas Cancer in the United States. *Cancer,* 42, 373-380.

Boots, B. and A. Getis, 1988, *Point Pattern Analysis* (Beverly Hills, CA: Sage Publications).

Brillinger, D.R., 1994, Examples of Scientific Problems and Data Analyses in Demography, Neurophysiology, and Seismology. *Journal of Computational and Graphical Statistics,* 3(1), 1-22.

Broome, F.R. and D.B. Meixler, 1990, The TIGER Database Structure. *Cartography and Geographic Information Systems*, 17, 39-47.

Byerly, E., 1990, *State and Local Agencies Preparing Population and Housing Estimates. Current Population Reports, Series P-25, No. 1063* (Washington, DC: U.S. Department of Commerce, U.S. Bureau of the Census).

Carrat, F. and A.-J. Valleron, 1992, Epidemiologic Mapping Using the "Kriging" Method: Application to an Influenza-Like Illness Epidemic in France. *American Journal of Epidemiology,* 135, 1293-1300.

Chen, R., 1996, Exploratory Analysis as a Sequel to Suspected Increased Rate of Cancer in a Small Residential or Workplace Community. *Statistics in Medicine,* 15, 807-816.

Chen, R., N. Mantel, and M.A. Klingberg, 1984, A Study of Three Techniques for Time-Space Clustering in Hodgkin's Disease. *Statistics in Medicine,* 3, 173-184.

Clayton, D. and J. Kaldor, 1987, Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics,* 43, 671-681.

Choynowski, M., 1959, Maps Based on Probabilities. *Journal of the Royal Statistical Association,* 54, 385-388.

Clark, W.A.V. and K.L. Avery, 1976, The Effect of Data Aggregation in Statistical Analysis. *Geographical Analysis,* 8, 428-438.

Clarke, K.C., S.L. McLafferty, and B.J. Tempalski, 1996, On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Diseases,* 2(2), 85-92.

Clayton, D. and J. Kaldor, 1987, Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping *Biometrics,* 43, 671-687.

Cleek, R.K., 1979, Cancers and the Environment: The Effect of Scale. *Social Science and Medicine,* 13D, 241-247.

Cliff, A.D. and P. Haggett, 1988, *Atlas of Disease Distributions: analytic approaches to Epidemiological Data* (London: Basil Blackwell).

Cliff, A.D. and J.D. Ord, 1975, Model Building and the Analysis of Spatial Pattern in Geography (With Discussion). *Journal of the Royal Statistical Society, Series B,* 37, 297-348.

Cliff, AD. and J. D. Ord, 1981, *Spatial Processes* (London: Pion Limited).

Cressie, N., 1977, On Some Properties of the Scan Statistic on the Circle and the Line. *Journal of Applied Probability,* 14, 272-283.

Cressie, N., 1992, Smoothing Regional Maps using Empirical Bayes Predictors. *Geographical Analysis*, 24, 75-95.

Cressie, N. and T. Read, 1989, Spatial Data Analysis of Regional Counts. *Biometrical Journal,* 6, 699-719.

Cuzick, J. and R. Edwards, 1990, Spatial Clustering for Inhomogeneous Populations. *Journal of the Royal Statistical Society, Series B,* 52, 73-104.

Dean, A.G., 1993, *EPIMAP* (Atlanta: Centers for Disease Control and Prevention).

Devine, O. J., and T.A. Louis, 1994, A Constrained Empirical Bayes Estimator for Incidence Rates in Areas with Small Populations. *Statistics in Medicine,* 13, 1119-1133.

Devine, O.J., T.A. Louis, and M.E. Halloran, 1994, Empirical Bayes Methods for Stabilizing Incidence Rates Before Mapping. *Epidemiology,* 5, 622-630.

Devine, O.J., T.A., Louis, and M.E. Halloran, 1996, Identifying Areas with Elevated Disease Incidence Rates Using Bayes Estimators. *Geographical Analysis,* 28, 187-199.

Diggle, P.J., 1983, *Statistical Analysis of Spatial Point Patterns* (New York: Academic Press).

Diggle, P.J., 1991, A Point Process Modeling Approach to Raised Incidence of a Rare Phenomenon in the Vicinity of a Pre-Specified Point. *Journal of the Royal Statistical Society, Series A,* 153, 349-362.

Diggle, P.J. and A.G. Chetwynd, 1991, Second-Order Analysis of Spatial Clustering for Inhomogeneous Populations. *Biometrics,* 47, 1155-1163.

Diggle, P.J. and B.S. Rowlingson, 1994, A Conditional Approach to Point Process Modelling of Elevated Risk. *Journal of the Royal Statistical Society, Series A,* 157(3), 433-440.

Dobson, A.J., K. Kuulasmaa, E. Eberle, and J. Sherer, 1991, Confidence Intervals for Weighted Sums of Poisson Parameters. *Statistics in Medicine,* 10, 457-462.

Dorling, O. and S. Openshaw, 1992, Using Computer Animation to Visualize Space-Time Patterns. *Environment and Planning A*, 24, 639-650.

Downer, R.G., 1996, An Introduction to Smoothing Incidence Rates by Penalized Likelihood. *Statistics in Medicine,* 15, 907-917.

Duncan, G.T. and D. Lambert, 1989, The Risk of Disclosure for Microdata. *Journal of Business and Economic Statistics*, 7, 207-217.

Duncan, G.T. and R.W. Pearson, 1991, Enhancing Access to Microdata while Protecting Confidentiality: Prospects for the Future. *Statistical Science,* 6(3), 219-239.

Ederer, F., M.H. Myers, and N. Mantel, 1964, A Statistical Problem in Space and Time, Do Leukemia Cases come in Clusters? *Biometrics,* 20, 626-638.

Elliot, P., J., Cuzick, D., English, and R. Stern (Eds.), 1992, *Geographical and Environmental Epidemiology: Methods for Small Area Studies* (New York: Oxford University Press).

Ericksen, E.P., 1974, A Regression Method for Estimating Population Changes of Local Areas. *Journal of the American Statistical Association,* 69, 867-875.

Feychting, M. and A. Ahlbom, 1993, Magnetic Fields and Cancer in Children Residing near Swedish High Voltage Power Lines. *American Journal of Epidemiology,* 7, 467-481.

Fisher, P.F. and M. Langford, 1995, Modeling the Errors in Areal Interpolation between Zonal Systems by Monte Carlo Simulation.. *Environment and Planning A,* 27, 211-224.

Flowerdew, R. and M. Green, 1989, Statistical Methods for Inference Between Incompatible Zonal Systems. In Goodchild M.F. and S. Gopal (Eds.), *Accuracy of Spatial Databases* (London, UK: Taylor and Francis), 239-247.

Fotheringham, A.S. and D.W.S. Wong, 1991, The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A,* 23, 1025-1044.

Fotheringham, A.S. and F. B. Zhan, 1996, A Comparison of Three Exploratory Methods for Cluster Detection in Spatial Point Patterns. *Geographical Analysis,* 28 (3), 200-218.

Frome, E.L. and H. Checkoway, 1985, Use of Poisson Regression Models in Estimating Incidence Rates and Ratios. *American Journal of Epidemiology,* 121, 309-323.

Gatrell, A.C. and T.C. Bailey, 1996, Interactive Spatial Data Analysis in Medical Geography. *Social Science and Medicine,* 42(6), 843-855.

Gatrell, A.C. and M. Loytonen, 1996, GIS and Health Research in Europe: A Position Paper. Paper Prepared for the Helsinki Workshop, January.

Geschwind, S.A., J.A.J. Stolwijk, M. Bracken, E. Fitzgerald, A. Stark, C., Olsen, and J. Melius, 1992, Risk of Congenital Malformations Associated with Proximity to Hazardous Waste Sites. *American Journal of Epidemiology,* 135, 1197-1206.

Getis, A. and J.K. Ord, 1992, The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis 24,* 189-206.

Ghosh, M., 1993, Constrained Bayes Estimation with Applications. *Journal of the American Statistical Association,* 87, 533-540.

Ghosh, M. and J.N.K. Rao, 1994, Small Area Estimation: An Appraisal. *Statistical Science,* 9, 55-93.

Glaz, J., 1993, Approximations for the Tail Probabilities and Moments of the Scan Statistic. *Statistics in Medicine,* 12, 1845-1852.

Glick, B., 1979, The Spatial Autocorrelation of Cancer Mortality. *Social Science and Medicine,* 13D, 123-130.

Goodchild, M.F. and N.S.-N. Lam, 1980, Areal Interpolation: A Variant of the Traditional Spatial Problem. *Geoprocessing,* 1, 297-331.

Goodchild, M.F., L. Anselin, and U. Deichmann, 1993, A Framework for the Areal Interpolation of Socioeconomic Data. *Environment and Planning A,* 25, 383-397.

Goodman, M.T., C.N. Yoshizawa, and L.N. Kolonel, 1989, Incidence Trends and Ethnic Patterns for Childhood Leukemia in Hawaii: 1960-1984. *British Journal of Cancer,* 60, 93-97.

Greenberg, B. and L. Voshell, 1990, Relating Risk of Disclosure for Microdata and Geographic Area Size. In *Proceedings of the Section on Survey Research Methods* (Alexandria, VA: American Statistical Association), 450-455.

Gribb, W.J., R.J. Czerniak, and J.A. Harrington, 1990, Rural Addressing and Computer Mapping in New Mexico. *The Professional Geographer,* 42, 471-480.

Grimson, R.C., 1993, Disease Clusters, Exact Distribution of Maxima and P-Values. *Statistics in Medicine,* 12, 1773-1794.

Grimson, R.C. and N. Oden, 1996, Disease Clusters in Structured Environments. *Statistics in Medicine*, 15, 851-871.

Grimson, R.C., K.C. Wang, and P.W.C. Johnson, 1981, Searching for Hierarchical Clusters of Disease: Spatial Patterns of Sudden Infant Death Syndrome. *Social Science and Medicine,* 15D, 287-293.

Haining, R., 1990, *Spatial Data Analysis in the Social and Environmental Sciences* (New York: Cambridge University Press).

Haining, R., S.M. Wise, and M. Blake, 1994, Constructing Regions for Small Area Analysis: Material Deprivation and Colorectal Cancer. *Journal of Public Health Medicine,* 16, 429-438.

Haining, R., S. Wise, and J. Ma, 1998, Exploratory Spatial Data Analysis in a Geographic Information System Environment. *The Statistician,* 47(3), 457-469.

Hall, H.I., C.V. Lee, and W.E. Kaye, 1996, CLUSTER: A Software System for Epidemiologic Cluster Analysis. *Statistics in Medicine,* 15, 943-950.

Haybittle, J., P. Yuen, and D. Machin, 1995, Multiple Comparisons in Disease Mapping. *Statistics in Medicine,* 14, 2503-2505.

Hills, M. and F. Alexander, 1989, Statistical Methods Used in Assessing the Risk of Disease near a Source of Possible Environmental Pollution: A Review (with Discussion). *Journal of the Royal Statistical Society, Series, A,* 152, 353-384.

Hjalmars, U., M. Kulldorff, G. Gustafsson, and N. Nagarwalla, 1996, Childhood Leukemia in Sweden: Using GIS and a Spatial Scan Statistic for Cluster Detection. *Statistics in Medicine,* 15, 707-716.

Holt, D., D.G. Steel, M. Tranmer, and N. Wrigley, 1996, Aggregation and Ecological Effects in Geographically Based Data. *Geographical Analysis,* 28, 244-261.

Hoover, R., T.J. Mason, F.W. McKay, and J.F. Fraumeni, Jr., 1975, Cancer by County: New Resource for Etiologic Clues. *Science,* 189, 1005-1007.

Hryhorczuk, D.O., L.J., Frateschi, J.W. Lipscomb, and R. Zhang, 1992, Use of the Scan Statistic to Detect Temporal Clustering of Poisonings, *Clinical Toxicology,* 30, 459-465.

Huel, G., J.-F., Petiot, and P. Lazar, 1986, Algorithm for the Grouping of Contiguous Geographical Zones. *Statistics in Medicine,* 5, 171-181.

Institute of Medicine, 1994, *Health Data in the Information Age: Use, Disclosure, and Privacy* (Washington, DC: National Academy Press).

Isserman, A., 1977, The Accuracy of Population Projections for Subcounty Areas. *Journal of the American Institute of Planners,* 43, 247-259.

Isserman, A., 1984, Projection, Forecast, and Plan: On the Future of Population Forecasting. *American Planning Association Journal,* 50, 208-221.

Jacquez, G.M., 1992, C2D: Spatial Autocorrelation in Two Dimensions. BioMedware, Inc., Ann Arbor, Michigan, 1992.

Jacquez, G.M., 1994a, Cuzick and Edwards Test when Exact Locations are Unknown. *American Journal of Epidemiology,* 140, 58-64.

Jacquez, G.M. 1994b, Stat! Guidelines and Procedures for Investigating Disease Clusters. BioMedware, Ann Arbor, Michigan.

Jacquez, G.M., 1996, Disease Cluster Statistics for Imprecise Space-Time Locations. *Statistics in Medicine,* 15, 873-886.

Jacquez, G.M. and L.A.Waller, 1996, The Effect of Uncertain Locations on Disease Cluster Statistics. In *Proceedings of the Second International Symposium on Spatial Accuracy Assessment*, 259-266.

Kelsey, J.L. and P.L. Horn-Ross, 1993, Breast Cancer: Magnitude of the Problem and Descriptive Epidemiology. *Epidemiologic Reviews,* 15, 7-16.

Kennedy-Kalafatis, S., 1995, Reliability-adjusted disease maps. *Social Science and Medicine,* 41, 1273-1287.

King, P.E., 1979, Problems of Spatial Analysis in Geographical Epidemiology. *Social Science and Medicine,* 13D, 249-252.

Kingham, S.P., A.C. Gatrell, and B. Rowlingson, 1995, Testing for Clustering of Health Events Within a Geographical Information System Framework. *Environment and Planning A,* 27, 809-821.

Klauber, M.R. and J.J. Angulo, 1974, Variola Minor in Braganca Paulista County, 1956: Space-Time Interaction Among Variola Minor Cases in Two Elementary Schools. *American Journal of Epidemiology,* 99, 65-74.

Knox, E.G., 1964, The Detection of Space-Time Interactions. *Applied Statistics,* 18, 25-29.

Kulldorf, M. and N. Nagarwalla, 1995, Spatial Disease Clusters: Detection and Inference. *Statistics in Medicine,* 14, 799-810.

Lagakos, S.W., B.J. Wessen, and M. Zelen, 1986, An Analysis of Contaminated Well Water and Health Effects in Woburn, Massachusetts (with Discussion). *Journal of the American Statistical Association,* 81, 583-614.

Langford, I., 1994, Using Empirical Bayes Estimates in the Geographical Analysis of Disease Risk. *Area,* 26, 142-149.

Lawson, A.B., 1993, On the Analysis of Mortality Events Associated with a Prespecified Fixed Point. *Journal of the Royal Statistical Society, Series A,* 156, 363-377.

Le, N.D., A.J. Petkau, and R. Rosychuk, 1996, Surveillance of Clustering Near Point Sources. *Statistics in Medicine,* 15, 727-740.

Louis, T.A., 1984, Estimating a Population of Parameter Values Using Bayes and Empirical Bayes Methods. *Journal of the American Statistical Association,* 79, 393-398.

Lovett, A.A., A.C. Gatrell, J. P. Bound, P.W. Harvey, and A.R. Whelan, 1990, Congenital Malformations in the Fylde Region of Lancashire, England, 1957-1973. *Social Science and Medicine,* 30(1), 103-109.

Mantel, N., 1967, The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research,* 27, 209-220.

Mantel, N., R.J. Kryscio, and M.H. Myers, 1976, Tables and Formulas for Extended Use of the Ederer-Myers-Mantel Disease-Clustering Procedure. *American Journal of Epidemiology,* 104, 576-584.

Manton, K., M., Woodbury, E., Stallard, W., Riggan, J. Creason, and A. Pellom, 1989, Empirical Bayes Procedures for Stabilizing Maps of U.S. Cancer Mortality Rates. *Journal of the American Statistical Association,* 84, 637-650.

Marshall, R.J., 1991, A Review of Methods for the Statistical Analysis of Spatial Patterns of Disease. *Journal of the Royal Statistical Society, Series A,* 154, 421-441.

Marx, R.W., 1990, The TIGER System: Yesterday, Today and Tomorrow. *Cartography and Geographical Information Systems,* 17, 89-97.

Maskarinec, G., 1996, Investigating Increased Incidence of Events in the Islands: A Hawaii Department of Health Perspective. *Statistics in Medicine*, 15, 699-705.

McAuliffe, T.L. and A.A. Afif, 1984, Comparison of Nearest Neighbor and Other Approaches to the Detection of Space-Time Clustering. *Computational Statistics and Data Analysis,* 2, 125-142.

McCune, B. and M.J. Mefford, 1995, PC-ORD. Multivariate Analysis of Ecological Data, Version 2.0. MjM Software Design (Gleneden Beach, Oregon).

McKnight, R.H., R.J. Kryscio, J.R. Mays, and G.C. Rodgers, Jr., 1996, Spatial and Temporal Clustering of an Occupational Poisoning: The Example of Green Tobacco Sickness. *Statistics in Medicine,* 15, 747-757.

Modan, B., Z. Blumstein, O. Luxenburg, I. Novikov, and J. Shemer, 1996, A Potential Risk of Cancer in a Central Laboratory. *Statistics in Medicine,* 15, 759-763.

Moellering, H. and W. Tobler, 1972, Geographical Variances. *Geographical Analysis,* 4, 34-50.

Mollie, A. and S. Richardson, 1991, Empirical Bayes Estimates of Cancer Mortality Rates Using Spatial Models. *Statistics in Medicine,* 10, 95-112.

Moran, P.A., 1950, Notes on Continuous Stochastic Phenomena. *Biometrica,* 37, 17-23.

Morris, R.D. and R.L. Munasinghe, 1993, Aggregation of Existing Geographic Regions to Diminish Spurious Variability of Disease Rates. *Statistics in Medicine,* 12, 1915-1929.

Moulton, L.H., B. Foxman, R.A. Wolfe and F.K. Port, 1994, Potential Pitfalls in Interpreting Maps of Stabilized Rates. *Epidemiology,* 5, 297-301.

Munasinghe, R.L. and R.D. Morris, 1996, Localization of Disease Clusters Using Regional Measures of Spatial Autocorrelation. *Social Science and Medicine,* 15, 893-905.

Nagarwalla, N., 1996, A Scan Statistic with a Variable Window. *Statistics in Medicine,* 15, 845-850.

National Research Council, 1980, *Estimating Population and Income of Small Areas, Assembly of Behavioral and Social Sciences. Panel on Small-Area Estimates of Population and Income* (Washington, DC: National Academy Press).

Naus, J.I., 1965, The Distribution of the Size of the Maximum Cluster of Points on a Line. *Journal of the American Statistical Association,* 61, 532-538.

Naus, J.I., 1966, A Power Comparison of Two Tests of Non-Random Clustering. *Technometrics,* 8, 493-517.

Neutra, R., S. Swan, and T. Mack, 1992, Clusters Galore: Insights About Environmental Clusters from Probability Theory. *The Science of the Total Environment,* 127, 187-200.

Oden, N.L., 1995, Adjusting Moran's I for Population Density. *Statistics in Medicine,* 14, 17-26.

Oden, N., G. Jacquez, and R. Grimson, 1996, Realistic Power Simulations Compare Point- and Area-Based Disease Cluster Tests. *Statistics in Medicine,* 15, 783-806.

Olson, J., 1975, Autocorrelation and Visual Map Complexity. *Annals of the Association of American Geographers,* 65, 189-204.

Openshaw, S., 1984, Ecological fallacies and the Analysis of Areal Census Data. *Environment and Planning A,* 6, 17-31.

Openshaw, S., 1994, Two Exploratory Space-Time-Attribute Pattern Analysers Relevant to GIS. In Fotheringham S. and P. Rogerson (Eds.), *Spatial Analysis and GIS* (London: Taylor & Francis), 83-104.

Openshaw, S., M. Charlton, and A.W. Craft, 1988a, Searching for Leukaemia Clusters Using a Geographical Analysis Machine. *Papers of the Regional Science Association,* 64, 95-106.

Openshaw, S., M. Charlton, A.W. Craft, and J.M. Birch, 1988b, Investigation of Leukaemia Clusters by Use of a Geographical Analysis Machine. *Lancet,* 1, 272-273.

Openshaw, S., M. Charlton, C. Wymer, and A.W. Craft, 1987, A Mark 1 Analysis Machine for the Automated Analysis of Point Data Sets. *International Journal of Geographical Information Systems* 1 335-358.

Pickle, L.W., T.J. Mason, N. Howard, R. Hoover, and J.F. Fraumeni, Jr., 1987, *Atlas of US Cancer Mortality Among Whites 1950-1980* (Washington, DC: U.S. Government Printing Office), DHHS Publ. No. (NIH) 87-2900.

Pickle, L.W., T.J. Mason, N. Howard, R. Hoover, and J.F. Fraumeni, Jr., 1990, *Atlas of US Cancer Mortality Among Nonwhites 1950-1980* (Washington, DC: U.S. Government Printing Office), DHHS Publ. No. (NIH) 90-1582.

Pompe-Kirn, V. and A. Ferligoj, 1991, Solving the Problem of Small-Population-Based Areas for the Analysis of Rare Disease by Clustering with Constraints Methods. *Cancer Detection and Prevention,* 15, 77-82.

Reynolds P., D.F. Smith, E. Satariano, D.O. Nelson, L.R. Goldman, and R.R. Neutra, 1996, The Four County Study of Childhood Cancer: Clusters in Context. *Statistics in Medicine,* 15, 683-697.

Ripley, B.D., 1981, *Spatial Statistics.* (New York: John Wiley & Sons).

Rizzardi, M., M.S. Mohr, D.W. Merrill, and S. Selvin, 1993, Interfacing U.S. Census Map Files with Statistical Graphics Software: Application and Use in Epidemiology. *Statistics in Medicine,* 12, 1953-1964.

Robinson, W.S., 1950, Ecological Correlations and the Behavior of Individuals. *American Sociological Review,* 15 351-357.

Rowlingson, B.S. and P.J. Diggle, 1993, SPLANCS: Spatial Point Pattern Analysis Code in S-Plus. *Computers and Geosciences,* 19, 627-655.

Rothman, K.J., 1990, A Sobering Start for the Cluster Busters' Conference. *American Journal of Epidemiology,* 132, S6-S13.

Rushton, G. and P. Lolonis, 1996, Exploratory Spatial Analysis of Birth Defect Rates in an Urban Population. *Statistics in Medicine,* 15, 717-726.

Schlattman, P., 1996, The Computer Package DismapWin. *Statistics in Medicine,* 15, 931.

Schlattman, P., E. Dietz, and D. Bohning, 1996, Covariate Adjusted Mixture Models and Disease Mapping with the Program DismapWin. *Statistics in Medicine,* 15, 919-929.

Schneider, D., M.R. Greenberg , M.H. Donaldson, and D. Choi, 1993, Cancer Clusters: The Importance of Monitoring Multiple Geographic Scales. *Social Science and Medicine,* 37, 753-759.

Schweder, D. and E. Spjotvoll, 1982, Plots of P-values to Evaluate Many Tests Simultaneously. *Biometrika,* 69, 493-502.

Silverman, B.W., 1978, Choosing the Window Width when Estimating a Density. *Biometrika,* 65, 1-11.

Silverman, B.W., 1986, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall).

Skelton, A.G., 1996, Quadrant Analysis Software for the Detection of Spatial or Temporal Clustering. *Statistics in Medicine,* 15, 939-941.

Smith, D. and R. Neutra, 1993, Approaches to Disease Cluster Investigations in a State Health Department. *Statistics in Medicine,* 12, 1757-1762.

Sokal, R.R., N.L. Oden, B.A. Thomson, and J. Kim, 1993, Testing for Regional Differences in Means: Distinguishing Inherent from Spurious Spatial Autocorrelation by Restricted Randomization. *Geographical Analysis,* 25, 199-210.

Stone, R.A., 1988, Investigations of Excess Environmental Risks Around Putative Sources: Statistical Problems and a Proposed Test. *Statistics in Medicine,* 7, 649-660.

Strassburg, M. and R. Williams, 1995, EpiCMR User's Guide: An EpiInfo 6 Application for Importing Data and Generating Reports, Graphs, and Maps from Confidential Morbidity Reports. Los Angeles County Department of Health Services, Los Angeles, CA, University of California, Community and Organization Research Institute, Santa Barbara, CA, April.

Tomlin, W.R., 1990, *Geographic Information Systems and Cartographic Modelling* (Englewood Cliffs, NJ: Prentice Hall).

Townsend, P., P. Phillimore, and A. Beattie, 1988, *Health and Deprivation: Inequality and the North* (London: Croom Helm).

Turnbull, B.W., E.J. Iwano, W.S. Burnett, H.L. Howe and L.C. Clark, 1990, Monitoring for Clusters of Disease: Application to Leukemia Incidence in Upstate New York. *American Journal of Epidemiology,* 132, S136-S143.

Upton, G. and B. Fingleton, 1985, *Spatial Data Analysis by Example: Point Pattern and Quantitative Data, Vol I* (New York: John Wiley & Sons)

U.S. Bureau of the Census, 1992, *TIGER/Line Files, 1992, Technical Documentation* (Washington, DC: The Bureau)

U.S. Office of Management and Budget, 1994, *Report on Statistical Disclosure Limitation Methodology, Statistical Policy Working Paper 22* (Washington, DC: U.S. Office of Management and Budget).

Waldhor, T., 1996, The Spatial Autocorrelation Coefficient Moran's I under Heteroscedasticity. *Statistics in Medicine,* 15, 887-892.

Wallenstein, S., 1980, A Test for Detection of Clustering Over Time. *American Journal of Epidemiology,* 111, 367-372.

Wallenstein, S., J. Naus, and J. Glaz, 1993, Power of the Scan Statistic for Detection of Clustering. *Statistics in Medicine,* 12, 1829-1843.

Wallenstein, S. and N. Neff, 1987, An approximation for the Distribution of the Scan Statistic. *Statistics in Medicine,* 6, 197-207.

Waller, L.A., 1996, Statistical Power and Design of Focused Clustering Studies. *Statistics in Medicine,* 15, 765-782.

Waller, L.A. and A.B. Lawson, 1995, The Power of Focused Tests to Detect Disease Clustering. *Statistics in Medicine,* 14, 2291-2308.

Waller, L.A. and B.W. Turnbull, 1993, The Effect of Scale on Tests for Disease Clustering. *Statistics in Medicine,* 12, 1869-1884.

Waller, L.A., B.W. Turnbull, L.C. Clark, and P. Nasca, 1992, Chronic Disease Surveillance and Testing of Clustering of Disease and Exposure: Application to Leukemia Incidence and TCE-Contaminated Dumpsites in Upstate New York. *Environmetrics,* 3, 281-300.

Waller, L.A., B.W. Turnbull, L.C. Clark, and P. Nasca, 1994, *Spatial y to Detect Rare Disease Clusters. Case Studies in Biometry* (New York: John Wiley & Sons), 3-23.

Walter, S.D., 1992a, The Analysis of Regional Patterns in Health Data: I. Distributional Considerations. *American Journal of Epidemiology,* 136, 742-759.

Walter, S.D., 1992b, The Analysis of Regional Patterns in Health Data: II. The Power to Detect Environmental Effects. *American Journal of Epidemiology,* 136, 742-759.

Walter, S.D., 1993a, Assessing Spatial Patterns in Disease Rates. *Statistics in Medicine,* 12, 1885-1894.

Walter, S.D., 1993b, Visual and Statistical Assessment of Spatial Clustering in Mapped Data. *Statistics in Medicine,* 12, 1275-1291.

Walter, S. D. and Birnie, S.E., 1991, Mapping Mortality and Morbidity Patterns: An International Comparison. *International Journal of Epidemiology,* 20, 678-689.

Wartenberg, D., 1992, Screening for Lead Exposure Using a Geographic Information System. *Environmental Research,* 59, 310-17.

Wartenberg, D. and M. Greenberg, 1990, Detecting Disease Clusters: The Importance of Statistical Power. *American Journal of Epidemiology,* 132, S156-S166.

Wartenberg, D. and M. Greenberg, 1993, Solving the Cluster Puzzle: Clues to Follow and Pitfalls to Avoid. *Statistics in Medicine,* 12, 1763-1770.

Wartenberg, D., M. Greenberg and R. Lathrop, 1993, Identification and Characterization of Populations Living Near High-Voltage Transmission Lines: A Pilot Study. *Environmental Health Perspectives,* 101, 626-632.

Webster, R., M.A. Oliver, K.R. Muir, and J.R. Mann, 1994, Kriging the Local Risk of a Rare Disease from a Register of Diagnoses, *Geographical Analysis,* 26, 168-185.

Weinstock MA., 1981, A Generalized Scan Statistic Test for the Detection of Clusters. *International Journal of Epidemiology,* 10, 289-293.

Weiss, K.B. and D.K. Wagener, 1990a, Changing Patterns of Asthma Mortality. Identifying Target Populations at High Risk. *Journal of the American Medical Association,* 264, 1683-1687.

Weiss, K.B. and D.K. Wagener, 1990b, Geographic Variations in US Asthma Mortality: Small-Area Analyses of excess Mortality, 1981-1985. *American Journal of Epidemiology,* 132, S107-S115.

Wise, S.M., R.P. Haining, and J. Ma., 1997, Regionalization Tools for the Exploratory Spatial Analysis of Health Data. In Fischer M. and A. Getis (Eds), *Recent Developments in Spatial Analysis: Spatial Statistics, Behavioural Modelling and Neuro-computing* (Berlin: Springer-Verlag), 83-100.

Wojdyla, D., L. Poletto, C. Cuesta, C. Badler, and M.E. Passamonti, 1996, Cluster Analysis with Constraints: Its Use with Breast Cancer Mortality Rates in Argentina. *Statistics in Medicine,* 15, 741-746.