2nd Edition

GIS and Public Health

Ellen K. Cromley
Sara L. McLafferty

ebook

GUILFORD PUBLICATIONS

GIS and Public Health

GIS and Public Health

SECOND EDITION

Ellen K. Cromley
Sara L. McLafferty

THE GUILFORD PRESS
New York London

A Division of Guilford Publications, Inc.
72 Spring Street, New York, NY 10012
www.guilford.com

Printed in the United States of America

This book is printed on acid-free paper.

Last digit is print number: 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Cromley, Ellen K.
GIS and public health / by Ellen K. Cromley and Sara L. McLafferty. — 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-60918-750-7 (hardback)
1. Geographic information systems. 2. Public health—Data processing.
I. McLafferty, Sara, 1951– II. Title.
RA566.C764 2012
610.285—dc23
2012019585

To Robert, Gordon, and Ed
and
Avijit, Smita, and Priya

Acknowledgments

In 2005, The Guilford Press invited us to update the first edition of *GIS and Public Health*. We were grateful for the success of the book and recognized how much significant change was occurring in the use of geographic information systems (GIS) in the sphere of public health. We set out to capture these new developments by building on the framework we adopted in the first edition. The second edition expands the scope of the work that motivated us to write about GIS and public health in the first place, and we remain grateful to all of the individuals we previously acknowledged.

We are grateful to Kristal Hawkins, editor at The Guilford Press, for her encouragement and support for this project and also thank everyone at Guilford who helped us see this project through to completion. In particular, we thank Guilford for publishing the online supplement, which is an important feature of the second edition. Jared Butler has our thanks for testing the exercises in that supplement and suggesting needed corrections and improvements. We also thank the reviewers for their careful reading of the first draft of the manuscript and their constructive and insightful comments and suggestions.

An important consideration that convinced us to attempt a second edition was the prospect of working together again. This edition, like the first, has been a true collaboration. We are also grateful to all of the people we have worked with for providing us new opportunities in the field to learn how we can work together to improve public health.

I (Ellen K. Cromley) especially thank my advisors Allen Fonoroff and Howard Biel (Case Western Reserve University), Kevin Cox (The Ohio State University), and Gary Shannon (University of Kentucky) for everything they taught me and for their contributions to their fields. I also acknowledge Bob Brems, Zanesville/Muskingum County Health Department; Julia Dickson, Medical College of Wisconsin; Bill Elwood, OppNet, National Institutes of Health; Dr. Carol Horowitz, Mount Sinai School of Medicine, and colleagues at the Center to IMPACT Diabetes in East Harlem; Blair Johnson, Center for Health, Intervention, and Prevention (CHIP), University of Connecticut; Stanley Lemeshow, Mary Ellen Wewers, and faculty and students at the Summer Program in the College of Public Health at The Ohio State University; Patricia Mabry, National Institutes of Health; Wendy Nelson and Pat Francis, URISA's GIS in Public Health Conference; Rachel Pruchno, University of Medicine and Dentistry of New Jersey; Lorraine Reitzel, Lorna Haughton McNeill, and Dave Wetter, Department of Health Disparities Research, MD Anderson Cancer Center; Timothy Sahr, Health Policy Institute of Ohio; Jean Schensul, Marlene Berg, and Emil Coman, The Institute for Community Research, and Kamla Gupta and S. K. Singh, International Institute for Population Sciences, along with members of the ASHRA project team in Mumbai; Philip Troped and Heather Whitcomb, Department of Health and Kinesiology, Purdue University; Francine Laden, Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School; Steve Melly, Exposure, Epidemiology and Risk Program, Harvard School of Public Health; Robin Puett, Department of Environmental Health Sciences, University of South Carolina; Joan Twiggs and Darcé Costello, Connecticut Department of Children and Families; Margaret Weeks, The Institute for Community Research; and the faculty and staff of the Department of Community Medicine and Health Care, University of Connecticut School of Medicine. I also thank Emily Marble, at The Institute for Community Research in Hartford, Connecticut, for assistance in research administration. Rich Mrozinski, Department of Geography, University of Connecticut, and Howard Sternberg, Connecticut Department of Environmental Protection, deserve special thanks for their great expertise in GIS and their willingness to share it.

I (Sara L. McLafferty) extend thanks to colleagues and current and former students at the University of Illinois, Hunter College, and elsewhere for their many contributions to this book. Many years ago, my PhD advisor, Gerard Rushton of the University of Iowa, challenged and encouraged me to explore the links between geographical analysis and social and health inequalities. A renowned leader in GIS and public health, Gerry's insights, intellect, and inspiration provided the academic foundation that made this book possible and is reflected in so many ways in my contributions to the book. I also consider myself fortunate to have worked with many colleagues in investigating diverse health issues in Illinois and in the New York metropolitan region. Special thanks to Fahui Wang of Louisiana State University; Vince Freeman of the University of Illinois at Chicago; Keith Clarke of the University of California at Santa Barbara; Victor Goldsmith of Pace University; Nick Freudenberg of Hunter College; Roger Grimson of SUNY Stony Brook; the members of the West Islip Breast Cancer Coalition; Dr. Christina Hoven of Columbia University; and Dr. James Childs of the Centers for Disease Control and Prevention. The New York City Department of Health generously gave permission to use individual-level rat bite data to prepare several figures in this book. I am also extraordinarily grateful to the many talented graduate students I have worked with on issues related to

GIS and health, including Lan Luo, Alisa Shockley, Miriam Cope, Trevor Fuller, Ranjana Chakrabarti, Travis Leonard, Jong-Hyung Lee, Carmen Tedesco, John Pan, Sue Grady, Doug Williamson, Barbara Tempalski, Henry Sirotin, Cheryl Weisner, Linda Timander, Brett Gilman, Delene Pratt, Julie Kranick, Sonia Tatlock, Colin Reilly, and Chris Hanson-Sanchez. This book benefited in innumerable ways from their insights, expertise, collaboration, and camaraderie as we probed the frontiers of GIS and public health.

Preface to the Second Edition

When we wrote the first edition of this book, the field of geographic information systems (GIS) and public health was in its infancy, a "new" field that was just beginning to attract attention among health researchers and policymakers. Hearkening back to that time, we approached the task of preparing the second edition as a relatively straightforward process of tweaking and updating. In short order, we were overwhelmed by the size of the task at hand. In the past 10 years, the field of GIS and public health has flourished to the point where literally hundreds of articles appear in the research literature each year. From infectious diseases to cancer to obesity to health care, researchers are embracing GIS in their efforts to understand health concerns and direct interventions to improve public health and reduce health disparities. The rapid expansion of the field is also reflected in new journals and conferences that facilitate interaction among researchers and practitioners.

In preparing this second edition, we have tried to convey the amazing breadth, diversity, and dynamism of these health–GIS applications without losing sight of basic concepts and earlier work that laid the foundations for more widespread adoption of GIS. Chapters 1 through 10 have been substantially revised, expanded, and updated to reflect developments in the research literature. A new chapter on health disparities (Chapter 11) considers neighborhood influences on health and the methods used to investigate contextual effects. The final chapter of the book (Chapter 12) addresses the institutional context of GIS by focusing on public participation GIS, a topic that we feel is of great importance in promoting community involvement in efforts to improve public health. We have also sought to expand the book's geographic scope beyond the United States to comprise research developments and applications throughout the world, reflected in the large section of references. In response to requests from readers of the first edition, we have prepared a series of GIS laboratory exercises with data to accompany the second edition. These exercises are available as an online supplement (at *www.guilford.com/p/cromley*) to the book.

The past decade has also witnessed major advances in geographic information science that are represented in the second edition. Especially noteworthy are the emergence of Internet-based geovisualization and mapping systems; the enormous increase in availability of georeferenced data from cell phones and other GPS-enabled devices; developments in spatial analysis methods that emphasize "local" patterns and processes; and the growing use of GIS in promoting community participation. All chapters have been revised to discuss how these developments in GIScience can contribute to public health research and practice.

Although extensively revised and updated, this book continues to reflect our belief that understanding core geographic concepts like space, place, location, and distance, and core principles related to spatial data, mapping, and spatial analysis, is essential in applying GIS to public health issues. The book retains many sections dealing with these topics, illustrated with maps, diagrams, and real-world applications, and with material written in a way that we hope is accessible to the diverse audiences interested in public health and GIS.

Contents

List of Figures

List of Tables

Introduction

Geographic information systems (GIS) continue to transform the way we describe and study the earth. Throughout history, geographers have attempted to understand the surface of the earth as the living environment of human populations and the forces of change that alter the earth's environments. The environment affects our health and well-being, and we, through our activities, reshape the environment. GIS provide a digital lens for exploring the dynamic connections between people, their health and well-being, and changing physical and social environments.

This second edition is an updated introduction to the use of GIS in analyzing and addressing public health problems. GIS are computer-based systems for integrating and analyzing spatial data. Our book considers how GIS can be used to map and analyze the geographical distributions of populations at risk, health outcomes, and risk factors, to explore associations between risk factors and health outcomes, and to address health problems. Targeting public health interventions to populations and places with greatest need is an essential and effective strategy for improving population health, and GIS are essential tools in these efforts.

The book is written for geographers, public health practitioners, epidemiologists, and community members interested in applying GIS to the study of human health problems. The main question we seek to answer for the reader is "what do I need to know about GIS for public health?" Our answer is that to use GIS to establish a geographic foundation for understanding and improving public health, we need to know about GIS data and systems, the methods for analyzing GIS data, and how and for whom GIS are used.

Geographic Foundations for Public Health

At its most basic level, a geographic foundation for public health looks at the question "Where?" Where do people live? Where are the agents of disease?

Where can we intervene to eliminate risks or to improve health services delivery? Both people and the agents that cause disease in humans are dispersed, often unevenly, across communities and regions. The processes that bring people into contact with disease agents and that impact their access to social and material resources are also geographically variable.

Health is not just the absence of disease, but a state of physical, social, and emotional well-being. Because people are affected by their environments, health has the environment of the person as its geographical context. This environment is connected to natural, social, and economic processes that operate on local, regional, and global scales. How people behave contributes to their health status, but we cannot divorce behavior from the environmental and social contexts in which it occurs. Not all of the factors that affect our well-being are under the immediate direct control of the individual. The environment of the person is one starting point for public health GIS.

Populations at Risk

Geographic information systems are being used in public health studies to model where people live and the environments they experience throughout their lives. GIS make it possible to view residential distributions in great detail (Figure I.1). Because of the economic and social processes that structure residential development, age, sex, and race–ethnicity of the population are usually not uniform throughout the region of settlement. Instead, different neighborhoods or communities often have different demographic characteristics, and GIS make it possible to view these differences in detail.

The distribution of population by residence is, perhaps, the most frequently considered geographical distribution in public health and epidemiology. Residence, however, is only one activity site in the environment of the person, albeit an important one. The ***activity space*** is the area where a person spends time (Golledge & Stimson, 1997). It comprises a home base or residence, other activity sites like workplaces, schools, stores, restaurants, and recreational areas that are regularly visited, and the pathways traveled to and from the home and other activity sites. The size and shape of an individual's activity space will vary depending on the activities the person is obligated or chooses to perform, the modes of transportation available, and the geographical locations where activities may take place. Individuals sharing the same home base may have very different travel and activity patterns and, as a result, different activity spaces. The activity space is important because it represents the zone where the individual can be exposed to risks—or resources. In addition to modeling population distribution by residence, GIS are being used to analyze travel diaries (McCormack, 1999; Kwan, 2000) and represent activity spaces (Figure I.2).

Migration is a process that results in the permanent or semipermanent relocation of the home (Golledge & Stimson, 1997). Although most people who move stay within the same community and relocate only a short distance from their previous homes, longer distance moves are possible. These moves result in a

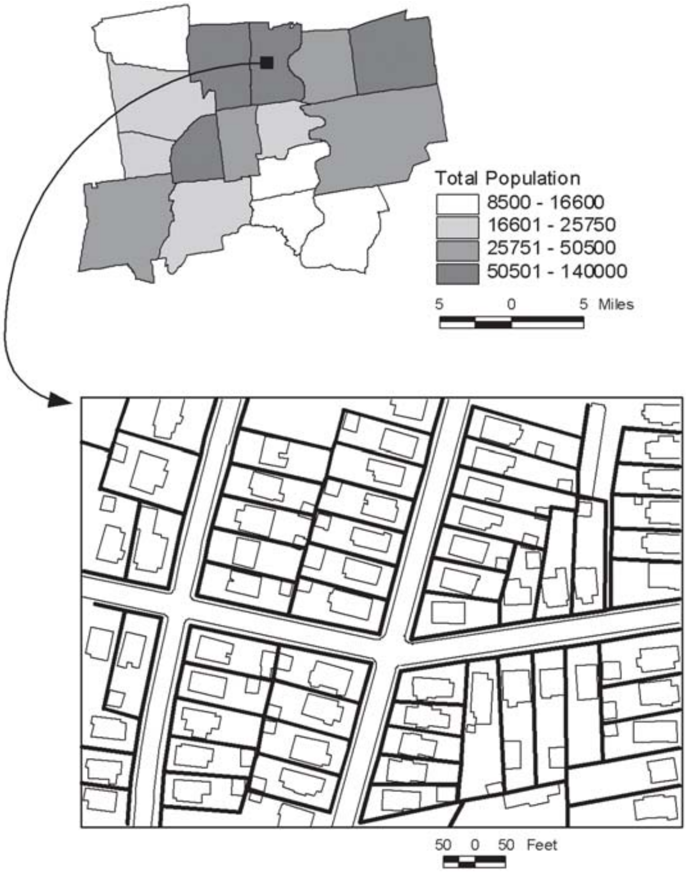


FIGURE I.1. GIS can be used to make conventional maps of population distribution, but they can also display maps showing the locations of buildings where people live.

complete displacement of the individual's activity space, exposing the person to a new set of risks. Furthermore, migration rates are rarely similar across all social and demographic groups. Migration complicates the study of disease when the time between exposure and onset is long, and it is an important process affecting health disparities.

The distributions of population and changes in population due to natural increase and aging are relatively easy to model because of the availability of census and vital statistics data. Migration flows and changes in population due to migration are more difficult to study in the United States because we do not normally maintain detailed residential histories for individuals. Nevertheless, some health records contain information on place of birth and place of residence at the time the health event is registered. Detailed surveys are also conducted to

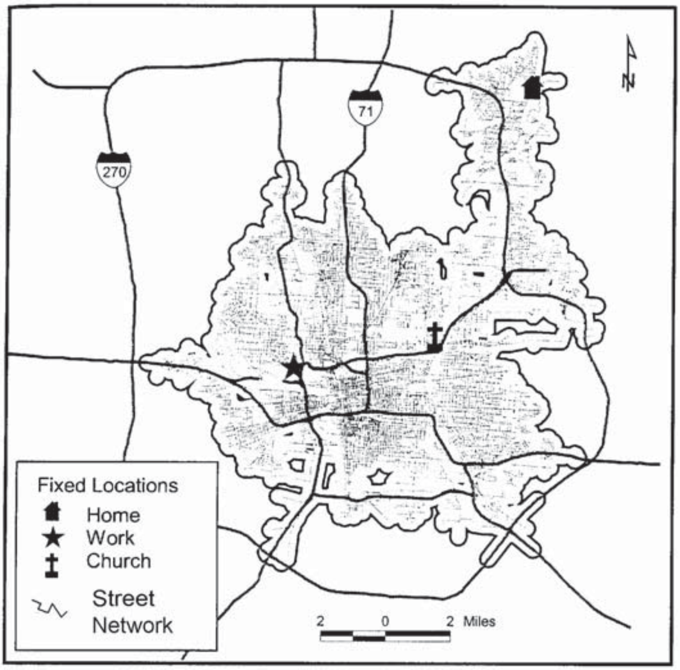


FIGURE I.2. The potential area where an individual could travel on the basis of the person's home location, workplace, and church. From: Gender and individual access to urban opportunities: A study using space-time measures, M. P. Kwan, *The Professional Geographer*, 1999. Association of American Geographers, reprinted by permission of the publisher.

re-create residential histories. These data can be used to explore differences in health outcomes in the context of migration (Greenberg & Schneider, 1992) and migration for medical care (Davis & Stapleton, 1991). In places where detailed migration data exist, the use of GIS has been proposed to improve estimation of populations at risk (Kohli et al., 1997). The use of GIS in public health enables us to describe more accurately the environment of the person and its temporal and spatial complexity.

Health Outcomes

Although the health event is the outcome of a complex process exposing vulnerable human populations to risk factors, many public health investigations and epidemiological studies start with outcomes. In many sources of data on mortality (death) and morbidity (illness), residential location at the time of death or diagnosis is reported. This makes it possible to map cases and rates and to search for clusters of health events.

Incidence and prevalence of disease vary geographically. ***Incidence*** is the number of new cases of the disease or health event observed within a specified period of time. ***Prevalence*** is the number of existing cases of disease at a particular point in time. Prevalence is related to incidence but is also influenced by the duration of the disease. A disease process that results in death shortly after onset would have a higher incidence relative to prevalence in a community. Pneumococcal disease, which causes pneumonia, bacteremia, and meningitis, can result in death within 10 days and the case-fatality rate among the elderly may be as high as 60%. A chronic incurable disease like diabetes resulting in death after many years might have higher prevalence than incidence.

Disease mapping has made contributions to public health and epidemiology for centuries (Gilbert, 1958; Shannon, 1981). As many of the examples in this book illustrate, GIS make it easier to explore and analyze large databases of health events at a high level of spatial disaggregation and to link data from surveillance systems to other information about the environment, including information on the distribution of risk factors (Figure I.3). These activities have also

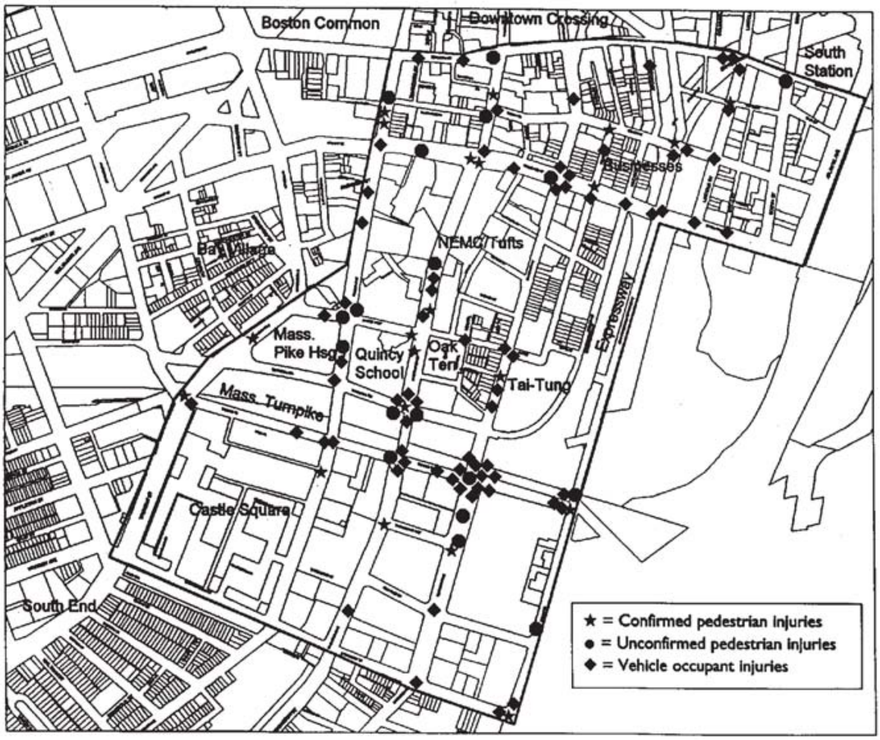


FIGURE I.3. Locations of confirmed pedestrian injuries in the context of reported total traffic injuries in a neighborhood in Boston. From Brugge, Leong, and Lai (1999). Reprinted by permission of Association of Schools of Public Health.

driven research on methods to protect the confidentiality of data on the health of individuals (Boulos, Curtis, & AbdelMalick, 2009).

A common practice in mapping incidence and prevalence has been to calculate disease rates for political or administrative units like towns or census tracts, primarily because population and health outcome data are often reported for these areas. A problem with this approach is that the boundaries of these units often arbitrarily partition the underlying distribution of population or cases, or both. GIS make it possible to overlay the distributions of cases and populations at risk and to display multiple views of the distribution of health outcomes.

GIS are also supporting analyses to search for disease clusters using new methods that do not rely on aggregated data. Areas of high and low incidence can be identified by searching around individual cases to find areas that have high numbers of cases relative to the local population. GIS have been important in the shift from global to local statistical methods for analyzing clusters of disease and health disparities.

Risk Factors

Like populations at risk, the risk factors for disease are also not usually concentrated at a single point. Contaminants and biological agents of disease are present in our ecosystem. GIS have proven to be powerful tools for modeling environmental conditions across the full ranges of geographical scales, from local to global, and related technologies are aiding in exposure assessment at the individual level.

Many GIS applications model temporal and spatial patterns of hazards to examine their associations with environmental health problems and infectious diseases (Figure I.4). These models are increasingly sophisticated, and the scale and quality of the spatially referenced data they incorporate continue to improve. Public health GIS applications may model the spread of contaminants and vector and host habitats.

The ability to model the distribution of known or potential risk factors for health problems is important for public health intervention activities. GIS are being used at the community level to notify people living in neighborhoods where hazards have been identified so that they can take appropriate action to prevent health problems from occurring. Although these analyses can be performed without reference to health outcomes, studies of the geographical patterns of risk factors can also be used to investigate the causes of disease.

Associations between Risk Factors and Outcomes

Epidemiological studies involve comparing the incidence rate of disease observed in a study group at risk against the incidence rate of disease in a comparison group. A number of measures are used to compare the rate of disease observed in the exposed group to the rate of disease observed in the comparison group (Woodward, 2004). The ***relative risk*** or ***risk ratio*** is the ratio of the risk

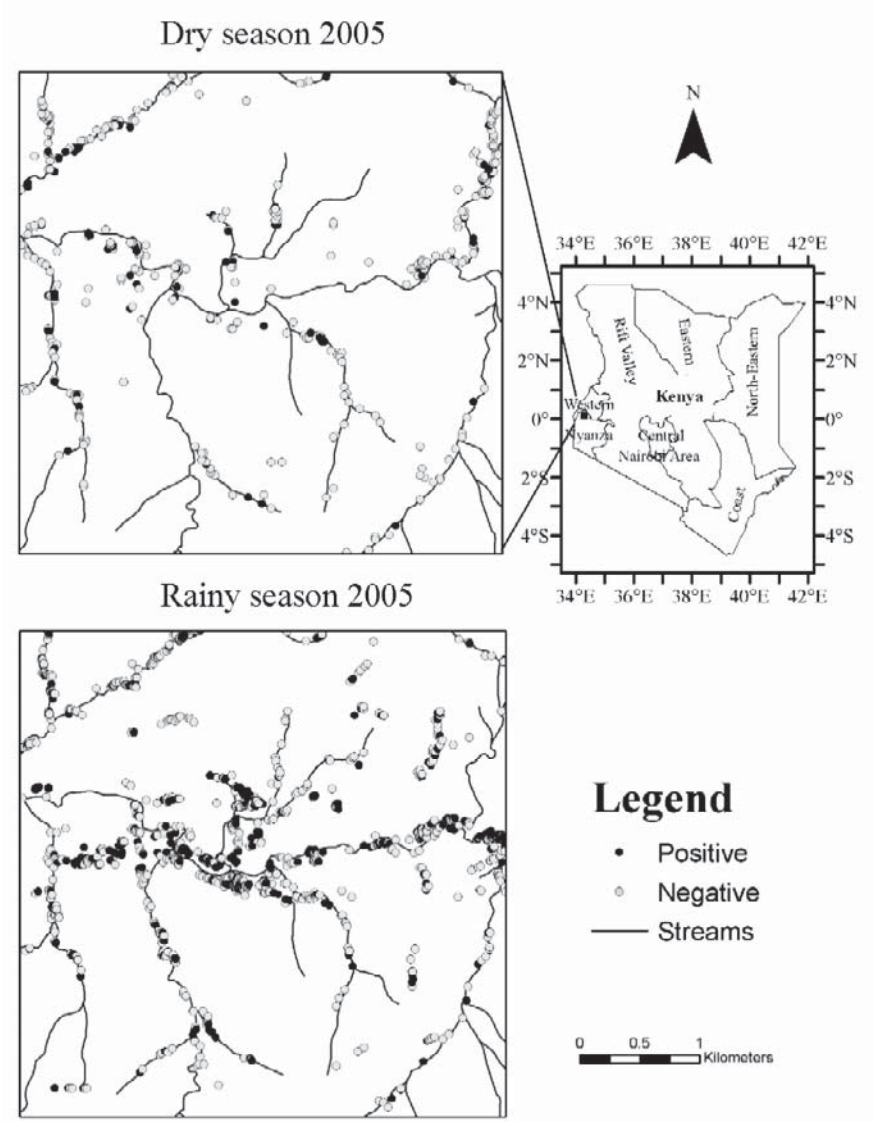


FIGURE I.4. Dry and wet season habitat areas for *Anopheles gambiae* larva in the Western Kenyan highlands based on identification of all identifiable aquatic habitats in the region. Each habitat was dipped up to 20 times or as many times as the volume of water would permit and the larval occurrence and species in the water samples were examined. Sites were coded as positive or negative depending on the presence of the *Anopheles gambiae* larva, the larva of the primary malaria vector species in the region. From Li, Bian, Yakob, Zhou, and Yan (2009). Originally published by BioMed Central in the *International Journal of Health Geographics*, Open Access.

of disease among those who have the risk factor to the risk of disease for those without the risk factor (Table I.1). ***Risk*** measures the number of times the health event occurs relative to the total number of people in the study group. A second measure, the ***odds ratio***, is also used by epidemiologists. The ***odds*** measures the number of times the health event occurs relative to the total number of times it does not occur. The numerator of the ***odds ratio*** is the ratio of health events/no health events (odds) in the exposed population. The denominator of the odds ratio is the ratio of health events/no events (odds) in the unexposed population. Standardized event ratios such as the ***standardized mortality ratio*** and the ***standardized incidence ratio*** are statistics frequently used for comparing mortality or incidence observed in exposed persons to mortality or incidence previously observed in a standard population (Fleiss, 1981; Kelsey, Whittemore, Evans, & Thompson, 1996; Selvin, 2004).

Although these measures are well established in epidemiological research, it is not always easy to determine meaningful numerators and denominators to calculate them. GIS are making it possible to explore some of the most important methodological issues in applying these measures to identify significant patterns of disease. Epidemiological surveys and case–control studies involve drawing samples from the population. Assumptions are made about the probabilities of inclusion in the sample. Because populations or population subgroups are not distributed evenly across communities, a random sample of all people with the demographic and health characteristics of interest will not be a random sample of all places (Goodchild, 1984). GIS can provide the means for exploring distributions of populations and health events to develop spatially stratified sampling techniques for conducting epidemiological surveys and for selecting cases and controls.

In order to obtain enough cases to ensure statistical power, analysts sometimes increase the study population, usually by increasing the geographic area of analysis. This approach can be misleading when the underlying geography of risk factors, exposure, and health outcomes are ignored. Enlarging the study area introduces other geographical differences than the health outcome in question, and it becomes difficult to tell whether there is an inherent difference between groups or between areas.

These problems are exacerbated by controlling for confounding factors through techniques like age standardization. A ***confounding factor*** is a variable that is causally related to the health problem under study or is a proxy for an unknown causal variable but is not a consequence of the exposure of interest. Age is a confounding factor commonly controlled for in epidemiological studies because "age almost always strongly influences disease risk" (Selvin, 2004). To compare disease incidence rates for two groups—say, white women and African American women living in a region—age-standardized rates would be calculated to adjust for differences in age between the two groups. As noted, the argument for this is that, since disease risk is higher in particular age groups, we would expect more cases of disease in a population that contains more people in those age groups.

TABLE I.1. Associations between Risk Factors and Outcomes for Disease Incidence in a Population

	Disease status		
Risk factor	Disease	No disease	Total
Exposed	a	b	$a + b$
Not exposed	c	d	$c + d$
Total	$a + c$	$b + d$	$n = a + b + c + d$

Risk = $(a + c)/(a + b + c + d)$

Exposure-specific risk for those with risk factor = $a/(a + b)$

Exposure-specific risk for those without risk factor = $c/(c + d)$

$$\text{Relative risk} = \frac{a/(a + b)}{c/(c + d)} = \frac{a(c + d)}{c(a + b)}$$

$$\text{Odds ratio for those with risk factor compared to those without} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

Example for Lyme disease among people who live near wooded areas (exposed) and people who live in towns (not exposed)

	Disease status		
Risk factor	Lyme disease	No disease	Total
Exposed	323	26,677	27,000
Not exposed	38	32,962	33,000
Total	361	59,639	60,000

Risk = 361/60000 = 0.006, or 6 per 1,000

Exposure-specific risk for those with risk factor = 323/27000 = 0.01196

Exposure-specific risk for those without risk factor = 38/33000 = 0.00115

$$\text{Relative risk} = \frac{0.01196}{0.00115} = \frac{10659000}{1026000} = 10.4$$

$$\text{Odds ratio} = \frac{0.01211}{0.00115} = \frac{10646726}{1013726} = 10.5$$

Note. Adapted from: Epidemiology: Study design and data analysis by Mark Woodward. Copyright 2004 by Taylor & Francis Group LLC – Books. Adapted with permission.

If, however, the cause of the disease is believed to be environmental, then we would expect disease risk to be higher in those geographic areas where environmental risk is higher. This would be particularly important in comparing rates for white and African American women if white and African American women of the same age were residentially concentrated in different areas within the larger community. Such areas would likely be associated with different levels of exposure to disease risk factors that are unevenly distributed in the environment.

An important criticism of adjustment of rates is "that if the specific rates vary in different ways across the various strata, then no single method of standardization will indicate that these differences exist. Standardization will, on the contrary, tend to mask these differences" (Fleiss, 1981, p. 239). An empirical study of age-specific death rates for males in the United States revealed that, up to age 40, rates in metropolitan counties were lower than rates in nonmetropolitan counties (Kitagawa, 1966). After age 40, the reverse was true. A single, summary comparison would fail to reveal this geographical pattern.

The spatial data-handling capabilities of GIS make it possible to identify exposed and unexposed groups and to explore geographical variations in health outcomes between those groups. In addition, spatial statistical methods are helping us to identify and account for spatial autocorrelation—spatial dependency or correlation among values of a variable in geographic space—in models of health outcomes and to investigate spatial variation in the relationships between factors contributing to disease and health outcomes. These methods are supporting the development of more meaningful comparisons of health outcomes across groups *and* areas.

Health Interventions

The geography of health services has been called "the *sine qua non* of medical geography" (Hunter, 1974). Setting aside the fact that health services utilization is one of the main, if biased, sources of information we have about health and morbidity in the population, what is the point of studying patterns of environmental contamination or uncovering the causes of disease if we are not willing to go the next step to intervene or to support the education, enforcement, environmental modification, and medical care intervention efforts of those committed to advancing human health? As long as our activities occur in time and space, knowing how patterns of health, disease, and health services characterize regions will be essential to this effort.

GIS are reviving interest in the location of health care services and the development of geographically based public health interventions to improve population health. From planning vaccination campaigns to remediating environmental hazards to providing social, health, and education services for high-risk populations, promoting public health involves targeting interventions to the places where they are most needed and where effectiveness will be maximized. Knowing how the characteristics of places intersect with the health of local populations is critically important. As illustrated in many of the chapters in this book, the mapping, spatial analysis, and data management capabilities of GIS are essential to these efforts.

The role of GIS in health services delivery and planning, however, is potentially broader than planning health services and public health interventions. A geographic foundation for public health also means considering the locational impacts of health policy. During the debate over health care reform at the national level in the 1990s in the United States, managed competition was offered as a policy for reorganizing medical care delivery. In "The Marketplace in Health Care Reform: The Demographic Limitations of Managed Competition" (Kronick, Goodman, Wennberg, & Wagner, 1993), some very straightforward concepts of economic geography were used to answer a simple question: Where are the places in the United States where managed competition as a market-driven approach to health care reform can work? Estimates of the minimum population required to support several independent competing provider groups were based on the extent to which the competing groups were independent, the number of organizations needed to provide competition, the threshold populations for services, and the geographic boundaries of health service markets. The results indicated that reform of the U.S. health care system by managed competition would be feasible only in major metropolitan areas (Figure I.5). More than one third of the population lived in places where managed competition could not be supported. With passage of the Patient Protection and Affordable Care Act of

FIGURE I.5. Health markets with populations greater than or equal to the threshold needed to support managed competition. From R. Kronick, D. C. Goodman, J. Wennberg, and E. Wagner, Special report: The marketplace in health care reform: The demographic limitations of managed competition, *The New England Journal of Medicine*, 328(2), 148–152.

2010, the United States is making the transition to a national health care system requiring people to have health insurance (Kline & Walthall, 2010). The number, capacities, and locations of medical care providers will be an important factor in the success of this system (Long & Stockley, 2010; MacDowell, Glasser, Fitts, Nielsen, & Hunsaker, 2010).

Efforts to improve health care systems illustrate the tensions between top-down versus bottom-up approaches to public health. In a top-down approach, priorities are identified at the national or state level based on aggregated data. Because of geographical variations in health problems or access to care, problems that are important at the state level may not be equally present in every community, and health care interventions that might work in some places may not work equally well in others. GIS are providing citizens at the local level with information to identify health problems of local concern, even if these are not the highest priority on a state or national level, and to advocate for public health policies. And they are creating an opportunity for policymakers at the national level to view and analyze health problems and policies in their full complexity.

When we use GIS to understand patterns of ill health and plan public health interventions, our efforts are rooted in place and space. Geographers have written about the concepts of space and place as overarching themes in the geography of health (Kearns, 1993). ***Space*** refers to position or location. It describes the geographical distributions of constraints and opportunities that influence and result from human activities and interactions. In contrast, ***place*** is a relational concept that addresses the human meanings and experiences associated with particular locations. Social relations and the physical sites of everyday life are intertwined in the concept of place.

By definition, GIS are concerned with space. The geographical coordinates that link diverse data layers together in a GIS identify location in space. However, GIS are also rooted in place. Each GIS is used in a particular context that is a composite of social, political, and historical conditions and trends. When people use a GIS, their "sense of place" defines what questions to ask and molds their interpretation of results. While our book emphasizes the spatial dimensions of GIS in public health, these cannot be divorced from the place contexts in which GIS are used.

The geographic foundations for public health we have briefly summarized have been explored in greater depth through the wider application of GIS technology in public health. People adopting GIS technology for public health need to understand GIS materials, GIS methods, and the institutional context of GIS in public health practice that affects the kind and quality of GIS data available. Our book is organized to foster this understanding.

Organization and Scope

The overall organization of the second edition is similar to the outline of the first book. The literature on health applications of GIS has grown considerably over the last decade, and individual chapters reflect these developments. A new chapter on health disparities considers neighborhood influences on health and the methods used to investigate contextual influences. The final chapter is devoted to the role of public participation in health applications of GIS.

The first section of the book focuses on GIS materials, including spatial databases and the generic hardware and software needed to manage them. Chapter 1 provides an overview of geographic information systems, their development, and their main functions. Chapters 2 and 3 describe important attributes of geographic databases and the foundation and other databases most commonly used in public health GIS.

The emphasis on geographic data is necessary. Many public health analysts maintain or use large databases of health events or health services. Most of these databases have some geographical information like a street address, a town name, or a county. Often, however, these databases are not structured in a way that makes it possible to map and analyze geographical relationships. GIS incorporate the foundation databases that make the mapping of health events and health services possible.

GIS implementation requires data, and adopting GIS for public health analysis therefore requires a commitment to acquiring, managing, and maintaining large spatial databases in digital form. Technological advances in computing and the growth of the Internet have made foundation databases widely accessible to public health organizations, provided they are willing to acquire the hardware and software necessary to develop applications or to access them over the Internet.

The second section of the book considers GIS methods for mapping and analyzing spatial data on population, health events, risk factors, and health services. Chapter 4 provides an overview of the mapping process, the different approaches available for mapping and querying health data, and the impact of developments in computer-assisted and web-based cartography on the kinds of maps that are being developed and published. Chapter 5 reviews the range of methods available for identifying areas of high and low incidence of disease, including spatial clustering methods. The uses of GIS in analyzing risk factors and health outcomes for environmental health problems, communicable, and vector-borne diseases are discussed in Chapters 6 through 8.

GIS implementation implies a commitment to spatial analysis methods as part of the research approach. Such methods may not always be emphasized in traditional public health curricula or research. Our book provides an introduction to a range of techniques used in spatial data analysis and points the reader to more detailed discussions of these techniques. Some of the studies we cite as examples use GIS in ways that can be criticized, often as a result of data limitations, but they have been included because they contribute to our understanding of GIS and public health in other ways.

The final section of the book deals with the institutional context of public health GIS, particularly as it affects public health policy and intervention efforts. Chapters 9 and 10 discuss how GIS are used to evaluate accessibility to health services and the geographical aspects of health care delivery. Chapter 11 looks at the question of health disparities and neighborhood effects on health, as well as the effects of neighborhood change and migration on health disparities. The final chapter, Chapter 12, focuses on public participation in health GIS applications.

GIS are being used to study public health issues around the world, and we have attempted in this second edition to broaden our coverage beyond the United States. Chapter 3, dealing with foundation data, is still focused mainly on the United States. The widespread public availability in the United States of digital spatial databases creates a foundation for GIS-based analyses of health problems that does not exist in many other countries. We are most familiar with these data, but we have provided references to selected sources of similar data in other countries.

Although the federal government has played a major role in developing these foundation databases and others of relevance to public health GIS in the United States, state and local governments have the main responsibility for health surveillance, public health intervention, and licensing and regulation of health providers and health services. As a result, public health GIS applications based in any region of the country are highly dependent for their success on the level of development of health and other GIS databases and systems in that region and state and local policies governing use and distribution of data. Other countries may have different policies governing the development and use of foundation databases. This is an important reason why our book does not attempt to tell readers how to use a particular GIS software package with their own data. We offer additional information on using GIS in an online supplement to the book.

GIS and Public Health

Everything that occurs on the earth that can be spatially referenced can be represented in a GIS. That is, we can use GIS to make maps of almost anything. Because of their visual power, GIS maps can become metaphors for the social and environmental conditions that are "contained" in geographic space, and we are often tempted to look for spatial solutions to problems that arise from these conditions. Although health problems have a spatial dimension and many can be addressed by changing locational relationships—by closing down a polluting manufacturing plant or opening a clinic, for example—GIS may not be an aid to analysis for other problems.

During the last decade, the use of GIS in public health has become firmly established. As new technologies continue to transform our ability to gather, analyze, and map health data, new types of GIS-based analyses in public health may emerge. In the broadest sense, GIS analyses bring to light places and populations for which ideas about how improvements in health can occur are—or are not—being adopted. GIS, as a means of exploring health problems and finding ways to address them, has taken its place in the conceptual and methodological foundations of public health.

CHAPTER 1

Geographic Information Systems

Geographic information systems (GIS) are computer-based systems for the integration and analysis of geographic data. ***Geographic data*** are spatial data that "result from observation and measurement of earth phenomena" referenced to their locations on the earth's surface (Tomlinson, 1987, p. 203). Whenever public health professionals or epidemiologists use disease registries with address information, consider the locations of toxic waste disposal sites, or look at air quality and water quality reports from monitoring stations, they are working with geographic data.

This chapter considers GIS as an "enabling" technology, applicable to the integration and analysis of many different types of spatial data—not just health data—by people in different organizational settings asking very different questions. One of the most striking features of GIS is their broad applicability. The first two sections of this chapter offer a definition of GIS and describe the major functions of GIS software: spatial database management, visualization and mapping, and spatial analysis. The last sections of the chapter trace the development of GIS and GIS applications in public health, including the role of GIS in distributed geographic information services on the Internet and wireless networks.

Definitions of GIS

In part because GIS is an enabling technology, a consensus definition of GIS has been difficult to achieve. The acronym has several usages: "as a technology, as a research field, and as a community" (Goodchild, 1995b, p. 35). In the 1990s, the term "GI Science" was coined to distinguish the research field from the technology of geographic information systems (Longley, Goodchild, Maguire, & Rhind, 2001, p. 438). Developments in GIS technology have clearly built upon and revived interest in theories and techniques of spatial analysis and cartography relevant long before the innovations in digital computing that made GIS possible. At the same time, technological developments in geographic information systems have opened and supported exploration of new avenues of scientific inquiry.

As a technology, GIS rely heavily on computer hardware and peripheral equipment like large-format scanners and printers that may not be part of the hardware configurations available to a public health analyst. The definition of GIS as "computer-assisted systems for the capture, storage, retrieval, analysis, and display of spatial data" (Clarke, 1986, p. 175) might lead one to assume that the acronym "is simply a catch-all for almost any type of automated geographic data processing" (Cowen, 1988, p. 1551). In fact, GIS are part of a larger constellation of computer technologies for capturing and processing geographic data as discussed in Chapters 2 and 3.

Some of these technologies, like the ***Global Positioning System*** (GPS) and remote sensing, are used to collect geographic data. The GPS developed by the U.S. Department of Defense relies on a series of at least 24 satellites in orbit and a network of satellite sensors or tracking devices located on the earth's surface (Laurini, 2001). Portable receivers capture signals from the satellites and calculate very accurate surface positions from the measurement of satellite positions.

Remote sensing is the analysis and interpretation of data gathered by means that do not require direct contact with the earth. ***Aerial photography*** of the earth's surface, taken with an aircraft-mounted camera, is an important source of up-to-date data (Jensen, 2007). Aerial photographs can be scanned and rectified for printing or viewing on a display screen, as described in Chapter 3. In addition to aircraft, satellites also serve as platforms for devices that capture information about the earth's surface. ***Digital image processing*** for geographic data collection involves the use of satellites with sensors capable of detecting electromagnetic energy reflected off or emitted from objects on the earth's surface (Jensen, 2005). The data are then enhanced for viewing and analysis.

Computer technology also plays a part in the collection of secondary geographic data from existing maps (Chang, 2009). ***Scanning*** is a technique for capturing map data in digital form. Scanners use an optical laser or other electronic device to "read" a map and convert its features to a computer database of dark and light values. To use the scanned image as more than just a backdrop in the GIS display, ***vectorization*** techniques implemented with GIS software can be used to recognize specific elements and convert them to points, lines, and areas representing features of interest. Pattern recognition techniques developed in artificial intelligence research are used in these processes. With scanning, the map manipulation into digital format happens relatively quickly but the processing of the scanned image to recognize cartographic objects takes much more time.

Digitizing requires use of a tablet and cursor to record coordinate locations of map features from a map placed on a digitizing tablet (Chrisman, 1997, p. 70). It is also possible to construct a GIS data layer by ***screen digitizing***. This process, also known as ***heads-up digitizing***, uses a pointing device like a mouse to move over and capture coordinate locations from a digital database or image file displayed on the monitor. Scanning and vectorization are becoming more important as procedures for data acquisition, and there is less emphasis on digitizing, at least with tablets. Screen digitizing, however, is still widely used.

Remote sensing, GPS, scanning, and digitizing are the main methods for spatial data collection. One or more of these methods may be used, either by the person using the GIS or by the government organization or commercial vendor from whom a spatial database is purchased. The nature of spatial data and important issues related to scale, resolution, accuracy, storage, and retrieval of health and health-related data that must be considered at the data capture stage are discussed at length in Chapter 2.

Once spatial data exist in digital form, computer graphics software supports the creation of cartographic displays. ***Computer-aided design*** (CAD) systems, as used with cartographic data, provide support for drafting and producing map-like displays of features of interest like roads or land parcels. As they originally developed, CAD systems like other computer graphics software made use of (X,Y) Cartesian coordinates local to the software rather than geographical coordinates like longitude and latitude. Furthermore, CAD systems did not link features to databases describing the features' attributes, so it was difficult to produce thematic or statistical maps using a CAD system. But developments in computer cartography have created systems for linking geographical and attribute databases to produce statistical maps. In computer mapping software systems, the ability to manipulate the geographic data—the location information—is usually quite limited.

The boundaries separating GIS from these other technologies are not rigid—in part because the technologies continue to evolve. Remote sensing, CAD, and the rest are valuable technologies that are related to GIS. But they are not, in and of themselves, considered GIS. Most efforts to define GIS as a technology emphasize the special nature of GIS software and the data structures on which the software is built.

These data structures and software functions are specifically designed to integrate and analyze data based on location. Spatial data representation and spatial analyses like map overlay have been implemented without the use of computer-based systems like GIS (Poiker, 1985; Charles, 2005). People thinking about using GIS who are not familiar with basic geographic and cartographic concepts and techniques can benefit greatly from learning these concepts before and as they become GIS users. The data and modeling requirements of most GIS research applications, however, necessitate the use of computers.

We define GIS, as a technology, as computer-based systems for integrating and analyzing geographic data. The locations of features on the earth's surface are stored so that neighborhood relationships among features can be analyzed and so that groups of different features sharing the same locations can be identified (Figure 1.1). The following software functions distinguish GIS as a technology (Goodchild, 1995b):

- The ability to store or compute and display spatial relationships between objects (e.g., the house is adjacent to the toxic waste site, the school is connected to the water supply system).
- The ability to store many attributes of objects.
- The ability to analyze spatial and attribute data in addition to simply managing and retrieving data.
- The ability to integrate spatial data from different sources.

Analysis and integration of geographic data requires a wide range of software functions.

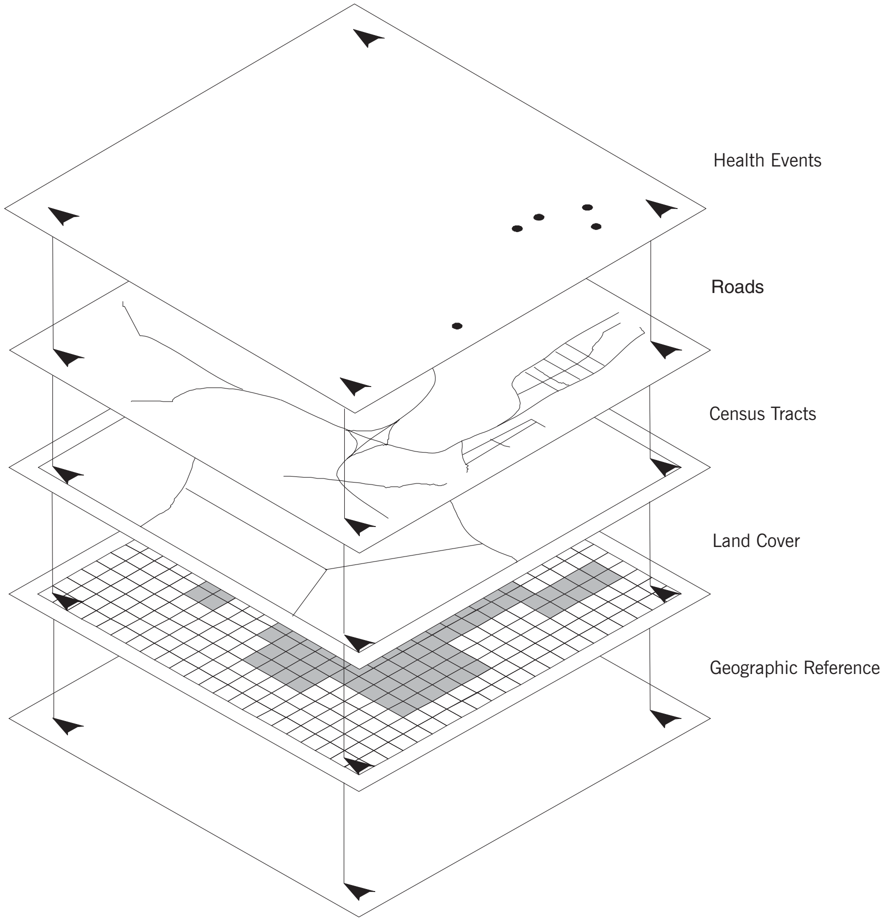


FIGURE 1.1. Digital geographic databases registered to a common geographic reference system. A composite of two or more layers can be produced because the geographic references match.

GIS Functions

What are the main functions of GIS software? The early lists of integrated software functions required to support input, analysis, and output of data (Tomlinson & Boyle, 1981; Dangermond, 1984) have been supplanted by textbooks and reference guides outlining GIS analytical methods (DiBiase et al., 2006; Chang, 2009). The applications selected for discussion in this book were intended to illustrate how some of the most important generic GIS functions can be used to support public health research and policy analysis. The "toolbox approach" to defining GIS as a set of software functions is useful for comparing the hundreds of software packages on the market, but it says little about the software functions in relation to the spatial data they process (Cowen, 1988).

The approach taken here classifies GIS functions in user terms, that is, based on what people want to do with spatial data. Three broad categories emerge: spatial database management, visualization and mapping, and spatial analysis. As noted earlier, health analysts need and want to manage spatial databases. This includes creating databases of health events located on the earth's surface that can be processed and stored by computers, keeping track of changes by adding and deleting events from those databases, and editing the location and public health attributes of these events, including data describing the contexts in which health events occur. Public health professionals and epidemiologists also want to visualize and map the spatial data they have acquired. This includes exploring visual representations of the patterns of health outcomes and risk factors and communicating information in the data to others in the form of maps. Equally important, public health researchers want to analyze the spatial relationships among the health events stored in the databases and to create new classes of health patterns based on those relationships. The development of the Internet has made it possible for some of these GIS functions to be supported in a network environment.

Spatial Database Management

One important function of GIS is managing spatial data. Identifying sources of, collecting, and preprocessing the spatial and health attribute data that would be managed in public health GIS are discussed in detail in Chapter 3. ***Database management systems*** (DBMS) are used to store, retrieve, and manipulate data in a database. A GIS software product, like other computer software systems, is built on an underlying data model. A ***data model*** is a detailed model that captures the overall structure of the data, independent of database management or implementation considerations (Elmasri & Navathe, 2007). A data model includes the relevant entities, relationships, and attributes, as well as constraints defining how the data are used. In the 1980s, a large body of research on the science of GIS explored the nature of spatial data models (Peuquet, 1984; Goodchild, 1992).

Spatial data embody complex and often hierarchical relationships that are not easily expressed in tables (Yearsley, Worboys, Story, Jayawardena, & Bofakos, 1994). As a result, GIS software packages are different from simple spreadsheets. Relational, object, and object-relational database management systems are all used in GIS (Longley, Goodchild, Maguire, & Rhind, 2001).

Relational database designs were dominant by the early 1990s (Healey, 1991; Longley, Goodchild, Maguire, & Rhind, 2001). ***Relational database management models*** organize data in the form of tables ("relations") where the rows or ***tuples*** represent units of analysis or objects of interest and the columns represent attributes (Elmasri & Navathe, 2007). Relational database management systems manage data as a collection of tables. Data in different tables can be linked when there are common attributes in the tables. In public health applications, the tables might include individuals diagnosed with a particular type of health problem and listed in a disease registry, public drinking water lines, or hazardous waste sites (Figure 1.2). The associated attributes may be quantitative or qualitative, numeric or alphanumeric, and reported as nominal, ordinal, or interval/ratio data.

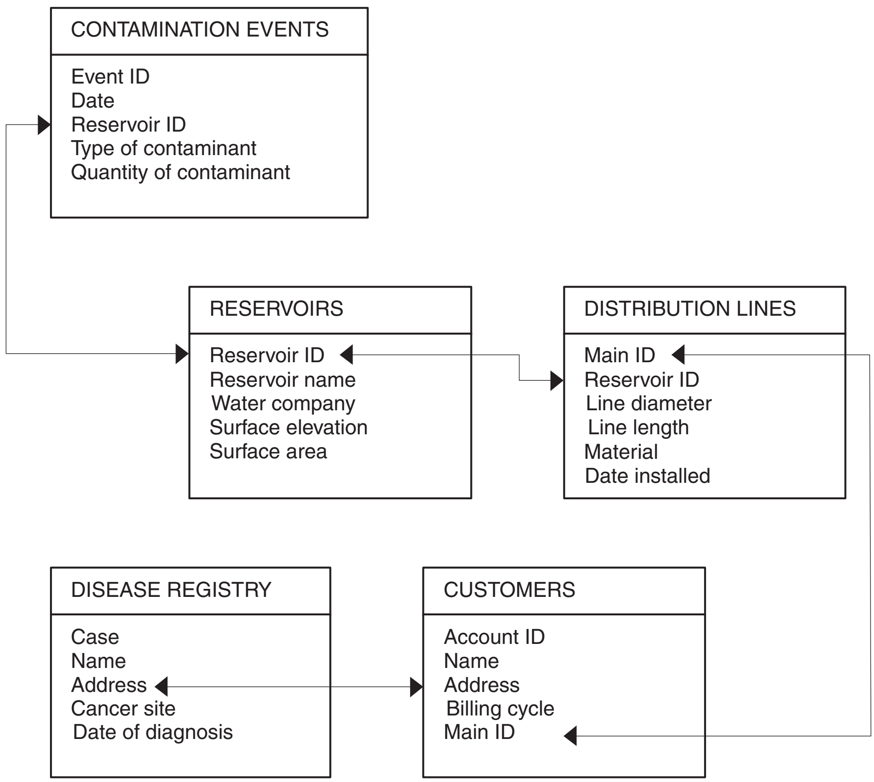


FIGURE 1.2. Tables of data from different sources containing fields that could be linked in a relational database to describe health outcomes associated with contamination events.

The geographic objects whose attributes might be described in a table can also be assigned spatial dimensions or topological properties (Laurini & Thompson, 1992). Objects represented as ***points***, like public drinking water well locations, are zero-dimensional, that is, they have specific locations but no sizes. ***Lines***—sometimes called arcs—are one-dimensional objects, like the route of an emergency medical response vehicle; they have lengths but no specified widths. ***Areas***—also called polygons—are two-dimensional objects enclosing a space, like health planning districts. Some GIS are also capable of representing three-dimensional objects. In a spatial database, the objects in the database have locations tying them as points, lines, or areas to the earth's surface (Figure 1.3). Thematic attributes, like reservoir yield or distribution line date of completion, are also stored and can be retrieved and analyzed.

In a relational database, data in different tables can be linked by matching values in a column in one table to values in another table until the data of interest have been retrieved from all tables (Worboys & Duckham, 2004). These

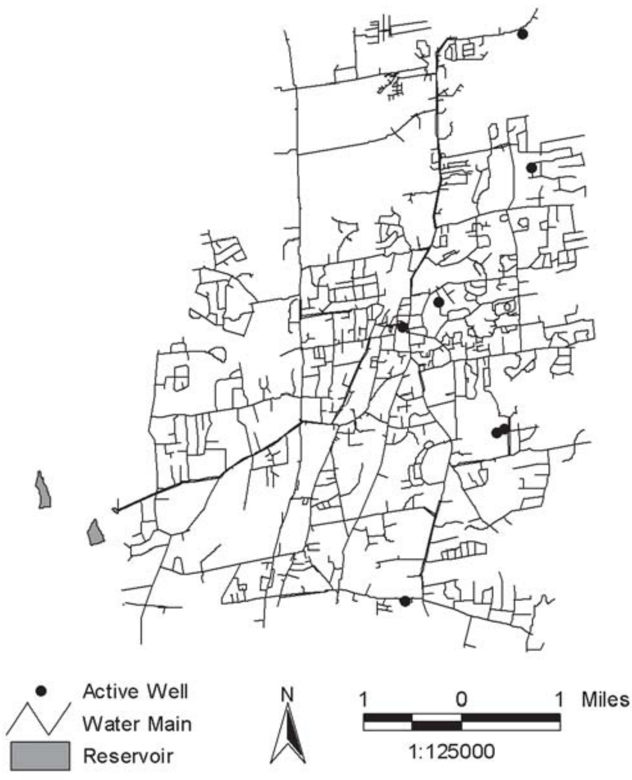


FIGURE 1.3. Three spatial databases used to model a public drinking water system. Active wells are modeled as points. Water mains are modeled as lines. Reservoirs are modeled as areas.

relationships can be one to one, one to many, many to one, or many to many. One-to-one relationships occur when a record in one table links to exactly one record in another table. For example, one private well serves only one private residence, and each such residence is served by only one private well. This kind of relationship is simple but rare. One-to-many and many-to-one relationships are more common. One water company serves many customers and vice versa. These relationships are relatively easy to model in relational databases.

Most spatial relationships are many to many. One point on the earth's surface can be enclosed in an infinite number of areas, and those areas contain many points. Because spatial databases are so large, it is necessary to retrieve many rows of data in a matter of seconds for cartographic display of objects on the screen. In an **"integrated" data model** (Healy, 1991; Worboys & Duckham, 2004), the location information or spatial data and the attribute information are stored together in individual rows of the database. The **"hybrid" data model** (Healy, 1991; Worboys & Duckham, 2004) separates storage of the location information or digital cartographic data (the points, lines, and areas) from the storage of the attribute information. The attribute information is stored in a commercial relational DBMS, but the digital cartographic data are stored in direct access operating system files that speed input and output. The GIS software links the two during processing. The integrated data model received renewed attention as a model that facilitated the use of Structured Query Language (SQL) and stimulated interest in object-oriented approaches (Worboys, 1999).

In contrast to relational databases, ***object databases*** include objects as the basic units and all of the properties that define both the state and the behavior of the object. An object is recognized by its ***oid*** or ***object identifier,*** which is always retained by the object and is independent of object attribute values. No explicit key is needed to identify an object in this type of database. The attribute values for a given object constitute the object's ***state*** at a point in time. The ***behavior*** of an object is the set of operations that can be applied to the object. A public drinking water service district as a region object in a geographic database might have a name, the district's population, and a set of points defining its boundary as attributes describing its state. The behavior of the district object might include operations like calculating the district's area and perimeter, operations that would not be appropriate behavior for a point object like a wellhead.

Objects having the same structure and behavior may be grouped into ***classes.*** The residence of a drinking water company customer is an object that could be in several classes, the *customers* class and the *point* class. As part of the class *customers*, the object might have the customer_name as a descriptive string attribute. Because the residence is represented as a point feature, its geometry attribute would identify it as a point referencing a spatial object belonging to the *point* class. As a member of the *point* class, the customer's residence object would behave like all points and could support a range of operations that can be performed on point features.

Classes are particularly useful for representing spatial hierarchies. Another *customers* attribute, district_in_customer, identifies the water district where the customer's residence is located. This attribute places the customer's residence in a spatial hierarchy by linking the *customers* points to the water districts serving them and the water district areas where they are located. The water district is an object in the class *districts*, and objects in this class can be linked to their component customers.

Using the object model (state plus behavior) is a different approach to database design from using the relational model (finite set of rows containing data values), and comparing the two approaches reveals their advantages and disadvantages for representing spatial data (Rigaux, Scholl, & Voisard, 2002). The object-oriented approach has gained some ground recently because of its advantages for managing, querying, and processing complex geographic data, but relational database systems have begun to incorporate some of the capabilities of object database systems. ***Object-relational database management systems*** adapt relational database management systems to handle objects.

The generic spatial database management functions of GIS include functions to create and maintain spatial databases. These functions enable users to add and delete records from databases, to edit information, to join databases, and to perform other functions familiar to most users of computer DBMS. The spatial database management functions of the GIS allow users to view and manipulate data in table format (Table 1.1).

This view of the data is useful for editing and for looking up information in a particular record. For example, a health researcher might need to edit information on a particular public drinking water main by modifying a field in a table (Table 1.1) to show whether the main was constructed before or after a particular year. The tabular view is not as useful for exploring the relationships among units of observation—for example, examining the spatial pattern of mains constructed before or after a particular year to identify patterns of water service at the time of a contamination event occurring at a particular location (Figure 1.4).

TABLE 1.1. Attributes of Public Drinking Water Mains

Main ID	Length (ft.)	Diam. (in.)	Date	Material
1	1097.51	12.00	Pre-1968	TR
2	2146.09	12.00	Post-1968	TR
3	1916.36	6.00	Pre-1968	DT
4	627.76	8.00	Pre-1968	DT
5	1913.46	6.00	Pre-1968	TR
6	538.11	12.00	Pre-1968	N-L
7	735.26	12.00	Pre-1968	N-L
8	554.65	12.00	Pre-1968	N-L
9	83.64	12.00	Pre-1968	N-L
10	47.63	12.00	Post-1968	N-L
.	.	.	.	.
.	.	.	.	.
1715	710.64	8.00	Pre-1968	TYT

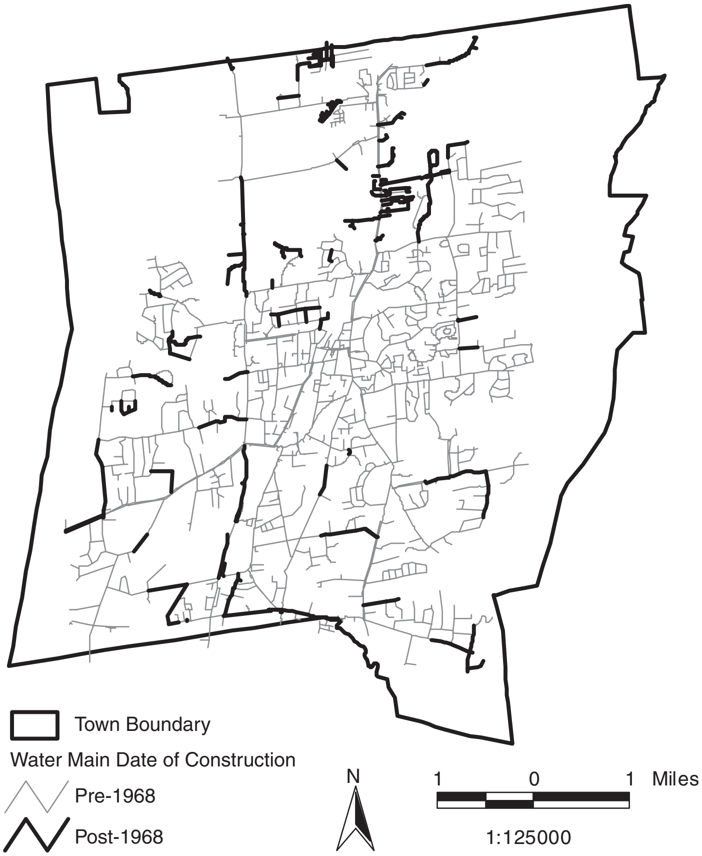


FIGURE 1.4. Mapping data on dates of water main construction to aid visualization.

Visualization and Mapping

Once the spatial database has been created and can be retrieved, the graphical display and mapping functions of GIS come particularly into play. Visualization as a form of human cognition—"making things 'visible in the mind' " (Wood, 1994)—has been considered a key to many important scientific advances. Components of this process include visual exploration and confirmation of data ("visual thinking") and synthesis and presentation of information ("visual communication") (MacEachren, Buttenfield, Campbell, DiBiase, & Monmonier, 1992).

Today's statistical graphics have their origins in the products of the late 16th century prepared as aids to abstraction in the research process (Buttenfield & Mackaness, 1991; Tufte, 2001). Like maps, other graphics were initially prepared and printed manually, with the data both stored and displayed in the printed product. The development of computers and computer graphics has made it possible to separate storage from display of large quantities of data. Computers have also made possible innovative cartographic and other kinds of displays that would be difficult or impossible to render by hand, including three-dimensional representations, animations, and scene generations. The development of GIS coincided with an increased interest in scientific visualization in general, a major topic in computing since the National Science Foundation report on "Visualization in Scientific Computing" was published in 1987 (McCormick, DeFanti, & Brown, 1987). The links between the visualization functions of GIS and scientific visualization in general are an ongoing area of interest (Hearnshaw & Unwin, 1994; Llobera, 2003).

GIS make it possible to view data in the form of tables, graphs, maps, and statistics. It is often difficult to compare and see trends in data in the form of tables without the use of graphs and maps. This point is illustrated using four fictitious data sets, each consisting of the same number of (*x*,*y*) pairs, presented in tables, statistics, and graphs (Table 1.2). Although statistical analysis of the four data sets yields the same standard output (ignoring residuals), the graphs highlight differences in the relationship between *X* and *Y* in the four data sets (Figure 1.5). Admittedly, three of the small data sets were designed to represent a particular effect in an extreme form. But the illustration is still effective in demonstrating that viewing the data in the form of a graph provides us with information that is more difficult to see in the tables or the statistics.

TABLE 1.2. Four Data Sets Each Comprising 11 (*x, y*) Pairs

	#1		#2		#3		#4	
Case ID	*X*	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*
1	10.0	8.4	10.0	9.1	10.0	7.5	8.0	6.6
2	8.0	7.0	8.0	8.1	8.0	6.8	8.0	5.8
3	13.0	87.6	13.0	8.7	13.0	12.7	8.0	7.7
4	9.0	8.8	9.0	8.8	9.0	7.1	8.0	8.8
5	11.0	8.3	11.0	9.3	11.0	7.8	8.0	8.5
6	14.0	10.0	14.0	8.1	14.0	8.8	8.0	7.0
7	6.0	7.2	6.0	6.1	6.0	6.1	8.0	5.3
8	4.0	4.3	4.0	3.1	4.0	3.1	19.0	12.5
9	12.0	10.8	12.0	9.1	12.0	5.4	8.0	5.6
10	7.0	4.8	7.0	7.3	7.0	8.2	8.0	7.9
11	5.0	5.7	5.0	4.8	5.0	6.4	8.0	6.9

Number of observations = 11	Sum of squares = 110.0
Mean of *X* = 9.0	Regression sum of squares = 27.50
Mean of *Y* = 7.5	Residual sum of squares = 13.75
Regression coefficient of *Y* on *X* = 0.5	Standard error of *b* = 0.118
Equation of regression line *Y* = 3 + 0.5X	*R*-squared = 0.667

Note. Adapted from Anscombe (1973). Reprinted with permission from *The American Statistician*. Copyright 1973 by the American Statistical Association. All rights reserved.

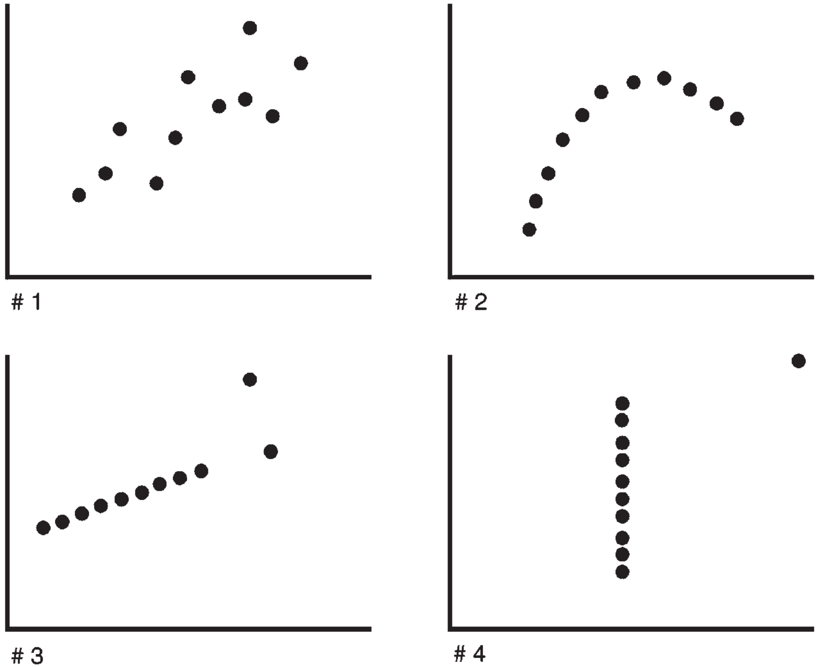


FIGURE 1.5. Scatterplots corresponding to the four data sets in Table 1.2. Adapted from Anscombe (1973). Reprinted with permission from *The American Statistician*.

The cartographic view also provides us with insights into the data that cannot be discovered through other means. Two pairs of variables (X_1, Y_1) and (X_2, Y_2) can have identical scatterplots and correlation coefficients, for example, but the variables can exhibit distinctly different patterns when mapped, even using the same class breaks (Figure 1.6). The statistical correlations are the same, but the geographical correlations are quite different. The maps for X_1 and Y_1 do not depict identical geographical distributions for the two variables, but they do point to an underlying geographic factor like proximity to a pollution source.

The visualization and mapping functions of GIS help us to see how observational units and their attributes are located in geographic space (Figure 1.7). In this example, the residences where arsenic measurements were taken are not evenly spread across the region. Across this set of residences, arsenic values could be arranged in many different ways. The map shows the observed pattern of arsenic measurement. The mapping and visualization functions of GIS are also useful in displaying the results of spatial statistical analyses to assess whether or not similar observed values are clustered together (Figure 1.8) (AvRuskin et al., 2004). Understanding the spatial distributions of the phenomena of interest is important for designing valid spatial sampling and spatial analysis strategies.

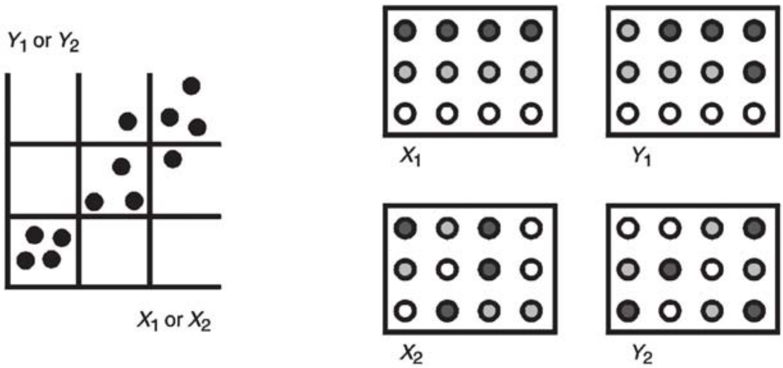


FIGURE 1.6. Two pairs of variables (X_1, Y_1 and X_2, Y_2) for 12 cases have identical scatterplots but very different spatial arrangements even when mapped using the same class breaks. Adapted from Monmonier (1996). Copyright 1996 by University of Chicago Press. Adapted by permission.

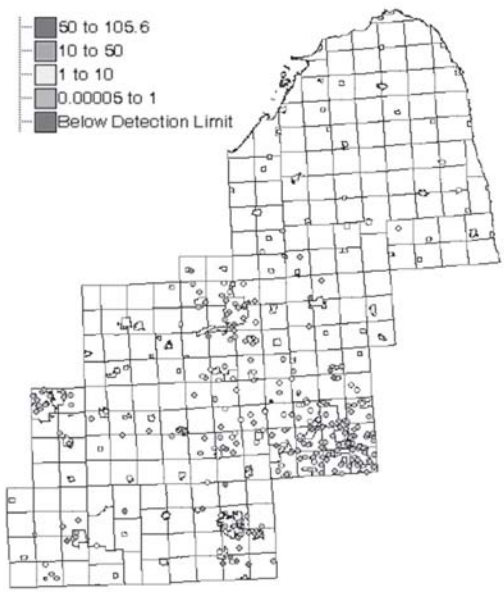


FIGURE 1.7. Each point represents an arsenic measurement in water taken from the tap at the current residence of participants in a study of arsenic exposure and bladder cancer in towns in Oakland and Genesee counties in Michigan. The map shows the spatial arrangement of the measurement sites. The measurement sites are clustered in more populated towns. From AvRuskin et al. (2004). Originally published by BioMed Central in the *International Journal of Health Geographics*, Open Access.

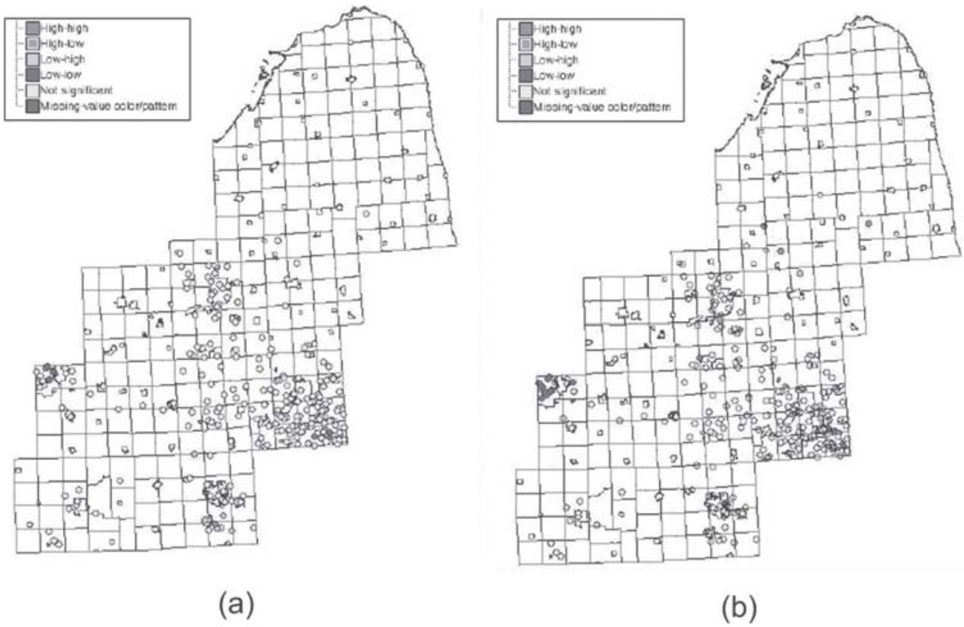


FIGURE 1.8. Maps showing the results of spatial statistical analyses to identify clusters of high- and low-arsenic measurements in a study of arsenic exposure and bladder cancer in southeastern Michigan. Mapping and visualization functions of GIS can be used to explore results when different parameters are used in an analysis. Figure 1.8a shows Local Moran values based on arsenic measurements for five nearest neighbors. Figure 1.8b shows the values based on 10 nearest neighbors. Measures of local spatial autocorrelation used in cluster analyses are described in detail in Chapter 5. From AvRuskin et al. (2004). Originally published by BioMed Central in the *International Journal of Health Geographics*, Open Access.

These spatial distributions are not easy to see or likely to be revealed in tabular or statistical displays of the data.

The importance of the visualization and mapping functions of a GIS must be understood in context. Throughout the development of computer-assisted cartography, cartographic research has investigated how traditional map representations can be accomplished in an automated environment (Buttenfield & Mackaness, 1991; Kraak, 1999). Cartographers have criticized GIS software companies for "a lack of attention" to principles of graphical design and for failing to adopt graphical defaults based on perception research. These problems affect visual communication through other types of graphics too (Tukey, 1977; Tufte, 2001). The growing importance of geographic visualization is evident in development of spatial multimedia and virtual reality systems, but the connections between these systems, cartography, and GIS are still being articulated (Thurston, Poiker, & Moore, 2003; Slocum, McMaster, Kessler, & Howard, 2009).

Setting aside issues related to the quality of the visual display, overreliance on visualization poses its own problems. "While the spatial perspective can be very powerful as a source of insight, it can also be highly misleading" (Goodchild et al., 1992, p. 409). Perceptions are context-dependent and change with experience. Some mechanism must be present for checking our perceptions and intuitions against other subjective and objective criteria. For example, the problems inherent in detecting and explaining meaningful spatial clusters of disease illustrate the tensions between visualization and analysis. A GIS offering only spatial database management and visualization capabilities would be incomplete. Despite these limitations, generic GIS visualization and mapping functions enable users to see the spatial relationships present in large and complex databases and to report the results of an analysis in cartographic and other graphic displays. Furthermore, GIS allows users to display available spatial databases quickly, easily, and interactively. These functions have provided spatial data analysts with a powerful mechanism for exploring spatial data.

Spatial Analysis

GIS software systems enable public health analysts to do more than simply manage and map data. GIS support a range of spatial analysis functions (de Smith, Goodchild, & Longley, 2007). ***Spatial analysis*** refers to "a general ability to manipulate spatial data into different forms and extract additional meaning as a result" (Bailey, 1994, p. 15). Specifically, spatial analysis comprises a body of techniques "requiring access to both the locations and the attributes of objects" (Goodchild et al., 1992, p. 409). The results of a spatial analysis are "not invariant" when locations of the objects being analyzed are changed. As such, spatial analysis covers a broad range of numerical methods.

The spatial analysis functions of GIS fall into five classes: measurement, topological analysis, network analysis, surface analysis, and statistical analysis (Table 1.3). The ***measurement*** functions of GIS allow the user to calculate straight-line distances between points, distances along curved paths or arcs, and areas. Although the measurement functions are relatively few in number, they are extremely important. Distance as a measure of separation in space is a key variable used in many other kinds of spatial analysis. Distance is often an important factor in interactions between people and places. Most GIS can create buffers around points, lines, and areas, depicting all of the area within a user-specified distance of the objects (Figure 1.9).

Topological analysis functions include the software functions used to describe and analyze the spatial relationships among units of observation. This category also includes spatial database overlay and assessment of spatial relationships across databases, including map comparison analysis. The public health analyst could identify the area within a specified distance of a public drinking water well or surface source, for example, and overlay the footprint of a proposed building to determine whether or not the building would be located far enough from the water source to meet legal requirements (Figure 1.9). Topologi-

TABLE 1.3. Spatial Analysis Functions of GIS

Function class	Function	Chapters
Measurement	Distance Length Perimeter Area Centroid Buffering Volume Shape Measurement scale conversions	Chapter 6, 9, 10, 11
Topological analysis	Adjacency Polygon overlay Point-in-polygon Line-in-polygon Dissolve Merge	Chapters 4, 5, 6, 8, 11
Network and location analysis	Connectivity Shortest path analysis Routing Service areas Location–allocation modeling Accessibility modeling	Chapters 9, 10
Surface analysis	Slope Aspect Filtering Line of sight Viewsheds Contours Watersheds	Chapter 5
Statistical analysis	Spatial sampling Spatial weights Exploratory data analysis Nearest neighbor analysis Global and local spatial autocorrelation Spatial interpolation Geostatistics Trend surface analysis	Chapters 5, 6, 7, 9, 11

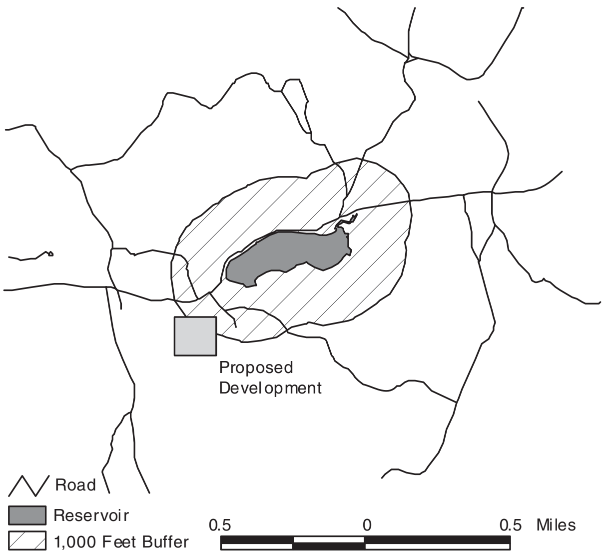


FIGURE 1.9. Buffering a polygon representing a public drinking water reservoir to show the area within 1,000 feet of the reservoir shoreline.

cal analysis functions can identify features in the landscape that are adjacent or next to each other such as adjacent census tracts. They can also be used to model containment by identifying features such as residences that are located within areas such as census tracts. Topology is also important in modeling connectivity in networks.

Network analysis is a branch of spatial analysis that investigates flows through a network. The network is modeled as a set of nodes and the links that connect the nodes. Once a network has been defined, it is possible to analyze the network and flows through it. These analytical models would be used, for example, to determine the shortest path through a street network for an emergency medical response (Figure 1.10). Network analysis functions are also used to model service areas of facilities and to locate facilities.

Surface analysis techniques are often used to analyze terrain or other data that represent a continuous surface. Filtering techniques include smoothing and edge enhancement. Smoothing removes "noise" from the data to reveal the broader trends. Edge enhancement accentuates contrast and aids in identifying linear features like highways or fault lines. Line-of-sight analysis, viewshed, and watershed analyses apply to digital elevation models. Public health analysts who need to model environmental conditions related to terrain would make the greatest use of these GIS functions.

Spatial data analysis is closely tied to spatial statistics (Haining, 2003), and this fifth category of spatial analysis functions is emphasized by researchers with an interest in spatial statistics and exploratory data analysis (Fotheringham &

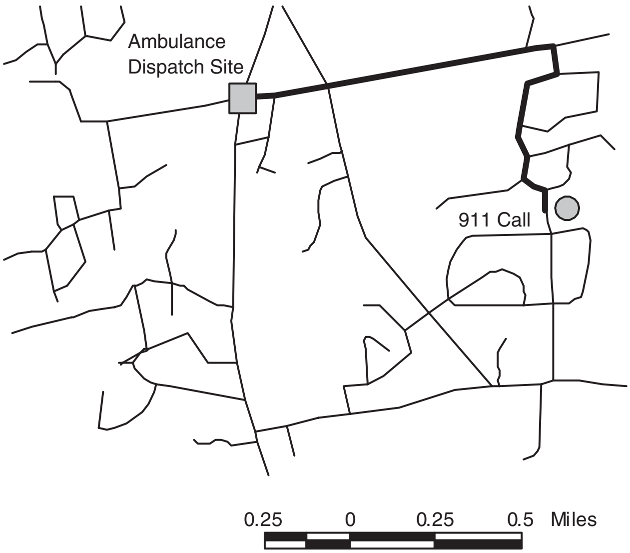


FIGURE 1.10. The shortest path through a street network from an ambulance dispatch site to the location of an emergency call.

Rogerson, 1994; Bailey & Gatrell, 1995). Epidemiologists make extensive use of multivariate inferential statistics (Gordis, 2004), but in many instances these methods are not forms of *spatial* statistical analysis because they assume that the observational units represent independent pieces of information about the relationships being modeled and they focus on the attributes of objects and events rather than their spatial relationships. The development of GIS has coincided with a growing interest in the application of spatial statistical methods in public health and epidemiology (Waller & Gotway, 2004; Lawson, 2006).

Software for performing many spatial statistical analyses was not initially part of GIS software packages (Goodchild et al., 1992). The measurement, topology, and network functions—along with some simple aspatial statistical functions—were more commonly featured. Considerable progress has been made in making spatial statistical software available, especially through the World Wide Web (WWW). GeoDa, a set of software tools developed by Luc Anselin, is one example (Anselin, 2003a). Several options for linking statistical and modeling software systems with GIS software are available: loose coupling, moderate coupling, tight/close coupling, or embedding (Goodchild, 1992; Nyerges, 1993).

When the statistical or modeling software and the GIS software systems are freestanding, separate software packages are developed for different analysis functions. One argument against this approach is that any spatial statistical analysis requires access to spatial data input, editing, management, and display functions, and it seems pointless to duplicate this functionality in stand-alone software. ***Loose coupling*** involves moving output from the GIS software analysis (e.g., a measure of whether or not a person is exposed to an environmental contaminant based on residential location) into another statistical software package where this input might be used to calculate and test the statistical significance of risk ratios. The results of the analysis might then be returned to the GIS for display and further analysis. ***Moderate coupling*** relies on techniques allowing indirect communication between systems, like shared database access between the systems. ***Tight/close coupling*** allows direct communication between the two systems during program execution so that the systems are operating simultaneously. Medium and tight/close coupling obviously require the necessary programming expertise (Westervelt, 2002).

Embedding is an alternative to coupling as an approach for linking GIS software to other software systems like those supporting agent-based modeling of processes like disease diffusion or other forms of simulation modeling (de Smith, Goodchild, & Longley, 2007). This approach chooses the GIS or the other modeling software system as the dominant system and embeds the needed software function within the dominant system using that system's underlying programming language. The Hazard Prediction and Assessment Capability (HPAC) System, an automated software system that predicts the effects of hazardous material released into the atmosphere, is an example. Alternatively, the embedding approach manages connections between the two systems. The Agent Analyst Extension of ArcGIS, an agent-based modeling extension, uses this approach (Johnston & Maguire, 2007). Expanding the set of functions available in the GIS package makes them available to all users of the software; the likelihood of functions being incorporated into commercial GIS packages is related to the number of customers who demand the particular function.

The increasing emphasis on open systems in computing has affected the GIS industry. Many software vendors have developed products that make it easier for professional software developers or system users within an agency to create specially designed GIS applications with modified user interfaces and functions. These modifications can be made available to other users by fully incorporating them into a software update distributed by the vendor, marketing them as new software products, or distributing them for free to interested software users via the Internet. The "open GIS" movement has been particularly concerned with developing a standardized interface to ease data transfer among geographic information systems that use different, proprietary, approaches to managing spatial data (Phair, 1997).

Trends in GIS Applications

Although GIS have developed rapidly, the history of the technology "is little more than anecdotal" (Coppock & Rhind, 1991, p. 21). An overview of GIS applications in the United Kingdom, the United States, and Canada gives a sense of GIS developments in those countries up to 1989 (Bracken, Higgs, Martin, & Webster, 1989). The growth and development of a European GIS community has also been described and analyzed (Masser, Campbell, & Craglia, 1996). A formal history considers the many origins of GIS and the various influences affecting GIS development (Foresman, 1998), although the work of the military is overlooked in this account (Smith, 1992; Cloud, 2002). Some of the earliest systems were developed for environmental management, particularly of land-based resources like forests (Tomlinson, 1987). Many early adopters were large public or quasi-public institutions that had access to or were responsible for compiling extensive spatial data sets in the form of maps (Coppock & Rhind, 1991). These organizations generally also had access to the hardware resources needed to support storage and analysis of large databases in the mainframe era of the 1960s and 1970s.

Adoption of GIS and related technologies also occurred relatively early, but selectively, at the local level in the United States. County and local governments are responsible for land registration, deed transfers, and property taxation. Local governments were interested in GIS for the opportunities they offered to create and manage ***cadastral databases***, digital land parcel or property databases. Up to the 1990s, effective use of these systems to monitor property changes was primarily at the local level where the databases involved were more manageable in terms of size (Dale, 1991; Dale & McLaren, 1999) and directly related to urban and regional planning (Parrott & Stutz, 1991; Yeh, 1999).

At the state level in the United States, transportation and environmental management agencies were early adopters of GIS. In most states, one of these two departments generally took the lead in early GIS implementation. Public utilities (Mahoney, 1991; Meyers, 1999) and civil engineering firms were among the earliest nongovernmental adopters of GIS. Utilities, transportation, and telecommunication companies have also been leaders in the development of "mobile" GIS, which supports spatial data acquisition and attribute data entry in the field (Wilson, 1998).

The rapid diffusion of GIS technology to new application areas has depended on a number of developments in hardware, software, and spatial databases. Hardware has been recognized as one of the major "drivers" of GIS development (Dangermond & Morehouse, 1987). Larger memories for lower costs, workstations and desktop computers with high levels of graphics performance, network architecture as an alternative to multiuser host architecture, and low-cost, reliable output devices like the inkjet printer were among the most important hardware developments affecting GIS in the late 1980s and early 1990s.

The higher levels of graphics performance have also made graphical user interfaces possible in GIS software (Buttenfield & Mackaness, 1991). Many of the software packages initially developed relied on command- or query-based user interfaces that presented fairly steep learning curves for the software. Hardware and software developments and new operating systems have also forced GIS software companies to broaden their product lines to operate on a variety of platforms.

Although hardware and software developments have supported the diffusion of GIS technology over the last three decades, development of "foundation" spatial databases was a key to the rapid adoption of GIS technology by the business community. Most commercial enterprises sit on large databases containing location information (e.g., retail outlet locations, customer locations, supplier locations, shipment routes) and make decisions about how to manage business operations in a geographic context. The development of the U.S. Bureau of the Census TIGER/Line® files (Callahan & Broome, 1984), a database covering the entire United States, for the 1990 census and their availability to the public at a relatively low price helped to create new markets for GIS software among census data users in the United States.

The ***TIGER/Line files,*** discussed in detail in Chapter 3, are a database containing line segments for streets and other linear features that can be used to create digital cartographic databases of census block, census tract, and other administrative and political boundaries (Marx, 1986). Address ranges on the street segments enable users to translate street addresses to locations on the earth's surface. The digital cartographic databases that could be created from TIGER could also be integrated with data from the Census of Population and Housing. New GIS-related businesses and products emerged as commercial vendors developed upgraded or customized versions of the TIGER/Line files. Within the research community, the availability of this database accelerated the expansion of GIS technology into the social sciences and public health.

Public Health Applications of GIS

There are only a few examples of GIS applications in public health, epidemiology, or health planning from the 1980s (Bracken et al., 1989). Applications expanded rapidly during the 1990s and growth continues into the 21st century (de Lepper, Scholten, & Stern, 1995; Gatrell & Löytönen, 1998). The hardware, software, and database developments that have brought other new users to GIS partly explain this diffusion into the public health sphere. Like many other organizations, public health agencies and public and private providers or insurers of medical care services manage large databases that contain geographic information that can be meaningfully integrated based on location.

The primary impetus for the diffusion of GIS technology into the public health field in the United States was the federal government. A workshop on automated cartography and epidemiology organized by the National Center for Health Statistics in 1976 brought together representatives from federal agencies and the research community in response to an increasing awareness of computer-based mapping and geographical analysis (National Center for Health Statistics, 1979; Aangeenbrug, 1997). In addition to those agencies using computer-based systems for mapping and spatial data analysis at the national scale, other federal agencies began to adopt GIS technology.

The growing interest in environmental health, including risk assessment, created one "market" for GIS applications in public health (Stockwell, Sorenson, Eckert, & Carreras, 1993). Given that environmental management was an early GIS application area, GIS data layers describing environmental conditions were available to support these efforts. In 1994, the Agency for Toxic Substances and Disease Registry (ATSDR) held a workshop on GIS applications in public health and risk analysis to reinforce its commitment to GIS technology as a key tool for assessing “real risks to real people.” The resurgence of infectious disease (National Science and Technology Council, 1995), particularly vector-borne disease, and the international efforts to address health problems like Lyme disease and rabies based on an understanding of the ecology of these diseases also led public health agencies to GIS.

Through the cooperative agreements that fund state and local health departments, these agencies stimulated the development of public health GIS applications at the state and local level and also brought the research community into collaboration with public health agencies at every level. By the mid-1990s, GIS sessions began to appear on the programs of public health conferences like the Public Health Conference on Records and Statistics and the annual meeting of the American Public Health Association (National Center for Health Statistics, 1995; American Public Health Association, 1996). Successful conferences focused solely on public health GIS, for example, the International Symposium on Computer Mapping in Epidemiology and Environmental Health held in 1995 (Aangeenbrug, 1997) and the Third National Conference on GIS in Public Health held in 1999 (Richards, Croner, Rushton, Brown, & Fowler, 1999), were supported by a wide range of sponsoring organizations at the national level.

In addition to organizing conferences, federal agencies like ATSDR and university-based researchers have taken the lead in developing GIS training programs specifically for public health professionals (Rushton, 1997; Richards, Croner, Rushton, Brown, & Fowler, 1999). During this period, ATSDR, the Centers for Disease Control and Prevention, and the National Center for Health Statistics initiated a GIS lecture series, an online public health GIS newsletter, and other GIS programs (National Center for Health Statistics, 2006a).

While federal leadership in supporting the development of public health GIS has continued, private-sector adoption of GIS for health is more difficult to track in the United States. The trends in GIS development described in the preceding section coincided with a period of increasing privatization of health insurance and health services delivery in the United States. Corporate concentration in the medical care industry itself and the movement of the insurance industry into managed care created a market for GIS technology (McManus, 1993). However, analyses of utilization and service organization conducted by private or nonprofit health care providers have rarely been shared in the published literature.

By 2000, international organizations like the United Nations had initiated GIS programs in the World Health Organization and other agencies (Department of Public Information, 2000). As in the United States, the use of GIS by health agencies in a number of other countries increased rapidly during the 1990s (Smith, Gould, & Higgs, 2003; Houghton, 2004). Changes in mapping and information technology have affected the development, design, and publication of international and national mortality and disease atlases. The first national small-area cancer mortality atlas of the United States was published in 1975 (Mason, McKay, Hoover, Blot, & Fraumeni, 1975). In the following decades, additional series of U.S. cancer atlases, general mortality atlases, and specialized atlases appeared in print and online (Pickle, 2009). The development of these atlases was informed by cognitive research on map reading, developments in cartographic design, and advances in web-based delivery of GIS-enabled atlases (Hermann & Pickle, 1996; Brewer & Pickle, 2002; MacEachern, Crawford, Akella, & Lengerich, 2008). Digital spatial data were used to create an innovative mortality atlas for the United Kingdom available in print (Shaw, Thomas, Smith, & Dorling, 2008) and online (SASI Group, 2008). Many of these atlases like the WHO Global Health Atlas and the CDC Interactive Atlas of Reproductive Health are interactive and make it possible for users who are not familiar with GIS or statistics to analyze demographic data, explore trends, and make comparisons across places (Centers for Disease Control and Prevention, 2008a; World Health Organization, 2011).

At the same time that the use of GIS in health has increased, applications have become more sophisticated. The *International Journal of Health Geographics*, an open access journal that began publication in 2002, has documented the breadth of GIS applications in health around the world. Nevertheless, there is room for expansion in the adoption and use of GIS by health organizations and agencies.

The development of public health GIS to date reflects, in part, lags in the availability of geocoded health data compared to other health-related GIS databases. Throughout the 1990s, "geocoded public health data have been in relatively short supply, limited to states with initiatives to geocode vital statistics data or to individual investigators who could geocode their own data" (Richards et al., 1999, p. 359). In 1997, only 21 of 49 state directors of vital statistics who responded to a survey reported that their states used some type of automated geocoding of vital statistics records. As part of its Healthy People 2010 initiative, the U.S. Department of Health and Human Services set Objective 23-3 to increase the proportion of all major national, state, and local health agencies that use geocoding, which would in turn promote the use of GIS at all levels (U.S. Department of Health and Human Services, 2000). The target level for meeting this objective was 90% of all public health data systems. At the time of the midcourse review in 2004, the use of geocoding in major health data systems had not increased significantly (U.S. Department of Health and Human Services, 2006).

Not all health agencies or organizations have the trained staff, software, and hardware necessary to apply GIS technology. Organizations developing a GIS for public health analysis will not necessarily require access to the full range of GIS functionality. With Internet tools for managing spatial data, even analysts with limited resources can acquire the hardware and software they need to geocode data and develop public health applications.

GIS and the Internet

Maps and geographic data have been an important category of online content since the introduction and growth in popularity in the early 1990s of the WWW as a tool for accessing the Internet. In addition to the spatial database management, visualization and mapping, and spatial analysis functions that in-house GIS users call upon, publication and distribution of spatial data are increasingly important GIS-supported activities that enable organizations to share data, maps, and information as files or images sent over the Internet. Distributed GIS systems may rely on the Internet to connect web clients to web, GIS, and data servers, or they may support mobile GIS using wireless networks and devices. Both forms of distributed GIS make it possible for many more users to obtain geographic information. Because geographic information is some of the most complex digital information available, distributed GIS services have special design requirements for representing and disseminating information.

Distributed GIS support four main applications (Peng & Tsou, 2003):

- Data sharing.
- Information sharing.
- Data processing.
- Location-based services.

Data sharing is accomplished in several ways. Data in the original format and metadata describing the geographic database may be published for downloading on an organization's website. ***Metadata*** are "data about data" (Green & Bossomaier, 2002, p. 95). The role of metadata and metadata standards are discussed in detail in Chapter 2.

MassGIS, the source of some of the databases used in the online supplement accompanying this book, is an example of a data sharing application (Office of Geographic and Environmental Information, 2011). Organizations may put data in a standard format like ***GML*** (Geography Markup Language), the XML (Extensible Markup Language) grammar defined by the Open Geospatial Consortium to express geographic features (Open Geospatial Consortium, 2011a), and distribute it on the web, or they may participate in a data clearinghouse network like the NSDI (National Spatial Data Infrastructure) Clearinghouse Network (Federal Geographic Data Committee, 2009) and create a clearinghouse node from which their collections can be searched through a portal like geodata.gov (Geospatial One-Stop, 2011).

A second type of application involves information sharing through an online GIS. Agencies can publish and periodically update static maps, or they can maintain web-based GIS that allow users to look up information, perform map queries, or obtain real-time information like traffic or weather information. These applications rely on a multitiered client–server design (Plewe, 1997; Gao, Mioc, Anton, Yi, & Coleman, 2008). In this system design, the public health analyst sitting at a client computer uses a web browser and sends a request to one or more web server computers where a GIS application is running (Figure 1.11). This request is translated by programs connecting the web server to the GIS, which processes the request and returns a result in the form of a map, text, or data file. The result is reformatted by the programs connecting the web server to the GIS into a format understood by the web browser and then returned to the public health analyst's client computer, where it is displayed. This pattern of requests and responses may be repeated many times during a single session.

A third application of distributed GIS involves serving GIS analysis tools themselves. This type of application enables users to edit spatial databases and perform analyses even though the databases and GIS software are not installed locally on their computers. The South Carolina Department of Health and Environmental Control supports a geocoding service accessible through a standalone web interface (Shoultz, 2005). This agency web service geocodes nearly 100% of all data submitted and outputs data in multiple coordinate systems, with variables describing the accuracy level of the geocoded records based on agency standards. The service uses the state street centerline database for the emergency 911 system. The service has been incorporated in the state's system for participating in the National Electronic Disease Surveillance System (NEDSS) and the state's Vital Records Reengineering project.

Applications to serve GIS analysis tools are most commonly found within agencies, but there are examples of tools that serve extended communities of users. The South Carolina Vital Records project makes it possible for all birthing hospitals in the state to register live births electronically (South Carolina Vital Record and Statistics Integrated Information Systems Project Team, 2005). The Office of Public Health Practice in the Public Health Agency of Canada supports both a Map and Data Exchange for networking and sharing information. It also developed and maintains the Public Health Map Generator (Public Health Agency of Canada, 2007). This tool enables users to link their health data to spatial data and to design map displays that can be saved, printed, and published in reports and presentations. The spatial data are drawn from the Canadian GeoSpatial Data Infrastructure (CGDI) warehouse of geographic data from networked data suppliers across the country.

Finally, location-based services allow users to access information about a location and its surrounding area (Boulos, 2003). A person who becomes ill in a particular city can use location-based services to find nearby walk-in medical care services, for example, and to determine the best route to reach a particular destination (Löytönen & Sabel, 2004). These services can also be used for real-time monitoring of environmental conditions such as air quality and notification of individuals suffering from health problems such as asthma.

The development of GIS on the WWW has broadened access to geographic data and GIS analyses because users do not necessarily need to have GIS software and databases resident on their own computers to access and display the information in a GIS application (Cromley, Cromley, & Ye, 2004). Other chapters in this book reference a number of websites where maps, text, and databases of relevance to public health GIS are distributed. The greater access to health

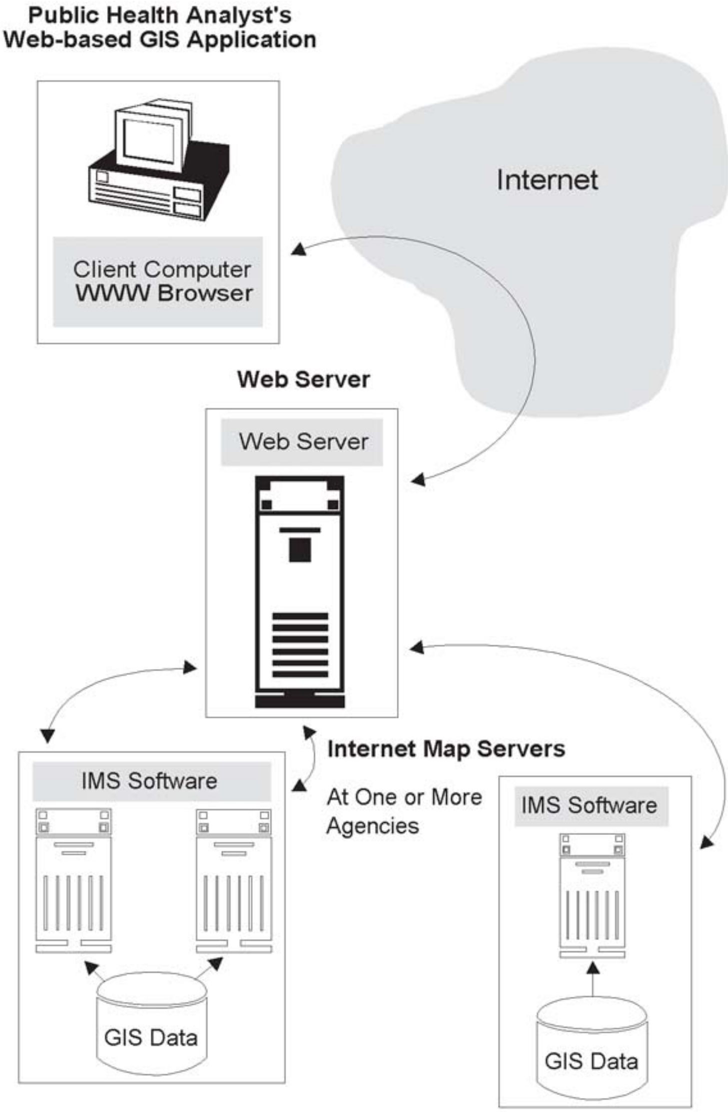


FIGURE 1.11. A client–server model for distributing geographic information over the Internet. This model is capable of providing real-time access to geographic data served by more than one agency and combined in one GIS application.

information made possible by delivering GIS over the WWW poses particular challenges for public health agencies and analysts because of the confidential nature of personal health records and concerns over privacy. These issues are discussed in detail in Chapter 3.

Distributed GIS implementation is probably beyond the capabilities of most health researchers. It involves partnerships with information technology professionals and other organizations. Nevertheless, the pace of technological change affecting GIS system design is rapid, and it is worth following new developments in GIS to consider their possible application in health research and health promotion.

The technological advance that has had the most significant impact on GIS since the development of the desktop computer was the launch of Google Earth® in 2005 (Lipowicz, 2006). Several weeks later, Paul Rademacher, a Dreamworks animation programmer, looked at the Google Earth source code written in JavaScript and created a demo that enabled users to pin their own information (in this case, housing ads) to locations on the map (Ratliff, 2007). Since 2005, Google Earth has been downloaded more than 250 million times. Most recently, Google has launched Street View, which incorporates street-level photography into Google Maps® for a number of cities in the United States.

Millions of individuals, as local observers, are now creating and posting their own map data. ***KML*** (*Keyhole Markup Language*) makes it possible for anyone to annotate online maps (Google, 2011a). ***Mashups***, web applications that combine data from more than one source into a single integrated tool, have become commonplace, and mapping mashups integrating data from a content provider or other source with Google Maps are a major type of mashup (Cho, 2007).

Mashups related to health and medical care are beginning to appear (Cho, 2007). Vimo, an integrated portal allowing users to research and compare health care providers, products, and services including health insurance, was launched in 2006 (Vimo, 2011). HealthMap is a global disease alert map application using Google Maps and unstructured Internet media reports on disease outbreaks around the world (Freifeld, Mandl, Reis, & Brownstein, 2008).

In addition to the wide range of geographically referenced content and maps now available on the Internet, GIS tools themselves are moving to the Internet. There are a number of geocoding service sites on the web, for example, many enabling users to geocode addresses without charge. The explosion of user-created map content and GIS services has wide-ranging implications for the geospatial web.

Conclusion

GIS as a technology is not the province of any single organization or group of users. Public health professionals and epidemiologists are now an important part of the GIS community. As they develop the potential of GIS in their own applications, the technology itself continues to evolve in response to the needs of the many and varied groups within the GIS community (Sui, 2008).

GIS implementation—even in the web environment—involves organizing people to use a collection of computer hardware, software, and spatial databases to answer questions or solve problems (Kessler, 1992). A GIS application reflects—at least implicitly—how the GIS user views the world, including what are appropriate and meaningful ways to represent reality, what are the subjects of interest, and how and by whom information will be used. The technologies of GIS and the ideas they represent "are vitally embedded in broader transformations of science, society, and culture" (Pickles, 1995a, p. 3). At present, these include the changes in how human beings communicate with each other through the use of electronic technologies and what the impacts of these changes will be for access to information, privacy, and individual and collective decision making. All of these changes have affected the collection and analysis of data on human health problems.

GIS implementation requires a significant commitment of time, money, and effort by the individuals and organizations adopting the technology. While the adoption of GIS technology has been advocated by public health professionals, epidemiologists, and medical care providers who see the value of geographical analysis of human health problems, the commitment of these resources may not seem justified to public health professionals and epidemiologists whose program activities and research are not primarily concerned with spatial analysis. Even for these public health professionals, however, GIS literacy and some GIS capability have become necessary because GIS have become the technology for managing and distributing spatial databases and most public health databases are spatially referenced. To the extent that census agencies, environmental protection agencies, vital statistics bureaus and disease registries, and medical care providers have adopted GIS technology, analysts who wish to access their data even for nonspatial analyses will become part of the GIS community.

CHAPTER 2

Spatial Data

Spatial data are observations with explicit locations. For geographers—and most people interested in studying human health problems—the relevant space is the surface of the earth. Geospatial data are obtained by observation and measurement of events or objects referenced to their locations in that space. GIS implementation requires access to geographic data. "The database is the foundation of a GIS" (Worboys & Duckham, 2004, p. 35).

This chapter highlights some of the important attributes of geographic data that the GIS user must consider in developing an application. Because analysts use GIS to integrate data from different sources based on location, they need to have a clear understanding of the data layers they are combining. Data in different projections cannot be meaningfully integrated in a GIS without additional processing to bring the data layers into a common projection, and data at different scales cannot be meaningfully integrated in a GIS without generalizing to the smallest scale. Because it is possible to overlay so many layers of data in a GIS display, the symbolization of data elements that produces a legible display can be difficult to design.

In the first part of this chapter we consider the two broad models of geographic data, field and object, and how these models are expressed in tessellation and vector databases. Next, we review the basic concept of location and how it is determined, either to define the set of locations that provide the spatial framework for field data or to georeference objects on the surface of the earth. Because geographic data come from a variety of sources including existing maps, we also briefly describe and discuss the importance of scale, projection, and symbolization. We then consider the quality of geographic data, how it can be assessed, and its implications for GIS applications. The last sections of the chapter discuss the role of metadata in documenting the characteristics of spatial databases and supporting online search and discovery of information.

Field and Object Data

There are two main approaches to modeling geographic information. From one perspective, we can think of phenomena that are continuously distributed just above, on, or below the surface. Precipitation, surface elevation, and soil are examples of ***continuous data*** or ***field data***. It is possible to visit any location on the earth's surface and ask "what is the elevation here?" or "what is the environmental quality here?"

Because these phenomena are continuous and can be observed everywhere, they are usually measured based on observations made at a set of sample sites on the surface. This network of sites creates a spatial framework for describing the distribution of the phenomenon. ***Tessellation*** is the geometric process of partitioning an area into smaller units that do not overlap but completely fill the entire area (Arlinghaus, 1994). Squares, triangles, and hexagons can be used as the basic units of a tessellation (Figure 2.1). When the units are the same shape and size, the tessellation is regular. Of the three, the regular square tesselation has probably been the most widely used because it can be represented easily in an array structure with row and column numbers identifying particular squares. Furthermore, it is compatible with hardware devices used for capture of spatial data like remote sensing data or used for output like the inkjet printer (Peuquet, 1984). Unlike triangles and hexagons, squares can be subdivided into smaller units with the same shape, area, and orientation. Irregular tessellations in which the partitions do not have the same shape, size, or orientation can also be constructed.

In field-based GIS applications, the GIS database generally contains several layers to enable comparison and integration of the various fields of information. A public health analyst might want to look at land cover in relation to soils to find where medium- and low-density residential development relying on septic systems is occurring in relation to soils that pose problems for septic system functioning. Because the spatial framework represents a set of locations on the surface, the observations made at these locations represent a sample of the phenomenon being modeled. Thus sampling error becomes a consideration in the GIS application.

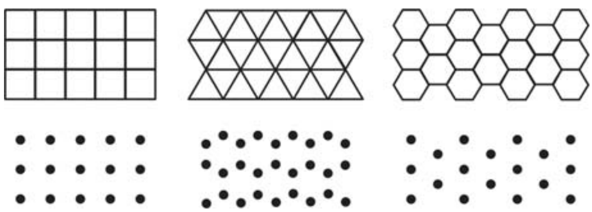


FIGURE 2.1. Regular square, triangular, and hexagonal tessellations. The centers of the tiles correspond to regular square, hexagonal, and triangular lattices, respectively.

From the other perspective, we can think of discrete objects that may be found on the earth's surface. A person, a hospital, and a public drinking water reservoir are examples of ***discrete data*** or ***object data***. Many health databases contain object data. ***Objects*** are entities that are identifiable, relevant to the particular public health problem at hand, and describable (Mattos, Meyer-Wegener, & Mitschang, 1993). It is possible to consider any object and ask "where is this object located?"

Each object of interest that can be located on the earth's surface has different types of attributes that are important for modeling purposes: textual/numeric, spatial, temporal, and graphical (Worboys & Duckham, 2004). A public drinking water reservoir, for example, will have a set of textual/numeric attributes indicating its name, surface area, and other attributes (Figure 2.2). The polygon representing the shoreline or perimeter of the reservoir surface is a spatial attribute. Attributes like the time the reservoir was developed as a public drinking water supply or the time it was first recorded in the database are temporal attributes. The symbol used to represent the reservoir as an object in cartographic displays is the graphical attribute of the underlying object. In this case, a solid-filled polygon symbol of a particular hue is used to indicate that the polygon is a reservoir. If the reservoir had been represented using a line to show the shoreline, it would not be possible to "fill" the area corresponding to the reservoir's surface. It would only be possible to change the style, thickness, and color of the line.

Tesselation and Vector Data Models

The field and object views are expressed in the two main data models used to implement GIS: tesselation and vector. These data models have important implications for the storage and processing of data (Worboys & Duckham, 2004). The

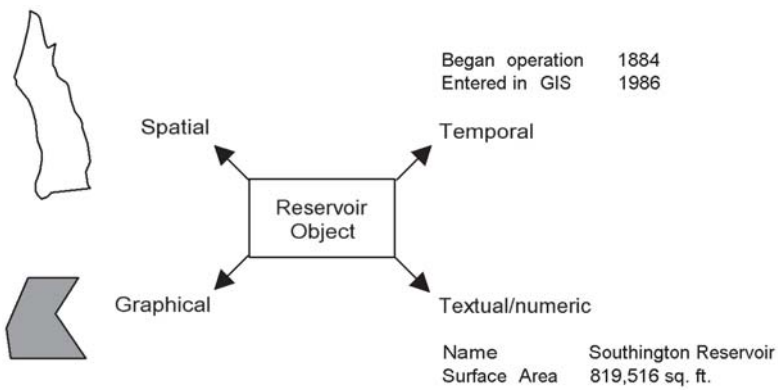


FIGURE 2.2. Attributes of a public drinking water reservoir as a spatially referenced object in a GIS.

most commonly used tessellation model is the raster model based on a regular tessellation. In a ***raster model***, every raster layer corresponds to a single measurement of a continuous field for a unit of space—for example, elevation or land cover at a particular place.

Stored in the computer, raster data are organized as an array of cells corresponding to particular places that are called ***pixels***—shortened from "picture elements" (Arlinghaus, 1994). A raster cell is typically assigned a real number associated with the measurement of the phenomenon at that location. Proprietary raster data storage formats may also use ***grids*** to represent nominal data coded as integers (Price, 2008, p. 524). These nominal data values often result from classification of the real number values in raster cells, for example, when a reflectance value is classified to represent a particular category of land cover in processing remote sensing data.

The GIS user defines the size of each pixel in terms of its area on the earth's surface to match the available data (Figure 2.3). A common ground dimension for remote sensing data in the United States is 30 meters × 30 meters, the pixel size of Landsat Thematic Mapper (TM) data (Jensen, 2005). The size of the pixel affects the size of an object that can be discerned in a digital image, thus determining the spatial ***resolution*** of the data. Some landscape features in an urban environment such as a single-family detached residence cannot be discerned in

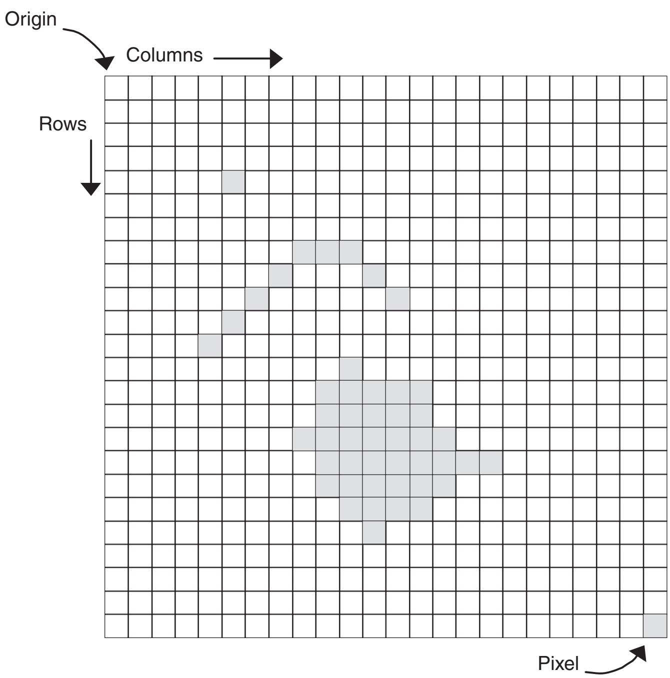


FIGURE 2.3. A raster data model.

low-resolution raster images like the 30-meter TM data because the features are smaller than the pixel size. On the other hand, such features would be discernible on a high-resolution image that uses a small pixel size. Low resolution is useful for large geographic areas where only limited detail can be represented.

The GIS user can define the geographic extent of the data by specifying the number of rows and columns of pixels in the data layer. The location of each individual pixel in the raster is given by its particular row and column numbers. Raster data structures make it easy to overlay data layers as long as pixel sizes are the same and corresponding pixels are registered to exactly the same position on the surface of the earth.

Perhaps the most commonly used irregular tessellation is the ***triangulated irregular network*** (TIN). A data value like elevation is observed at a set of sampling points on the surface. These points form the vertices of triangles in the TIN. Each triangle in the TIN connects three neighboring points so that the plane of the triangle approximates the surface between the points and calculating the slope and aspect of the surface is straightforward (Figure 2.4).

In a ***vector model***, every vector layer corresponds to a single class of objects that have the same dimensionality (point, line, or area), although data layers of different dimensionality can be used in a vector GIS application. A ***vector*** is a finite straight-line segment that can be described by the locational coordinates of its endpoints. In a vector data structure, a point such as a hospital location would be represented by a single ordered pair (x,y), a line or an arc such as a meals-

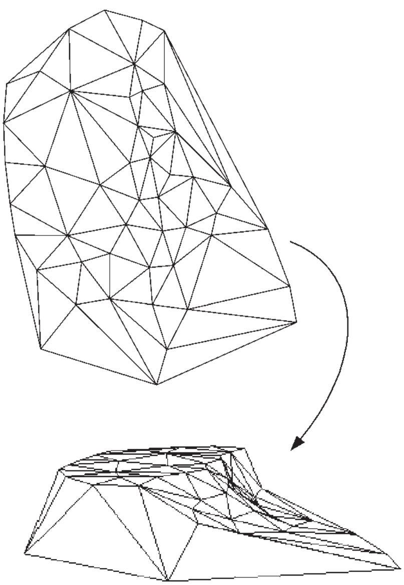


FIGURE 2.4. A triangulated irregular network (TIN) and the corresponding surface representation.

on-wheels route by a sequence of straight-line segments, and an area or polygon such as a health service district by the vectors enclosing it. Every vector layer corresponds to a set of objects located in space and described by many attributes (Figure 2.5).

Vector data may or may not be topological. A vector database is ***topological*** if it contains information on the neighborhood relationships among objects (Figure 2.6). Specifying a "start node" and an "end node" for each arc indicates the direction of the arc. Areas to the left and to the right of a particular directed arc can then be identified. The U.S. Census Bureau's GBF/DIME file was a milestone in the development of topological data structures for digital spatial data (Broome & Godwin, 2003).

Network data are a type of vector data that model space as a set of connected links and nodes (Rigaux, Scholl, & Voisard, 2002). Network databases consist of nodes and arcs (Figure 2.7). A ***node*** is a point of connection for two or more arcs or an endpoint of an arc. An ***arc*** is a link that connects two nodes. The nodes and the arcs of the network comprise the entire space of interest and locations of interest exist only on the network. In a ***planar network***, a node exists whenever two arcs intersect. In a ***nonplanar network***, arcs may cross each other without resulting in a node where arcs are connected. Ground transportation networks with overpasses and underpasses are examples of nonplanar networks.

GIS designed to model transportation systems (Miller & Shaw, 2001) and public utilities use network databases. A table stores the topological relationships among nodes and arcs in the network and a measure of impedance for each arc. The impedance value measures the effort associated with moving from one endpoint to the other endpoint along the arc. In transportation networks, the length of the arc (travel distance), the time required to travel the arc, or the cost associated with traveling the arc such as a toll are commonly used impedance measures.

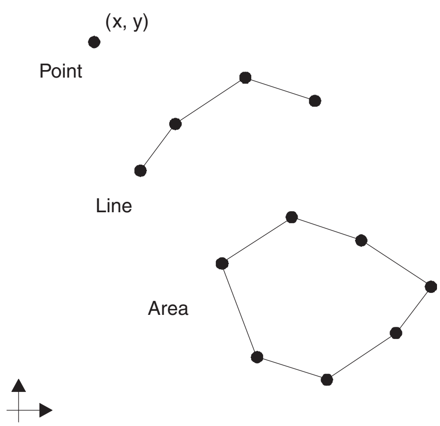
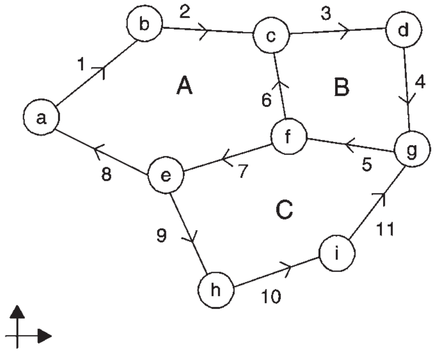


FIGURE 2.5. A vector data model.



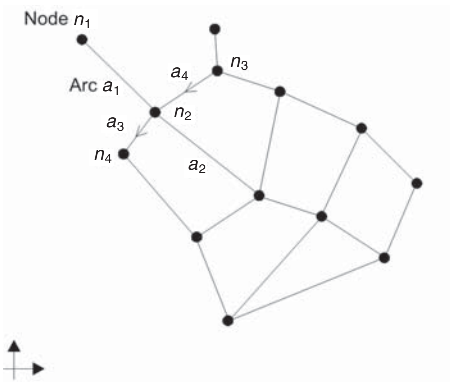
Directed Arcs, Nodes, and Areas

Arc	Start Node	End Node	Left Area	Right Area
1	a	b	X	A
2	b	c	X	A
3	c	d	X	B
4	d	g	X	B
5	g	f	C	B
6	f	c	A	B
7	f	e	C	A
8	e	a	X	A
9	e	h	C	X
10	h	i	C	X
11	i	g	C	X

FIGURE 2.6. A topological vector data model.

Transportation network databases often include a turn table and a reference address table. The ***turn table*** includes a tuple or row (see Chapter 1) for each direction of travel through an intersection from one segment to another segment. An ***impedance measure*** is associated with each "turn" indicating whether or not it is possible to move from one street segment to another in a particular direction. Turn and arc impedances provide some measure of the travel cost associated with travel in a particular direction. A ***reference address table*** stores information about address ranges for street segments.

Vector data that are not topological are sometimes referred to as "spaghetti" data (Rigaux, Scholl, & Voisard, 2002). In ***spaghetti data***, lines and areas are independent features, like strands of spaghetti on a plate, and need not correspond to any actual spatial object. For example, a GIS analyst could digitize several lines capturing distinct parts of the boundary of Connecticut. There may be enough strands on the plate or lines to represent the boundary, but there is no information on how to connect these strands. Intersections or connections between the strands of spaghetti are not explicitly modeled. GIS functions make it possible to create line and polygon topology for vector data that are not topological.



Turn Table

From Arc	To Arc	Impedance
a_1	a_2	1
a_1	a_3	2
a_1	a_4	-1
.	.	.
.	.	.

FIGURE 2.7. A network data model. The turn table shows that it is very easy to go straight from Arc a_1 to Arc a_2 (no traffic sign or signal). It is more difficult to turn right from Arc a_1 to Arc a_3 (the driver must slow down to make the turn). Because Arc a_3 and Arc a_4 are one way, the turn from Arc a_1 to Arc a_4 cannot be made.

In a vector system, "resolution" again refers to the smallest feature that can be discerned. The minimum length of a line object, the minimum separation required to display objects as separate and distinct, and the minimum size below which a long narrow area will be represented as a line and a small area will be represented as a point are affected by the scale of the database (Veregin & Hargitai, 1995). The extent of a vector data layer is described by a ***bounding rectangle*** within which all of the objects in the database can be placed.

It is possible to represent a range of spatial phenomena—field and object—with various data models and to convert from raster to vector and vice versa (Figure 2.8). Raster/vector conversion is a function commonly found in GIS software systems. It is also possible to think of regular square, triangular, or hexagonal tessellations in terms of the square, hexagonal, or triangular lattices formed, respectively, by their center points (Figure 2.1) (Gold, 1990; Boots, 1999).

The choice of data model is more than just a technical data management issue. Spatial statisticians recognize that observations associated with point locations may have a range of meanings and require different analytical methods (Cressie, 1993). ***Geostatistical data*** are measured for a sample of points in a fixed area, and ***lattice data*** are measured for a fixed collection of points or area centers

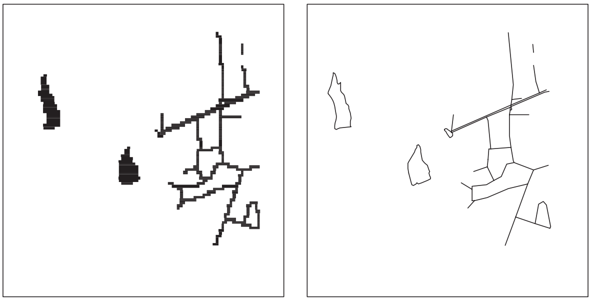


FIGURE 2.8. A raster database and a vector database representing the same situation of three reservoirs and an adjoining network of water distribution mains.

(Haining, 2003). In a geostatistical database, the point locations are considered to be a sample of the set of all points in an area, and a value for a continuously distributed phenomenon is measured at each point. In a lattice database, the point locations are fixed points or centers of areas, and the observation describes the value of a phenomenon at or aggregated to the discrete location. Lattice data are supplemented with information about the neighborhood structure or spatial arrangement of the points used in the analysis of the data (Anselin, 1999).

However represented in the GIS, the field and object views of data are recognized as inverse ways of looking at geographic information (Peuquet, 1984; Worboys & Duckham, 2004). The field view starts with a spatial framework over which the geographical domain of an attribute like precipitation or environmental quality is represented. The object view starts with objects that can be "embedded" in an otherwise empty space based on their locations. Both views require us to operationalize location.

Measuring Location

Location means position in space. Location is the basis for integrating geographic data in a GIS. ***Absolute location*** refers to position with respect to an arbitrary grid system like the geographic grid of parallels and meridians. Absolute location gives the position of a point so that its unique position on the earth is clear. "The burner stack is 41°48′N and 72°15′W" is a statement of absolute location. ***Relative location*** refers to position with respect to other objects in the geographic space. "The burner stack is 300′ northwest of the intersection of Park and Broad Streets" is a statement of relative location.

Positional data in a GIS can come from several sources; these are discussed in detail in Chapter 3. Geodetic, photogrammetric, and digital image processing data are all primary sources for positional data because positions are determined by direct or indirect measurement of the earth's surface (Drummond, 1995). These primary sources of positional data provide an essential and accurate foundation for other spatial data. Maps and archival documents like gazetteers are secondary sources for positional data.

Geodesy is the science of observation and measurement of the shape, size, and dimensions of the earth as a whole and of the earth's surface (Arlinghaus, 1994). Position can be measured directly from the earth's surface by surveying (Moffitt & Bossler, 1998). With the development of satellite technology, the Global Positioning System (GPS) is now used to ascertain position and for primary data capture of spatial data. The GPS was designed and built and is operated and maintained by the U.S. Department of Defense (Kaplan & Hegarty, 2006). The primary purpose of GPS was to aid navigation, not to record locations per se. In 1978, the first operational GPS satellite was launched. By the mid-1990s, the system was fully operational with the required 24 satellites.

The system today has several components: 27 satellites orbiting earth at an altitude of 12,500 miles, five monitoring stations, and the receivers that individuals use to determine their locations. The satellites send continuous radio signals and receive correctional data from monitoring stations. Signals are received by GPS units held at different locations on the earth. The receiver measures how long it takes the signal to travel from the satellites. By measuring the distance from three or more satellites, the location of the receiver can be obtained by triangulation. Signals from at least three satellites are needed to obtain horizontal positions (lon/lat) at a particular location; a signal from a fourth satellite is needed to obtain vertical position.

Until now, GPS has been the only system of its kind, although the former Soviet Union developed the GLONASS system that was operational in 1995. GLONASS was affected by the collapse of the Soviet Union and the economic situation in Russia. It is operated for the Russian government by the Russian Space Forces. In December 2005, India and Russia agreed that India would share the development costs of the newest series of satellites and launch them from India. GLONASS was fully restored with 24 operating satellites in September 2010. Also in December 2005, the European Union launched the first satellite in the Galileo system. Full deployment of the Galileo system is not expected until 2014. The Chinese are also moving toward developing a system called BeiDou.

Using GPS, a field researcher can obtain accurate locations for a variety of features in the landscape. Many GPS devices allow storage of the coordinates of the measured points, and these data can then be loaded directly into the GIS. Using the GPS has become a standard method for surveying. Individuals and organizations requiring highly accurate positional data usually hire a surveyor or GPS technician to make the necessary measurements. Small, inexpensive GPS receivers capable of storing up to 1 megabyte of GPS data with 2.5-meter horizontal accuracy and of downloading the data to a computer via the ***Universal Serial Bus*** (USB) port are making it easier to develop an in-house capability to obtain GPS data and to integrate it with other spatial data in a GIS.

Surveying and GPS involve collecting data directly from the surface of the earth. ***Remote sensing*** is the analysis and interpretation of data gathered by means that do not require direct contact with the subject. Aerial photography is a method of photographing the earth's surface from an aerial platform (Jensen, 2007). When the photo is taken directly above the surface in the image, the vertical photograph is known as an ***orthophotograph***. Remote sensing images like air photos taken at an angle are ***oblique***. ***Photogrammetry*** is the measurement of aerial photographs to determine locations for mapping.

Digital image processing of geographic data relies on satellites with sensors capable of detecting electromagnetic energy reflected or emitted from objects on the earth's surface (Jensen, 2005). The energy detected is converted into a data value for a specific location and transmitted to receiving stations either directly or via tracking data and relay satellites. The data are then enhanced for viewing and subsequent analysis by using digital image processing algorithms. It is possible to obtain a variety of data from a single flight. Acquisition of photogrammetric and digital image data usually involves purchasing a database from a government or quasi-governmental agency that has the means to produce these data.

In many GIS applications, positions are estimated from an existing map of the earth's surface created at a particular scale. For example, we could estimate the coordinates of a hospital by digitizing from an appropriately annotated topographic map. ***Digitizing*** requires the use of a tablet and cursor to record Cartesian coordinate locations of map features from a map placed on the digitizing tablet or the use of a cursor to screen digitize from a visual display. We could also estimate the coordinates by using a GIS to address-match geocode the hospital's address against a digital, address-ranged street network database. ***Geocoding*** is "the process by which an entity on the earth's surface, a household, for example, is given a label identifying its location with respect to some common point or frame of reference" (Goodchild, 1984, p. 33). Address-match geocoding as a GIS function is described in detail in Chapter 3.

Finally, coordinates for particular places are published in gazetteers and other archival sources in both paper and digital formats (Abate, 1991; Armstrong, 1995; U.S. Geological Survey, 2011; Ordnance Survey, 1999; Ordnance Survey, 2011a). Many of these gazetteers are available online. When using secondary sources of positional data, the GIS analyst should read the database documentation to understand the primary source of positional information for the published data and its accuracy. The scale and projection of the digital model or map and the symbols used to represent objects affect the quality and accuracy of the positional data obtained from it.

Scale, Projection, and Symbols of Cartographic Data Sources

As models, maps are generalized representations of reality. Maps distort reality by simplifying the complex, three-dimensional surface of the earth for representation on a flat sheet of paper or video screen. The "cartographic paradox" is that "to present a useful and truthful picture, an accurate map must tell white lies" (Monmonier, 1996, p. 1). Scale, projection, and symbolization are three basic components of maps. Each is a potential source of distortion. Public health analysts using GIS need to understand these concepts and how to identify scale, projection, and symbols by reading the details of the description printed on a paper map or the data description supplied with a digital spatial database.

Scale

Map scale tells the user how much smaller the map is than the reality it represents. Map scale can be stated as a ratio, a simple bar graph, or a phrase (Figure 2.9). A ratio scale of 1:5,000 means that one unit on the map represents 5,000 units on the ground. Ratio scales are particularly useful for comparison. A 1:5,000 map is a ***large-scale map***, a map of a small area showing high detail. In contrast, a 1:1,000,000 map is a ***small-scale map***, a map of a large area showing limited detail. The graphical scale is useful for representing scale on paper and in digital displays because the scale will be correct even if the map is enlarged or reduced during reproduction.

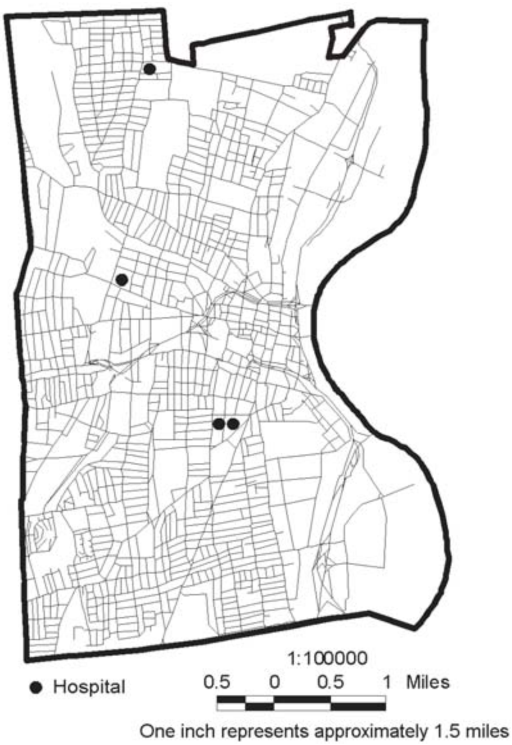


FIGURE 2.9. A map of hospital locations showing different methods for representing map scale.

Scale is an important attribute of maps because scale affects the amount of detail that can be captured and represented. At the map scale in Figure 2.9, the locations of hospitals in Hartford, Connecticut, and the street network can be represented. When the map is compiled at a smaller scale to depict the state, the street network cannot be shown clearly. Similarly, meanders or curves in the Connecticut River are generalized as the scale decreases and areas like the town of Hartford shrink to points and eventually disappear. It is important to understand that simply enlarging a small-scale map or zooming in the visual display of a spatial database does not enable us to see features that are not present in the database. Enlargement by zooming often aids visualization when features are clustered together, but it does not affect the level of generalization of the data.

When data layers are at different scales, the larger scale data should be recompiled to match the scale of the smallest scale data layer. This is accomplished through generalization of the larger scale data layers using techniques like line generalization. It is possible to develop GIS applications that call on different *scale* databases for the same region; in this case more detail can be observed as the user shifts through to higher scale data sets. Although scale can be constant at all points and in all directions on a globe as a true scale model of the earth, scale varies on a paper map because map projection stretches some distances and shrinks others.

Projection

Projection, a second basic component of maps and spatial databases, refers to the mathematical function that transforms locations from the curved, three-dimensional surface of the earth to a flat, two-dimensional representation (Pearson, 1990; Maling, 1992; Iliffe, 2000). There is more than one system for describing positions of places such as environmental monitoring sites or objects such as hospitals. The network of meridians and parallels, which form a geographic grid, is used to reference locations on the earth's surface.

Meridians are true north–south lines connecting the poles. Each is half of a great circle, a circle created by passing a plane through the center of the earth. They are spaced wide at the equator and converge at the poles. ***Parallels*** are true east–west lines. They intersect meridians at right angles. The equator is the only parallel that is a great circle. ***Longitude*** is the position of a place east or west of the prime meridian, which passes through Greenwich, England. It is measured as the arc of the parallel between that place and the prime meridian. Longitude ranges from 0° to 180°E or 180°W. ***Latitude*** is the position of a place north or south of the equator. It is measured as the arc of the meridian between the place of interest and the equator. It ranges from 0° to 90°N or 90°S. A meridian is all points having the same longitude; a parallel is all points having the same latitude.

Lon/lat can be reported in degrees, minutes, and seconds (DMS), with 60 minutes in a degree and 60 seconds in minute. Computer processing of lon/lat reported in degrees, minutes, and seconds and by direction poses some problems. For computer analysis of lon/lat data, the preferred measurement is the radian, a unit defined so that there are 2π radians to a circle, with one radian equal to approximately 57.296°. The computer functions used to generate trigonometric function values are based on radians and not on degrees. Most users of spatial data find it convenient to manage the conversion from degrees to radians by converting their data to decimal degrees (DD). This enables storage of the lon or lat coordinate as a single real number rather than as separate integer values for degrees, minutes, and seconds. But it also retains relevance to the geographic grid so that users can easily conceptualize point locations and reference objects to paper maps. The GIS software converts the decimal degree lon or lat to its radian equivalent internally.

In addition, the computer cannot easily recognize E and W or N and S as directions. Instead, + and – are used to indicate directions, with quadrants I, II, III, and IV organized around the origin as in a Cartesian coordinate system (Figure 2.10). Thus, the coordinates for the burner stack would be found in a digital spatial database not as 41°48'N 72°15'W but as (–72.25, 41.80), so that the longitude (east–west or "x" in Cartesian space) is given first (i.e., lon/lat rather than lat/lon), values are in decimal form, and directions are correct. For most places in North America, longitude will be negative but latitude will be positive.

Because meridians converge at the poles and are spread apart at the equator, the distance between two meridians is not constant over the range of lati-

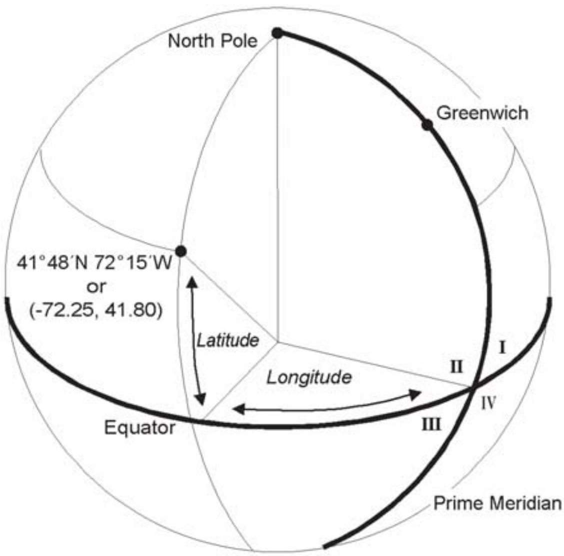


FIGURE 2.10. The geographic grid. Positive and negative values of lon/lat in decimal degrees are based on "quadrants" I, II, III, and IV organized around the origin as in a Cartesian coordinate system.

tude. At the equator, the distance between two meridians 1° apart is 69.17 statute miles. At 60° it is 34.67 statute miles. Similarly, because the earth flattens near the poles, the distance between two parallels is not constant. At the equator, the distance between two parallels 1° apart is 68.71 statute miles. At 60° it is 69.23 statute miles. The geographic grid is not a planar grid like the Cartesian coordinate system, and distances between points on the geographic grid should not be calculated using the Euclidean distance metric. Chapter 9 describes different measures for calculating distance.

Longitude and latitude defined with reference to a spheroid are ***geodetic coordinates*** (Iliffe, 2000); some sources also refer to these as ***geographic coordinates***. A ***spheroid*** is an approximation of the shape of the earth as a sphere flattened at the poles. The combination of shape and size of the earth, as given by the spheroid, with a fixed position used as a point of origin defines a ***datum***. There are global, regional, and local datums used in a range of geographic information technologies. Countries may use the same spheroid for mapping, but they are on different datums if the systems have different points of origin.

Because the surface of the earth is curved and not flat, lon/lat represents location in a three-dimensional space. Direct plotting of lon/lat coordinates results in an image that does not match what would be observed on the earth or a globe (Figure 2.11). Map projection provides a method for making the transformation from three dimensions to two. Most paper maps and many digital spatial databases represent projected spatial data. Map projection may be a more serious issue for users of small-scale maps and projected digital spatial databases than for users of large-scale paper maps. Distortion will not be as great on large-scale maps because the areas being mapped are relatively small, covering only a small portion of the earth's curved surface. Because one of the most important capabilities of GIS is integrating spatial data, however, a basic understanding of map scale and projection and their implications for geodata processing is essen-

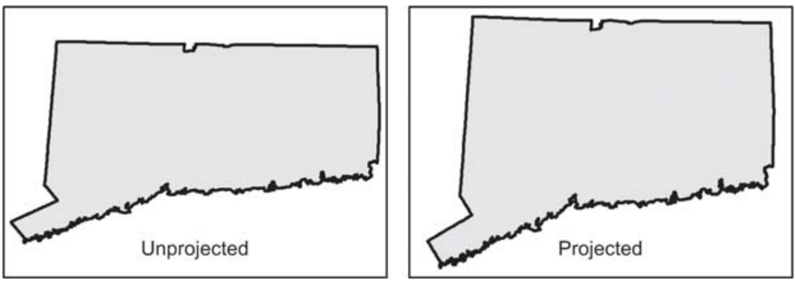


FIGURE 2.11. The importance of map projection is demonstrated by displaying a view of the unprojected boundary of a study area, the state of Connecticut, and a view of the projected boundary. At the latitude of the study area (around 41°N), a direct plot of latitude against longitude results in considerable distortion of the study area size and shape in the east–west dimension compared to boundary data projected in state plane coordinates.

tial. Databases cannot be properly overlayed or integrated if they are not in the same projection. Projecting spatial data from lon/lat or from one projection to another is a function commonly found in GIS software.

Three plotting surfaces have been used to develop practical map projections (Figure 2.12). Map projections based on these objects preserve different spatial properties of objects and their relationships. On a perfect map, areas on the map would be in correct proportion to corresponding areas on the earth, distances on the map would be true scale, directions and angles on the map would remain true to the corresponding directions and angles on the earth, and shapes of objects on the map would not be distorted. Not all of these properties can be achieved when we move from three to two dimensions. One way of classifying map projections is by the relationship they preserve. ***Conformal*** projections preserve shapes but not areas of land masses. ***Equal*** area projections preserve areas but may distort shapes. ***Cylindrical*** projections like the Mercator projection preserve true directions and angles but not shapes and areas.

The usefulness of the conformal projections is evident in the widely used State Plane Coordinate System in the United States. The ***State Plane Coordinate System*** provides a convenient means of locating mapping positions on a two-dimensional plane. The system is based on a rectangular grid defined for each state of the United States (Figure 2.13). The grids permit the methods of plane surveying to be extended over great distances at high precision.

In the continental United States, two conformal map projections are used for the state plane coordinate systems (Figure 2.14): the transverse Mercator secant and the Lambert conformal with two standard parallels (Pearson, 1990). The transverse Mercator secant projection defines a grid zone roughly 158 miles east–west (Figure 2.15). There are two true-scale lines running north and south in each transverse Mercator zone. Between these lines of secancy, the distances are less than true scale. Outside of these lines, the distances are greater than true scale. The distortion increases with increasing distance east or west from the secant lines. Within the zone, distortion is not a function of latitude, so north–

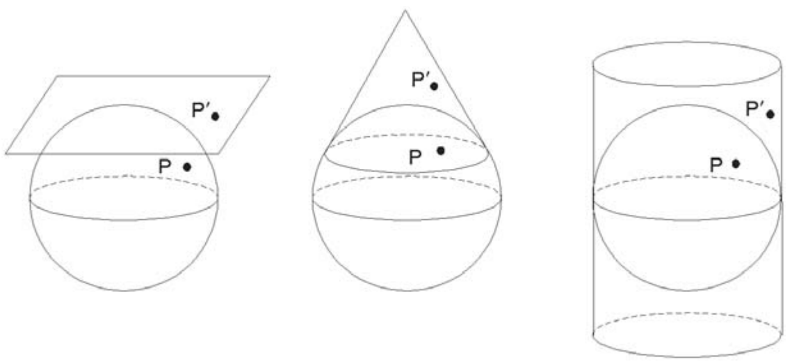


FIGURE 2.12. Three plotting surfaces used to develop practical map projections.

FIGURE 2.13. The State Plane Coordinate System of 1983 zones. Each zone is defined by an origin and a projection system, either a central meridian and scale factor if transverse Mercator or standard parallels if Lambert. Zones of the Universal Transverse Mercator projection system differ from the zones of other state plane transverse Mercator projections by only the zone-defining constants used. Adapted from Stern (1989).

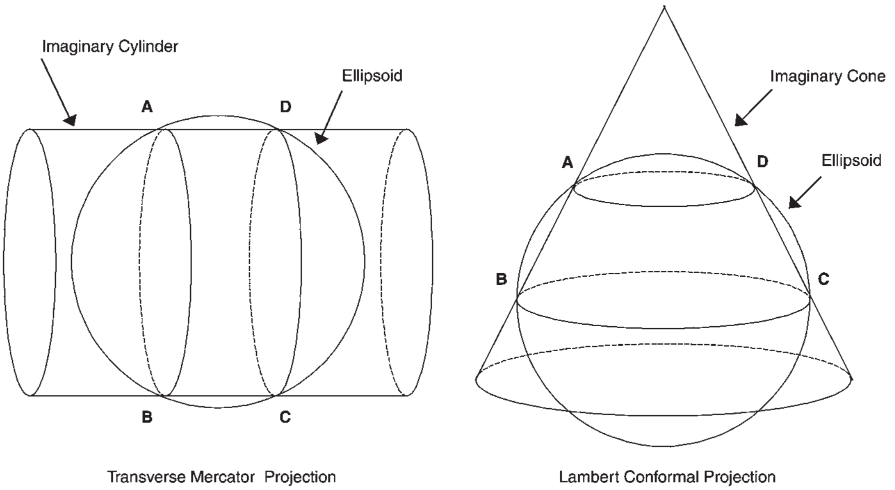


FIGURE 2.14. Surfaces used in state plane coordinate systems. The transverse Mercator secant projection provides the closest fit to the datum surface for a rectangular (**ABCD**) zone greatest in north–south extent. The Lambert projection provides the closest approximation to the datum surface for a rectangular (**ABCD**) zone greatest in east–west extent.

south extension is unlimited. The transverse Mercator projection is well suited for states of major north-to-south extent like Illinois.

The Lambert conformal projection with two standard parallels defines a grid zone roughly 158 miles north–south (Figure 2.15). The parallels are two true-scale lines running east and west in each zone. Again, between the true-scale lines, distances are less than true scale. North and south of the lines, the distances become more than true scale. Within the zone, distortion is not a function of longitude, so east–west extension is unlimited. The Lambert conformal is suited for states of major east–west extent like Tennessee. For each state plane grid, an origin is established to the south and west of all points so that the coordinates can be given as a ***false easting*** and a ***false northing*** (Figure 2.16). This means that none of the coordinates will have negative signs. Distance units in state plane coordinate systems are feet or meters depending on the datum plane and the data distribution practices in various states.

The planes produced by these projections are like tiles approximating parts of a sphere. Because they represent data on a plane, these systems support spatial analysis like the measurement of distance based in Euclidean geometry. Different states in the United States rely on different state plane coordinate systems for their paper and digital spatial databases (Warnecke, Johnson, Marshall, & Brown, 1992). A number of documents are available that describe the specifics of the systems used in each state (Snyder, 1987; Stern, 1989). To make a map of the entire state in a state with more than one state plane zone (Figure 2.13), analysts

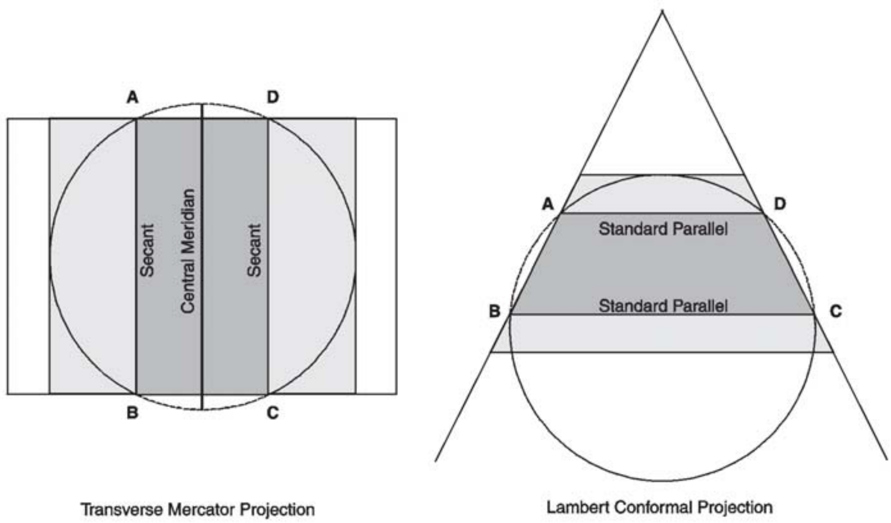


FIGURE 2.15. Scale relationships in State Plane Coordinate System projections. Along the standard lines of the projection (secants for transverse Mercator or standard parallels for Lambert conformal), scale is exact. Between the standard lines, distances are less than true scale.

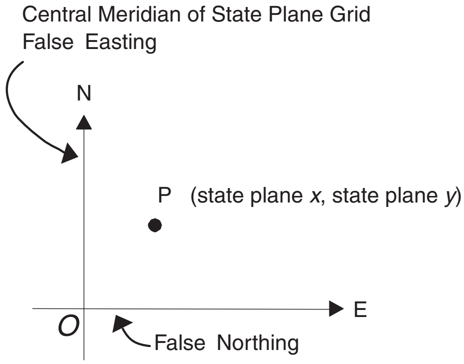


FIGURE 2.16. State Plane Coordinate System geometry. The central meridian is assigned a false easting. The origin *O* for measuring the state plane coordinates is located where the false easting intersects a false northing. The location of the origin *O* forces the eastings (state plane *x* coordinate values) and northings (state plane *y* coordinate values) to be positive numbers. The coordinates of point P are both positive.

select one of the zones and project data for the entire state according to that zone or they select some other map projection.

In addition to the various projections used in data layers available at the local level, other map projections may have been used in the various map series published by national governments (Parry & Perkins, 1987; Böhme, 1993). This situation poses problems for public health analysts who need to integrate data from various sources across jurisdictions. In order to use the GIS software functions to project or to change the projection of a digital spatial database, the analyst needs to know what projection the data are in to start. GIS software allows users to *display* data in different projections without actually reprojecting the data, but GIS spatial analytic procedures may require data layers to have a common projection or spatial reference. Techniques for integrating data layers are discussed in greater detail in Chapter 3, which describes foundation databases for public health GIS and how health data can be linked to them.

Symbolization

The visualization and mapping functions of GIS require data objects to be represented with some kind of graphical symbol. Bertin (1979) identified six dimensions of visual variability of map symbols: size, shape, value, texture, orientation, and hue (Figure 2.17). These aspects of symbolization can be and are manipulated to achieve certain objectives in cartographic communication (Monmonier, 1996). The range of symbols supported will vary from system to system, depending on software and hardware configurations. Cartographic design is discussed in greater detail in Chapter 4. Standard cartography texts provide useful guidelines for map compilation and design. One body of cartographic research eval-

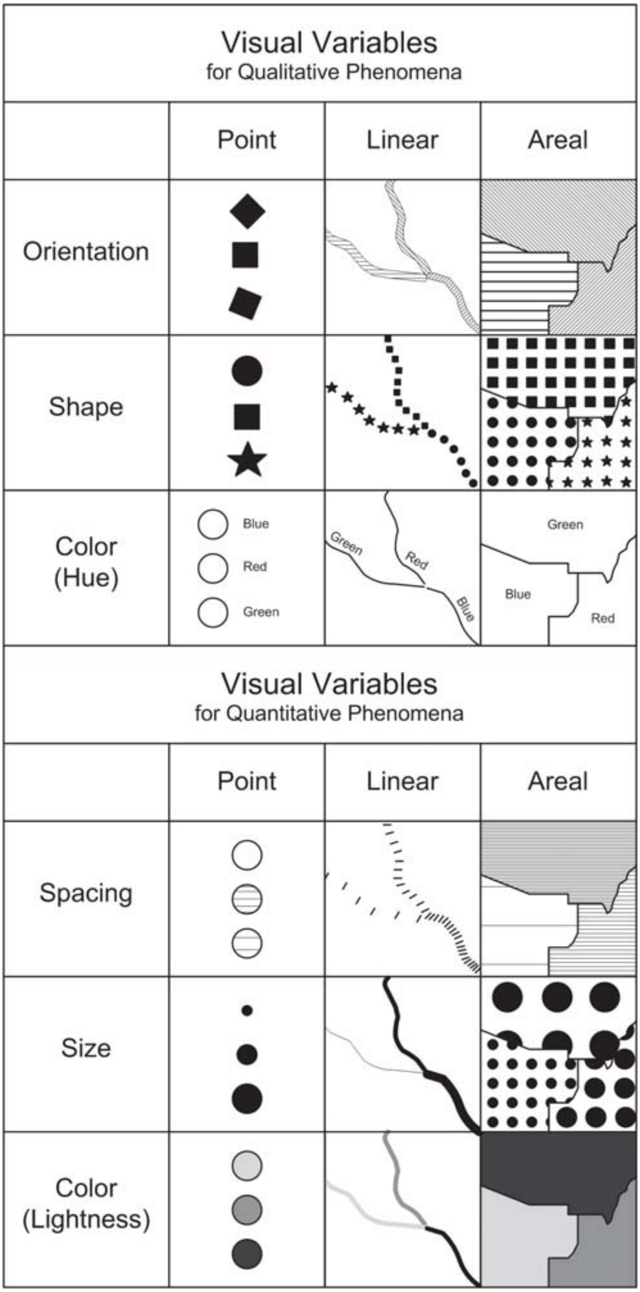


FIGURE 2.17. Visual variability of map symbols. Adapted from Slocum, Terry A., McMaster, Robert B., Kessler, Fritz C., Howard, Hugh H., *Thematic cartography and geovisualization, 3rd Edition*, © 2009, pp. 82, 83. Adapted by permission of Pearson Education, Inc., Upper Saddle River, NJ.

uates the impact of different symbolizations on the perceptions of map users (MacEachern, 1994; Dykes, MacEachern, & Kraak, 2005).

Geographic Data Quality

Because the database is the foundation of any GIS, the quality of the geographic data that goes into the system is paramount. The United States established the National Committee on Digital Cartographic Data Standards in 1982 (Morrison, 1995). Its draft report identified five important aspects of spatial data quality (Moellering, 1987); these dimensions have since been accorded a degree of international consensus (Moellering, 1991; Servigne, Lesage, & Libourel, 2006).

Lineage is a "description of the source material from which the data were derived, and the methods of derivation, including all transformations involved in producing the final digital files" (National Institute of Standards and Technology, 1994, p. 21). To an extent, lineage is not so much a measure of data quality as the information needed to assess data quality based on other factors. The contents of a lineage describe data at various stages in its existence. Development of standards for lineages is an ongoing process. The difficulty of developing useful lineages creates a problem for data suppliers, although the importance of lineages to spatial data quality assessment cannot be denied.

Accuracy, in general, refers to the level of error present in a database. An "accurate" database is one that is free from error. Because spatial databases contain both locational and thematic data (e.g., pixel and elevation, lon/lat and health outcome), users must be concerned with both positional accuracy and attribute accuracy.

Positional accuracy refers to the nearness of the values describing the position of a real-world object to the object's "true" position. Positional error may be introduced at the initial measurement of location. Analysts should pay particular attention to the precision (the number of significant digits) with which geographic coordinates are measured and reported. Rounding the lon/lat of a place from (-72.24952,41.80443) to a coordinate pair with fewer decimal places (-72.25,41.80) affects position. A second source of error is the chain of processing between the initial measurement or observation and its final "resting place" in a GIS database (Drummond, 1995). Because GIS analysis involves manipulations of databases like projection change and overlay, errors propagate.

The preferred test of positional accuracy is comparison to a data source of higher accuracy (Antenucci, Brown, Croswell, Kevany, & Archer, 1991). These tests are often made at the time of data capture, for example, when using a GPS receiver or digitizing or scanning an existing database. In this approach, positional accuracy is measured by comparing the final positional information to a known higher standard, perhaps a set of geodetically or photogrammetrically observed checkpoints. The standard database contains checkpoints whose (x,y) coordinate values are considered "true" coordinates, having been determined using a measurement system of higher quality. For each i of the n checkpoints available, the difference or error measured between the "true," or accurate coordinate (x_{ia},y_{ia}) and the corresponding database coordinate (x_{id},y_{id}) is recorded. The values are used to determine the ***root mean square error*** (RSME) as follows:

$$RMSE = \sqrt{\frac{\left(\sum_{i=1}^{n}\left[(x_{id} - x_{ia})^2 + (y_{id} - y_{ia})^2\right]\right)}{n}}$$

GIS users are familiar with an aspect of RMSE in the digitizing process. At the start of the digitizing session, a set of control points whose coordinates are known in the map's projection grid system are marked. From these control points, the transformation constants used to convert all digitized coordinates to the projection coordinates are produced. The projection coordinates of selected control points can then be compared to their known coordinate values, and the differences used in the calculation of RMSE. In many GIS systems, the user will be presented with this value after digitizing and asked either to accept it or to repeat the process. Public health analysts can, then, be able to assess positional accuracy of spatial databases in the early stages of developing a GIS application, either from the description of positional accuracy accompanying a database acquired from another source or during the digitizing process.

Attribute accuracy is an aspect of data quality that considers the nearness of the values describing the real-world entity in the database to the entity's "true" attributes. In many public health applications of GIS, the attribute data for objects will already have been collected in a disease registry or surveillance system. The amount of information available about uncertainty or error in these attribute data will vary depending on whether the agency collecting the data has carried out and described procedures for determining the level of error in the data.

In public health GIS applications, consistent definitions of what constitutes a health event or a health service are needed to ensure attribute accuracy. As discussed in detail in Chapter 3, it is not always easy to define what is meant by a "case"; moreover, case definitions may change over time. Attributes of cases, like race, ethnicity, or International Classification of Disease (ICD) diagnosis, also need to be coded consistently to meet standards for attribute accuracy.

When the attributes are measured as interval/ratio level data, the normal or Gaussian distribution and log normal distribution may be appropriate models of the relative frequencies of error. Sometimes these distributions are truncated to reflect maximum or minimum possible values. When the attributes are nominal, places or objects have been assigned to simple classes, and differences, means, and standard deviations are not meaningful. Uncertainty may exist when the place or object is assigned to the wrong class, when two or more analysts do not agree on the assignment, or, in the case of remotely sensed data, when the class does not agree with the class assigned based on direct observation.

As with positional accuracy, attribute accuracy assessment relies on comparison with a source of higher accuracy. In remote sensing, for example, a sample of locations is selected, and the class assigned in the digital image processing is compared to the directly observed "ground truth." The results can then be tabulated in a misclassification matrix where the diagonal represents correct classifications and the off-diagonal elements record incorrect classifications (Table 2.1). A number of indices can be calculated to summarize the overall level of accuracy (Congalton, 1991).

A major issue in the assessment of attribute accuracy is the degree to which errors are spatially systematic. ***Spatial dependence*** measured as ***spatial autocorrelation*** means that there is a correlation between errors for features located near each other. In health outcome data, for example, there may be geographically systematic variations in surveillance or diagnosis. These "errors" are not necessarily connected to positional errors in the data. "Unfortunately, we know very little about the spatial structure of uncertainty in geographic data" (Goodchild, 1995a, p. 76). Research investigating both the spatial structure of data errors and methods of visualizing uncertainty is ongoing.

Completeness as a measure of data quality refers to "the relationship between the objects represented in the data set and the abstract universe of all such objects" (Brassel, Bucher, Stephan, & Vckovski, 1995, p. 82). Two types of completeness can be considered. First, are all of the relevant objects captured in the database? Health outcome databases derived from voluntary versus mandatory screening programs might have widely varying degrees of completeness in representing the true universe of all individuals with a particular health problem. Second, are all the records for an individual unit in the database complete? Completeness as a measure of data quality addresses presence and absence of data for the specified universe.

Finally, ***logical consistency*** is a measure of data quality that considers the "structural integrity" of a database (Kainz, 1995). A spatial database is logically consistent when it is compatible with the attribute data and when it complies

TABLE 2.1. An Example Error Matrix for Classification of Areas Based on Land Cover

Data classification	Ground truth classification				
	Deciduous	Conifer	Shrub	Barren	Row total
Deciduous	82	4	16	18	120
Conifer	6	23	6	5	40
Shrub	3	8	97	2	110
Barren	0	3	9	87	99
Column total	91	38	128	112	369

Overall accuracy 289/369 = 78%

with the requirements of the selected raster or vector data model. The health analyst would want to make sure, for example, that all of the health events in a database were modeled in the same way—as fields or as objects—and described by the same set of attributes. Consistency rules prevent invalid changes to the database and ensure consistency in data handling—for example, treatment of missing values—throughout a series of transformations.

Most GIS software systems can test for topological consistency in a spatial database. In a vector database, this would include checks for missing nodes, dangling arcs, duplicate centroids, and other inconsistencies in the node–arc–area relationships depicted in Figure 2.6. For a particular GIS database, the consistency checks would verify that each directed arc or line segment has exactly one start and one end node, that each node is a start or end node of at least one directed arc, that each area is bounded by one or more directed arc, and so on.

Logical inconsistencies present in a database may not prevent the health analyst from producing a graphic display or map of the areas, but would almost certainly affect any spatial analysis performed using the database. GIS applications that involve more spatial data analysis require higher levels of consistency in spatial databases. The topological consistency checks can be performed when spatial databases are created so that errors can be detected before the public health analyst incorporates the database into an application.

A GIS application involving multiple layers of data may meet the test of logical consistency for every layer but lack consistency across data layers. Data fields in vital statistics databases may change over time as fields are added or modified. Records from an earlier time period would not, therefore, be logically consistent with records from the later reporting period unless the earlier records were modified to reflect the new format with the added or modified fields. Two GIS databases collected at different scales or using different projection systems may be logically consistent as individual databases, but would cause serious problems if they were integrated without appropriate modification. Methods for performing logical consistency checks across data layers are still being researched and have generally not been incorporated into GIS software packages.

A final important issue in assessing the quality of a spatial database is the handling of ***temporal*** information (Guptill, 1995). Because GIS applications involve assembling data from many different sources, information about the temporal attributes of the data is extremely important. Even when each particular data layer represents the most current information available, the layers may not mesh temporally because census data, land use data, and data on other elements in the universe of geographic features are not updated on the same schedules.

An important approach to handling temporal information is to model it as an attribute (Table 2.2). Both positional information and attribute information can be assigned temporal attributes describing the date when the data were observed and the date when the position or attribute "expired." In Table 2.2, a public drinking water well changed from "Active" to "Inactive" status as a source of drinking water on October 13, 1997. The importance of historical information, as distinct from information just recording change as it occurs, is obvious in the

TABLE 2.2. Example of Temporal Description Attributes for a Public Drinking Water Well

Basic feature	Spatial object		Attributes	
Well	Well ID:	101	System Value:	Manchester
	Feature observed:	*07/01/1989*	*Value observed:*	*07/01/1989*
	Feature expired:	*Current*	*Value expired:*	*Current*
			Status value:	Active
			Value observed:	*07/01/1989*
			Value expired:	*10/13/1997*
			Value:	Inactive
			Value observed:	*10/13/1997*
			Value expired:	*Current*

case of process studies. Does the researcher need to reconstruct past patterns of land use or public drinking water supply as part of an epidemiologic investigation?

Aside from the quality of the geographic data itself, there are additional desirable properties for digital geographic databases (Worboys, 1995). Secure databases prevent unauthorized access or allow different levels of access. Security is particularly important for many databases containing health records for individuals. The issue of confidentiality of health data and its implications for mapping is discussed in greater detail in Chapter 7. Reliability of a database (as opposed to reliability of measurement) means that systems and data will be up and accessible when users need information. Finally, use of the data is facilitated when technological change is transparent to the user. "Technology-proof" systems insulate users from the technical aspects of the database system so that databases are not required to change with each new technological advance in hardware and software.

The Role of Metadata

The characteristics of a spatial database can be described in ***metadata***, "data about data" (Green & Bossomaier, 2002, p. 95). Metadata provide information in four key areas:

- Availability of data including information that makes it possible to search for and discover the data.
- Fitness for use of data including information on special features of the data to allow analysts to assess the appropriateness of data for a given use.

- Access to data including information on how to acquire the data.
- Transfer information including technical specifications for data handing.

Metadata Standards for Geospatial Data

The Federal Geographic Data Committee (FGDC) developed metadata standards for spatial databases in the United States, and producers of digital spatial databases have been expected to prepare metadata that complies with these standards (Federal Geographic Data Committee, 1998). The FGDC standard was designed as a ***content standard***. This means that the standard requires certain kinds of information to be included in the metadata but does not govern the format in which the content is presented. According to the standard, there are several categories of information covered in a metadata file for a digital spatial database (Table 2.3). Other countries have also developed national metadata standards for digital spatial data (Green & Bossomaier, 2002).

International standards for digital geospatial metadata have also been developed. In 1995, the Geographic Information Technical Committee (TC 211) of the International Organization for Standardization (ISO) set out to develop a metadata standard for geographic information. This effort built on the work of the FGDC, whose standard was recognized by the American National Standards Institute (ANSI) under the auspices of the InterNational Committee for Information Technology Standards (INCITS). INCITS Technical Committee L1 serves as the U.S. advisory group to ISO TC 211 (Hill, 2006). ISO 19115, the ISO's content standard for geographic information metadata, was published in 2003 (Table 2.4) (International Organization for Standardization, 2003). In addition, ISO 19139 created an XML schema that prescribes the format of the metadata record. These standards are part of the general family of ISO 19100 standards dealing with various aspects of geographic information and transfer (Larsgaard, 2005).

Georeferences in Metadata

Metadata are the key to search and discovery of information using the Internet. The effort to develop digital libraries of spatial data in initiatives like the Alexandria Project brought together experts from the fields of geographic information science, library science, information science, and museum informatics (National Research Council Panel on Distributed Geolibraries, 1999; Hill, 2006) An outgrowth of these efforts was the search for a unified georeferencing approach that would not limit searching by location to place name-based referencing. In its broadest sense, ***georeferencing*** is "relating information to geographic location" (Hill, 2006, p. 1). The growth of the World Wide Web as a tool supporting the distribution of and search for information has led to an interest in using geographic references as a way of drawing together all kinds of information—not just geospatial data—based on location.

TABLE 2.3. FGDC Content Standard for Digital Geospatial Metadata

Metadata content area	Content	Mandatory
Identification information	Basic information about the data set including citation, description, time period of content, status, spatial domain, keywords, and access and use constraints	Yes
Data quality information	General assessment of the quality of the data set including attribute accuracy, logical consistency, completeness, and positional accuracy	As applicable
Spatial data organization information	Representation of spatial information in the data set including direct spatial reference method and point and vector object information or raster object information	As applicable
Spatial reference information	Description of the reference frame for and means of encoding coordinates in the data set including the horizontal and vertical coordinate system definitions	As applicable
Entity and attribute information	Information content of the data set including the entity types, their attributes, and attribute domains	As applicable
Distribution information	Basic information about the distributor of the data set and options for obtaining it	As applicable
Metadata reference information	Description of the metadata information including metadata date, standard, metadata access and use constraints, and identification of responsible party for metadata preparation	Yes

The ***Dublin Core Metadata Initiative*** (DCMI) began with a workshop held in Dublin, Ohio, in 1995 that brought together experts on web-authoring tools and HTML and library science (Dublin Core Metadata Initiative, 2011). DCMI promotes widespread adoption of interoperable metadata standards and develops specialized metadata vocabularies to facilitate the finding, sharing, and management of information. Subscribers from more than 50 countries participate in the organization. The Dublin Core Metadata Element Set used to describe resources has 15 elements, and spatial information is included in the coverage element. There are four recommended ways for describing a place in a DCMI metadata entry: by the ISO 3166 code for the representation of country names, by a listing from the Getty Thesaurus of Geographic Names, by a description of a point, or by the description of a box (Table 2.5).

TABLE 2.4. ISO 19115 Metadata Core Elements

Metadata element	Mandatory	Metadata element	Mandatory
Data set title	Yes	Spatial representation type	
Data set reference date	Yes	Reference system	
Data set responsible party		Lineage statement	
Geographic location		Online resource	
Data set language	Yes	Metadata file identifier	
Data set character set		Metadata standard name	
Data set topic category[a]	Yes	Metadata standard version	
Spatial resolution		Metadata language	
Abstract	Yes	Metadata character set	
Distribution information		Metadata point of contact	Yes
Additional extent information (Vertical, temporal)		Metadata date stamp	Yes

[a]ISO 19115 identifies these topic categories: farming, biota, boundaries, climatologyMeteorologyAtmosphere, economy, elevation, environment, geoscientificInformation, health, imageryBaseMapsEarthCover, intelligenceMilitary, inlandWaters, location, oceans, planningCadastre, society, structure, transportation, utilitiesCommunication. Topics may be entered with the truncation and capitalization shown as theme keywords in FGDC metadata records to ensure compliance with ISO standards.

Among producers of information for the web, there has been growing use of ***geotagging***, adding geographical identifiers to HTML meta elements for various media types including webpages, ***RSS*** (Really Simple Syndication) feeds used to publish frequently updated digital content like news feeds, blogs, or podcasts, and images like map images and photographs. One or more meta elements in HTML can be nested inside head elements and can include information that browsers use to find information about the web content. Web developers have used a variety of formats for geotags (Ruiz, 2005; Bausch & Bumgardner, 2006).

Other types of spatial referencing approaches are tied to national grid systems. The ***National Grid*** in Great Britain systematically breaks down the area covering the region into progressively smaller squares identified by letters and numbers (Ordnance Survey, 2011b). The largest unit is 500 kilometers, and it takes only four of these squares to cover Great Britain. The smallest grid is 1 kilometer. By estimating positions within a cell of this grid, it is possible to develop a grid reference accurate to 100 meters on the ground. The National Grid is based on the Universal Transverse Mercator projection, and grid lines can be applied to all Ordnance Survey maps of Great Britain at all scales.

The ***U.S. National Grid***, also based on the Universal Transverse Mercator, is a similar system that can be used to create addresses in the United States

TABLE 2.5. Dublin Core Metadata Elements and Spatial Coverage Examples

Element	Content	Example
Contributor	An entity responsible for making contributions to the resource.	U.S. Geological Survey National Mapping Program

Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant. Temporal period may be a named period, date, or date range.	Spatial:
		ISO 3166 Code: US
		Getty Thesaurus: Connecticut
		Dublin Core Point:
		east = 1006384.471918;
		north = 787698.776589
		units = feet;
		projection = Connecticut State Plane Coordinate System NAD83 Feet
		Dublin Core Box:
		northlimit = 944279.125000;
		westlimit = 730512.187500;
		eastlimit = 1263094.375000;
		southlimit = 554854.687500;
		units = feet;
		projection = Connecticut State Plane Coordinate System NAD83 Feet
		Temporal:
		1968–1994

Creator	An entity primarily responsible for making the resource.	State of Connecticut, Department of Environmental Protection
Date	A point or period of time associated with an event in the life cycle of the resource.	2005 edition
Description	An account of the resource in the form of an abstract, graphical representation, or similar information.	Town is a 1:24,000-scale, polygon and line feature-based layer that includes state, county, and town (municipal) boundary features.
Format	The file format, physical medium, or dimensions of the resource.	Shapefile 604Kb distributed online at *www.ct.gov/dep/gis* with Zip compression 1.35Mb
Identifier	An unambiguous reference to the resource within a given context.	The name Town identifies the data sets in the GIS data sets published by the Connecticut Department of Environmental Protection.
Language	A language of the resource.	en [English]
Publisher	The entity responsible for making the resource available.	State of Connecticut, Department of Environmental Protection

(cont.)

TABLE 2.5. *(cont.)*

Element	Content	Example
Relation	A related resource.	Connecticut Department of Environmental Protection GIS Base Map data sets
Rights	Information about rights held in and over the resource.	No restrictions or legal prerequisites for using the data. The data are in the public domain.
Source	The resource from which the described resource is derived.	USGS 7.5-minute topographic quadrangle maps for the state of Connecticut
Subject	The topic of the resource.	Connecticut—Administrative and political divisions—Maps
Title	A name given to the resource.	Town
Type	The nature or genre of the resource.	Data set

(Terry, 2004). GIS software makes it possible to overlay grid lines on layouts in a mapping application (Figure 2.18). Like the National Grid in Great Britain, the U.S. National Grid has its origins in military mapping. Problems encountered in responding to Hurricane Katrina have provided an impetus to use the U.S. National Grid in disaster and emergency planning and response, but its use is not yet widespread.

Another system that is used to describe the locations of places in selected states in the United States is the ***Public Land Survey System*** (PLSS). The PLSS does not cover the New England states, coastal states from New York to Georgia, Kentucky or Tennessee, Texas, or Hawaii. The system is a rectangular survey system in which regions are divided into 6-mile square townships subdivided into 1-mile square sections. Township designations indicate location north or south of a baseline, and range designations indicate location east or west of a principal meridian. The full legal description of a property includes the state, principal meridian, township and range designations with directions, and section number (*nationalatlas.gov*, 2010). PLSS designations have a wide range of uses, including describing the locations of events relevant to public health. The California Pesticide Use Reporting program described in Chapter 6 uses PLSS designations to identify the locations where pesticides have been used.

Geographic information technologies have made it easier to assign a range of spatial identifiers to describe digital spatial databases and to georeference a range of media. These identifiers are used in metadata in a variety of ways, and metadata standards continue to evolve. The FGDC, DCMI, and other organizations involved with metadata have developed software and training materials

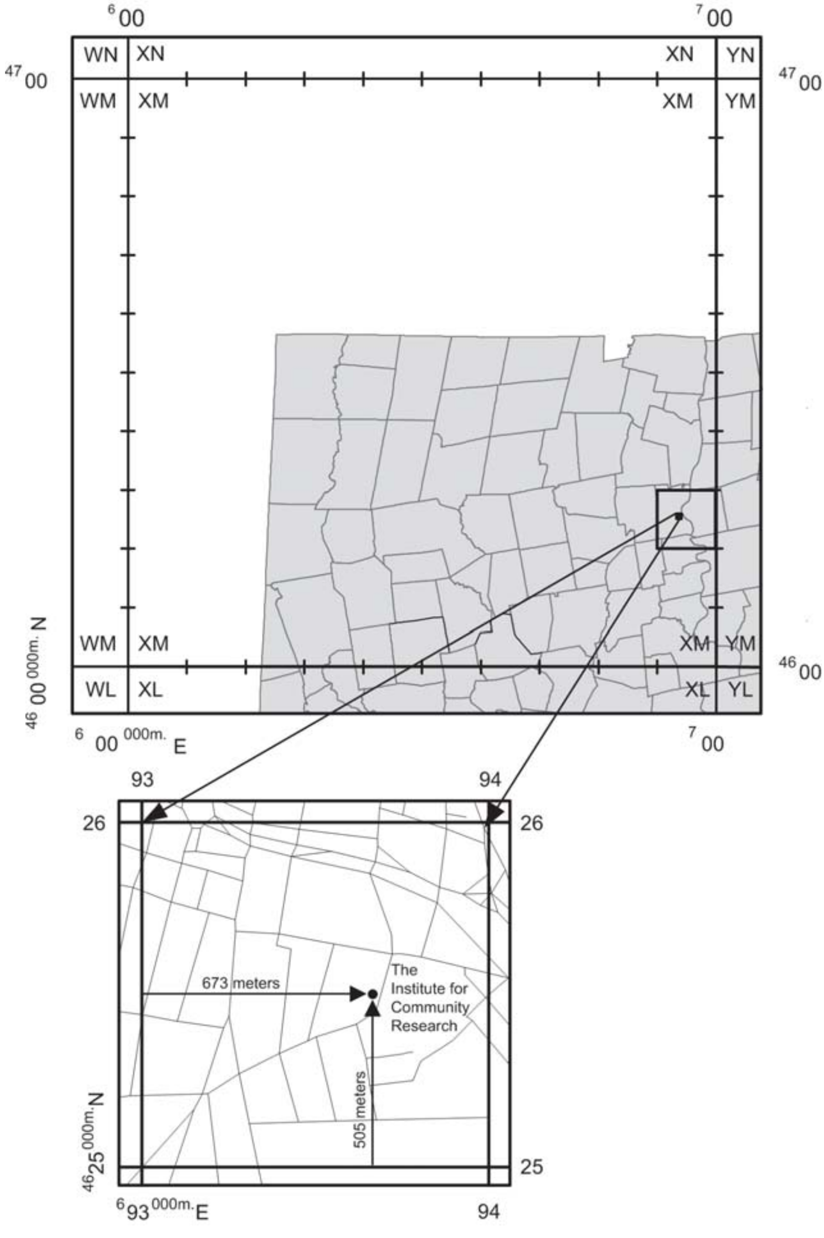


FIGURE 2.18. The U.S. National Grid location for the Institute for Community Research in Hartford, Connecticut, is 18TXM9367325505. The Institute is located in Grid 9325 in the 18T UTM grid zone in the XM 100,000 meter square. Measuring right from grid line 93 another 673 meters and up from grid line 25 another 505 meters results in the full 10-digit designation. This designation locates a point to an area about the size of a parking space.

to assist data producers in preparing metadata files that meet standards. GIS software packages now provide support for metadata preparation in a variety of formats.

Conclusion

Many public health professionals and epidemiologists analyze databases that contain some geographic information, such as a residential location. Use of these data in a GIS requires re-creating the database as a spatial database in a format that the GIS software can recognize. As GIS technology has developed, concern for the accuracy of spatial databases has grown and efforts have been made to develop standards for describing spatial databases that enable users to search for and access data, assess the appropriateness and accuracy of the data, and decide whether the spatial database can be used to answer the questions the user is asking.

GIS implementation involves creation, transformation, and analysis of potentially many spatial databases. In most GIS application areas, however, there are a number of "foundation" databases commonly used. The next chapter describes these databases and how they have been used in public health and epidemiological research.

CHAPTER 3

Spatial Databases for Public Health

Spatial data sets are fundamental components of GIS. The success of health-related GIS projects depends critically on having access to accurate, timely, and compatible spatial data. For organizations embarking on GIS projects, spatial data can be viewed as both a cost and a resource. Developing spatial data sets is expensive; it is estimated that well over half the cost of GIS projects goes to database creation, updating, and improvement. Yet, database development is also an investment that creates long-term value for organizations and the people they serve. Spatial data sets are often useful for addressing a wide range of policy and planning issues. Their value extends well beyond the scope of the original projects for which they were created, and it increases as the data sets are used.

This chapter describes the major types of spatial databases for public health GIS. We begin by discussing the concept of foundation data and summarizing major types of foundation data sets like geodetic control, digital orthorectified imagery, and address-ranged street network data files. We then consider the diverse types of population and health data sets that can be incorporated in GIS by geocoding data on individual health events or by joining aggregated population and health data for areas to spatial databases for mapping and analysis. The final sections examine issues related to spatial data integration and sharing.

Foundation Spatial Data

In generating spatial databases for public health GIS, the key linkage among data layers is the spatial linkage. Layers are tied together by their common geographical location. If a house is located a quarter-mile east of a park and adjacent to a hospital, these features should appear in the same relative positions in a GIS that connects data layers of houses, parks, and hospitals. In a GIS, we cannot link the locations of features directly to their positions on the earth since we are working at a scale much smaller than the earth. Therefore, spatial data layers must be connected to a foundation that makes spatial integration and linkage possible. Foundation data provide a geographical frame of reference to which other data layers are tied.

Foundation spatial data are "the minimal directly observable or recordable data to which other data are spatially referenced" (National Academy of Sciences, 1995, p. 16). We use the term here to apply to the spatial data layer to which other data layers are linked in a public health GIS project. As in constructing a building, the foundation supports the other data layers and defines the ***footprint***, or geographic extent, of the GIS database. Many different types of spatial data can serve as foundation data, for example, digital imagery from aerial photographs or satellites, street centerline data, or property boundary data. These databases differ in their scale, resolution, degree of positional accuracy, and ease and cost of use. The choice of a foundation data set will be influenced by the scale of the analysis. A study of health problems at the neighborhood scale requires foundation data at an equivalent spatial scale.

This section explores the various types of foundation data and their characteristics. National mapping agencies guide the development of foundation data in different countries of the world. The discussion that follows focuses on geodetic control and foundation databases widely used in the United States. The International Cartographic Association website provides an interactive map to access contact information for national mapping agency members and also offers national reports of mapping activities (International Cartographic Association/Association Cartographique International, 2010). The African Geo Information Research Network (AGIRN), an initiative of the Human Sciences Research Council of South Africa and EIS-Africa, maintains a site with links to national mapping agencies in Africa (African Geo Information Research Network, 2011).

Geodetic Control

Geodetic control is a system for registering location information to a set of well-defined points on the earth's surface. It includes a set of ***survey monuments*** on the ground and a ***reference datum*** that gives geographic coordinates for those monuments based on our knowledge of the size and shape of the earth, as discussed in Chapter 2. The reference datum is a key feature of geodetic control. In North America, the currently accepted reference datum is the North American Datum, 1983 (NAD-83). This datum is linked to the World Geodetic System, 1984, a geodetic control system for geographical coordinate use worldwide. The reference datum for North America has changed in recent years. For decades, the reference datum was the North American Datum of 1927 (NAD-27), replaced by NAD-83 after its publication in 1986. Spatial databases that were created in the United States, Canada, and Mexico before the mid-1980s often use NAD-27.

In developing GIS databases, it is critically important that all data layers use the same reference datum. Longitude/latitude coordinates based on NAD-27 and NAD-83 can differ by up to 100 meters in the lower 48 states, leading to positional errors and inconsistencies (Keating, 1993). When linking different spatial data layers, analysts should check the reference datums associated with each data set and, if necessary, convert all data sets to a common datum. Most GIS include commands for converting among NAD-27, NAD-83, and other common reference datums.

In most GIS applications, geodetic control is not used directly as a foundation data layer. Geodetic control is transparent, never displayed or connected with attribute information. However, understanding geodetic control and reference datums is vital for developing GIS data sets and ensuring consistent, accurate data linkage. In addition, the growing use of GPS receivers for generating coordinates heightens the importance of geodetic control because GPS coordinates are directly tied to geodetic control. Furthermore, online systems like Google Earth® that are used for mapping incorporate imagery that is tied to specific reference datums.

Digital Orthorectified Imagery

Digital orthorectified imagery (DOI) comprises pictures of the earth's surface that show the locations of features like roads, coastlines, and buildings. The pictures are raster images generated from aerial photography or satellite data. Digital images are encoded records of spectral reflectance or emittance intensity for objects or areas. Sensors on satellites record energy reflected from the earth's surface for different wavelengths or "bands" of the electromagnetic spectrum. For each band, an individual pixel corresponding to a place on the earth's surface has a digital number representing the intensity of spectral reflectance.

Image files are generally very large and difficult to store. Compression reduces the size of the image file. ***Lossless compression***, as the name implies, results in a compressed image that can be reconstructed to produce an image identical to the original. Its main advantage is the ability to reconstruct the original image. Its main disadvantage is limited compression ratio. Wavelet compression is a ***lossy compression*** method, which means that some information is lost in order to achieve higher compression rates. The compressed image cannot be used to reconstruct the original image. A wavelet compression method commonly used with geographic imagery is ***MrSID (Multiresolution Seamless Image Database)*** (LizardTech, 2004). ***JPEG 2000*** is another wavelet compression technique used with geospatial imagery and many other types of images (Taubman & Marcellin, 2002).

Tied to geodetic control to permit matching with other spatial data layers, the images have the geometric properties of a map. The information necessary to make this tie may be stored in separate so-called world files for use with images in MrSID or JPEG 2000 format. The Open Geospatial Consortium has also created a metadata standard for georeferencing JPEG 2000 images with embedded Geography Markup Language (Open Geospatial Consortium, 2011b). Similarly, ***GeoTIFF*** image metadata allows georeferencing information to be embedded in a TIFF file so that the image displays properly when added to a GIS application. The GeoTIFF format is being adopted by a wide range of data providers including the U.S. Geological Survey, SPOT Image Corporation, and other agencies in the United States and other countries (Ruth, 2010).

DOI does not incorporate specific feature or attribute information: it simply provides an image of some part of the earth's surface. Identifying and recording features on the images requires image interpretation, field checking, or linkage with an attribute-based spatial data layer for the area. However, many significant landscape features are clearly visible on DOI.

An important kind of DOI for public health GIS is the ***digital orthophoto-quarterquad*** (DOQQ). A DOQQ covers a "quarterquad," an area roughly 4 miles × 4 miles, at 1:12,000 scale. Produced by the U.S. Geological Survey in conjunction with other federal agencies, the DOQQs depict roads, houses, trees, and other detailed features (Figure 3.1). With their high resolution and high degree of positional accuracy, DOQQs form a useful foundation data layer for localized, large-scale public health assessments, such as mapping individual exposures to environmental contaminants. Other data layers can be matched to the DOQQs for detailed mapping and analysis.

The ***National Agriculture Imagery Program (NAIP)***, which began in 2003, is a source for high-resolution aerial photography imagery acquired during the

FIGURE 3.1. A portion of a digital orthophotoquad for the area around downtown Hartford. The dark area running north–south just east of the center of the view is the Connecticut River. The town boundary between Hartford to the west and East Hartford is the center of the river.

agricultural growing seasons for the continental United States (U.S. Department of Agriculture, 2010). NAIP imagery is acquired at one-meter ground sample distance. The imagery shows "leaf-on" conditions with no more than 10% cloud cover per quarterquad tile. Images correspond to the USGS quadrangles and are distributed in GeoTIFF format. The program makes imagery available to government agencies and the public within a year of acquisition. Many public agencies and private entities at every level are using these data for mapping, land classification, environmental monitoring, and a wide range of other activities including public health and safety (U.S. Department of Agriculture, 2008).

Smaller scale DOI includes ***satellite imagery*** from systems like SPOT and Thematic Mapper. Satellite images typically cover scales ranging from 1:50,000 to 1:100,000 at positional accuracies ranging from ± 25 meters to ±70 meters (Keating, 1993). Although scale and positional accuracy vary widely across satellite imagery, generally the images show major features such as roads, rivers, fields, and water bodies (Figure 3.2). As in other forms of imagery, features are not labeled or identified. However, methods for digital image interpretation that

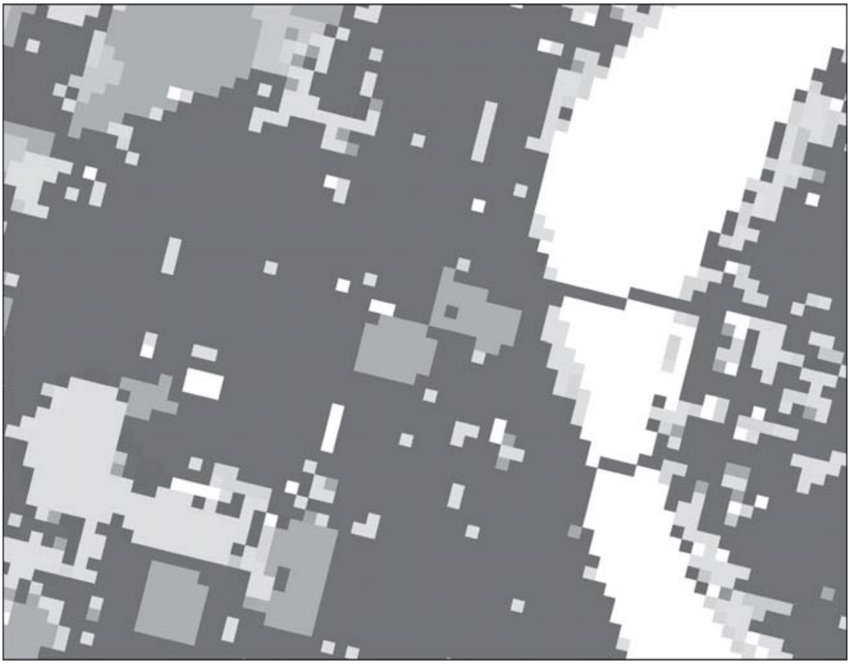


FIGURE 3.2. A portion of the land cover database for the area around downtown Hartford derived from Thematic Mapper Imagery. The areas shaded dark gray were classified as commercial/industrial/transportation. The areas shaded medium gray were classified as residential. The areas shaded light gray were classified as urban/recreational grasses. The areas shaded white were classified as open water. This figure shows roughly the same area as Figure 3.1.

distinguish land use/land cover features based on their distinct spectral characteristics are well developed and available in specialized computer software (Jensen, 2005). Some visible features in a satellite image vary seasonally because of changes in vegetation and precipitation. Cloud cover can also obscure features, complicating the interpretation of satellite images. In choosing a satellite image, the analyst should think through carefully the appropriate time of year for the image and the maximum allowable cloud cover. Detailed information is available for satellite imagery covering the United States to aid the analyst in selecting useful images (U.S. Geological Survey, 2010a).

Satellite images offer an important foundation data layer for regional-scale health analyses covering states or parts of states and for local analyses. The images have been widely used in displaying and analyzing land use, land cover, and natural resource patterns. In the public health field, the images have been utilized to analyze and predict outbreaks of vector-borne diseases such as Lyme disease (Glass et al., 1995; Ford et al., 2009).

Digital Line Graphs

Vector data also provide a foundation for regional-scale GIS development. ***Digital Line Graphs (DLGs)*** are vector databases that show transportation lines, water bodies, political boundaries, and elevation contour lines. Unlike imagery, DLGs include attribute information. Attribute codes describe the physical and cultural characteristics of points, lines, and areas on the DLG. DLGs are derived from the large- and intermediate-scale topographic maps created by the U.S. Geological Survey. They exist for all of the United States, excluding Alaska, at a scale of 1:100,000 (U.S. Geological Survey, 2010b). Large-scale DLGs, generated from the 7.5-minute topographic maps, have been created for many areas of the United States (Figure 3.3).

One concern in using DLGs is the accuracy and recency of attribute information. The sources of information for DLGs are topographic maps which may be years out of date. The Geological Survey has updated its topographic map series through a procedure known as "limited update," focusing on features that are most likely to have changed such as roads and hydrography (Lemen, 1999). DOQQs from aerial photography are the basis for limited update revisions. The efficient, limited-update procedure has generated more timely information for topographic maps and DLGs, but time lags, naturally, exist. For GIS, these issues are especially relevant in communities experiencing rapid population and commercial development where feature and attribute information changes frequently.

As the national topographic mapping program of the United States has developed, data in DLG format are being incorporated into a new generation of topographic maps and spatial data products built in collaboration with local and state agencies. For example, DLG data have been used in the creation of the National Hydography Dataset. This and other data layers are available as part of The National Map and its developing Digital Map program (U.S. Geological Sur-

FIGURE 3.3. A portion of the 1:24,000 digital line graph database for the area around downtown Hartford, including roads, hydrographic features, and town boundaries. This figure shows roughly the same area as Figure 3.1.

vey, 2010b). These new spatial data products will be, like DLGs, used in health applications of GIS in the future.

TIGER/Line® Data

Another form of vector foundation data, compiled at 1:100,000 scale, is TIGER/Line data. The ***Topologically Integrated Geographic Encoding and Referencing (TIGER)*** data set was developed for the 1990 census (Marx, 1986). Since 1990, the TIGER/Line database has evolved into the ***MAF/TIGER® (Master Address File/Topologically Integrated Geographic Encoding and Referencing)*** database, which is the Census Bureau's set of digital files storing all of the geographic and attribute data necessary to conduct the census. The MAF portion contains a record for each potential housing unit. The TIGER portion contains all of the points and lines identifying the features used to form the areas for which the Census Bureau tabulates data. TIGER/Line data are an extract of selected geographic and cartographic information from the MAF/TIGER database (U.S. Census Bureau, 2009a). MAF/TIGER also included a redesign of the original TIGER/Line files database (U.S. Census Bureau, 2005). Although earlier versions of TIGER/Line data were distributed in Vector Product Format and required special utilities to convert them to formats that could be used in GIS software, TIGER/Line data have been converted to shapefile format in preparation for the 2010 census (Table 3.1). Various versions of TIGER/Line data and technical documentation can be downloaded from the census website. TIGER/Line data have been used widely in population, health, political, and transportation mapping.

TIGER/Line shapefiles may contain landmark point features, line features including ***street centerline*** data and other line features like political boundaries or rivers forming boundaries of census areas, or area features including states, counties, and census tracts. Depending on the data, shapefiles can be downloaded for the entire nation, a state, or a county within a state. Shapefiles containing line segments are distributed for individual counties (Figure 3.4). For TIGER/Line segments that are street centerlines, attributes for the left- and right-hand sides of the street segment are coded. These include a wide range of attributes for each side: street name, address range, ZIP Code, census and political unit identifiers, and congressional district identifiers.

A major benefit of the TIGER/Line files is that they provide a connection between street address ranges and locations on the ground. This makes it possible to locate or geocode address-based information such as hospital discharge records, birth certificates, and clinic locations. However, the TIGER/Line files do not record a precise location for each address, just an address range along a street segment; therefore, address locations can only be approximated by interpolation, as described later in this chapter. This may pose a few problems in urban and suburban areas where addresses are spread relatively evenly along street segments, but in rural areas TIGER/Line files should be used with caution for locating addresses if a high degree of positional accuracy is required.

Despite their wide coverage and applicability, the TIGER/Line files have had several important limitations. First, street and address coverage is incomplete and in some cases inaccurate in earlier versions of the data. Streets may be missing or misnamed. Address ranges may be missing, include incorrect values, or identify the wrong side of the street. These problems are especially relevant in rapidly growing communities where new residential development has taken

TABLE 3.1. Selected Shapefile Components

File extension	Content	Status
.shp	Geographic feature geometry	Mandatory
.shx	Index of feature geometry	Mandatory
.dbf	Attribute table with variables describing features	Mandatory
.sbn	Spatial index of features	Optional
.prj	Description of coordinate system and projection	Optional
.shp.xml	Metadata in XML format	Optional

FIGURE 3.4. A portion of the TIGER/Line database showing road features in the area around downtown Hartford. This figure shows roughly the same area as Figure 3.1.

place. Many local governments enhanced and improved the TIGER/Line files for their local areas, and this information has been used to update TIGER in the years following its introduction (Sperling, 1995).

A second problem is that the positional accuracy of TIGER files as they were originally developed from multiple sources was unknown and varied from place to place. Positional accuracy of the original TIGER files was "no greater than the established National Map Accuracy Standards for 1:100,000-scale maps" (U.S. Census Bureau, 1992), and the files were logically consistent, but positionally inaccurate for large-scale mapping in some cases. These positional inaccuracies have implications for health studies. They may be of little consequence for a study of lead screening programs, for example, but extremely important in analyzing a problem like radon exposure.

Finally, TIGER data layers often did not match perfectly with layers generated from DLGs or DOQQs. Rubbersheeting techniques—techniques to adjust features to a foundation data layer—were often needed to combine TIGER data with data from other sources. Rubbersheeting is discussed in greater detail in the section on database integration later in this chapter.

In response to these problems, the Census Bureau has worked with various state and local agencies to improve the quality of address range information and

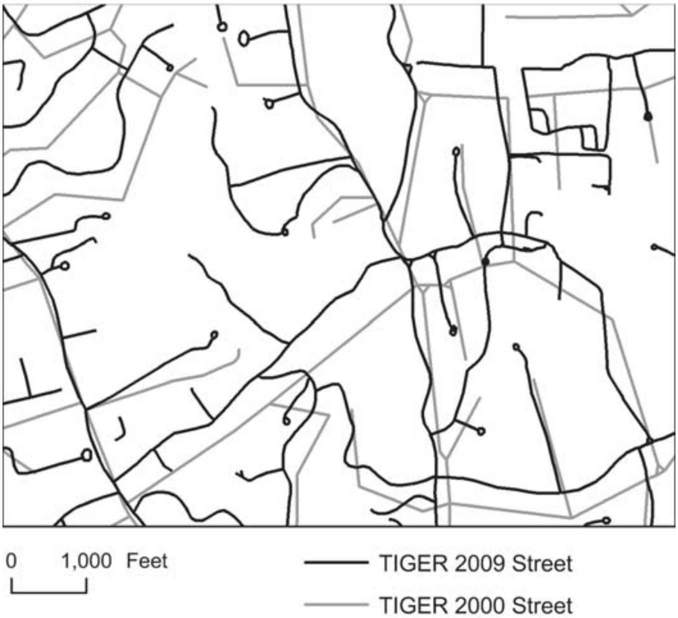


FIGURE 3.5. Overlaying street segments from the 2009 edition of the TIGER/Line database and street segments from the 2000 edition shows improvements in the positional accuracy of the data.

the positional accuracy of the TIGER/Line files for the 2010 census (Figure 3.5). The MAF/TIGER Accuracy Improvement Project, completed in 2008, resulted in the realignment of many, though not all, features based on data submitted by local and state agencies, imagery, and GPS data collected in the field (U.S. Census Bureau, 2009a). Even before these improvements, the TIGER/Line database has been one of the most important and widely used foundations for GIS-based health and socioeconomic analysis. Furthermore, the development of TIGER has prompted commercial firms to sell corrected and updated versions of the data. In fact, many GIS software packages come bundled with TIGER-based spatial data to facilitate mapping of census data. In developing a TIGER-based database for a GIS, it is well worth seeking out the most accurate and updated version. Analysts should also make sure that data used for geocoding health events is consistent with the data used for mapping census data by census tracts or other units so that health events will be correctly allocated to areas.

Cadastral Data

Another source of address-based spatial data that is generally more accurate than TIGER/Line for small geographic areas is cadastral information. ***Cadastral data*** are data associated with land ownership, and they are a matter of public record in the United States. Cadastral features are not visible on the ground, but are legally defined to specify ownership and administration of land parcels (Huxhold & Levinsohn, 1995). Digital cadastral data files contain property boundaries and a wide range of attribute data including land title, address, sale/resale information and building type, size, and characteristics (Figure 3.6). Property boundaries are stored in a vector format, with property attributes attached. Because the files describe land ownership, they often have a high degree of positional accuracy and represent large spatial scales—1:12,000 or larger. Address information is generally accurate and complete. Cadastral data also show street widths and thus better depict the built environment of a local area than do TIGER/Line files.

Despite these advantages, cadastral data have important limitations. Although most communities collect and maintain cadastral spatial data, conversion to digital form has been a relatively recent development. Some communities still rely on paper maps and written descriptions of property boundaries, some decades old. Furthermore, the quality and accuracy of cadastral data vary widely, depending on the recency and quality of the surveying or historical information on which it is based. Errors can creep into cadastral databases over time

FIGURE 3.6. A portion of a cadastral database for the area around downtown Hartford. The figure shows roughly the same area as Figure 3.1. Property databases are generally maintained by local governments, so this database covers only Hartford and does not include properties in East Hartford.

depending on how well the registry is kept as boundaries are resurveyed, landscape change occurs, and properties are subdivided. In addition, communities use different formatting systems for digital cadastral information and capture different attributes, so conducting studies across community boundaries can be challenging. Efforts are currently underway to develop common standards for cadastral information. Cadastral data files can also be large, unwieldy, expensive to create, and unnecessarily detailed for some kinds of spatial analysis. Network models such as those discussed in Chapter 10, for example, require transportation routes to be represented as arcs, as in TIGER, rather than as double lines. Still, cadastral data offer an excellent foundation for address-matching and mapping in small areas.

Choosing a Foundation Database

The foundation data sets described in this section each offer a unique set of advantages and disadvantages for public health GIS. They differ in scale, resolution, positional accuracy, and display of features, as well as in their raster or vector structure. They are also evolving over time.

The choice among foundation data sets depends on the scale and scope of the project, the resources available for data creation, and the types and scales of other data sets to which the foundation data will be linked. Projects that are national or regional in scope are more likely to utilize intermediate scale foundation data such as satellite imagery, DLGs, and TIGER/Line data. In contrast, studies of single communities or neighborhoods can take advantage of the detail and positional accuracy of cadastral data and DOQQs.

Population Data

Foundation data create a platform for integrating spatial data layers that contain population and related health, social, and environmental information frequently used in health applications of GIS. The TIGER/Line files store spatial information for the geographical units the Census Bureau uses in tabulating and publishing the population data it collects. An understanding of census geography is important for any public health analyst who uses data compiled by the Census Bureau (U.S. Census Bureau, 2008a). Not all of the population data tabulated by the Census Bureau is published for every level of census geography (Peters & MacDonald, 2004).

The smallest unit is the census ***block***, and each block is bounded by a set of connected street segments or other linear features such as rivers, railroad tracks, or municipal boundaries (Figure 3.7). A ***block group*** is a cluster of blocks, typically containing from 600 to 3,000 people. Census ***tracts*** comprise groups of contiguous blocks (and block groups) and have populations ranging from 1,200 to 8,000. TIGER shapefiles are also available for state, county, and local political

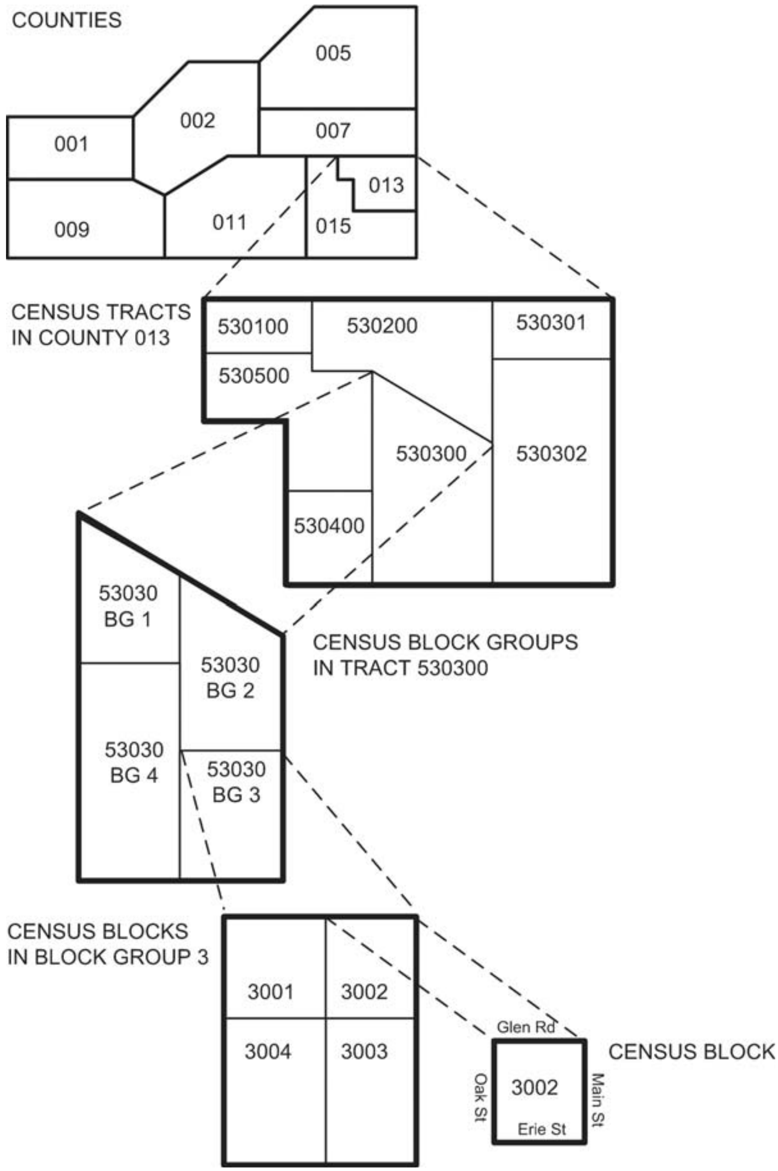


FIGURE 3.7. Geographic subdivisions for the U.S. Census. The smallest unit is the block. Each county is divided into census tracts, which are divided into block groups, and then into blocks. The first digit of the census block identifier corresponds to the block group.

boundaries, along with ZIP Code boundaries. As described in another section in this chapter, database tables describing the population, health, and socioeconomic attributes of these areas can be joined to databases describing the geography of the areas.

The hierarchy of census units in relation to other political and administrative units is complex, given the federal nature of the U.S. system of government (Figure 3.8). Census blocks, block group areas, and tracts nest perfectly within counties, but other local areas do not necessarily follow this pattern. In addition to data provided for counties, data are provided for places or minor civil divisions that are legally incorporated and bounded areas such as cities, towns, and villages. Census blocks nest perfectly within these units. Census block group and

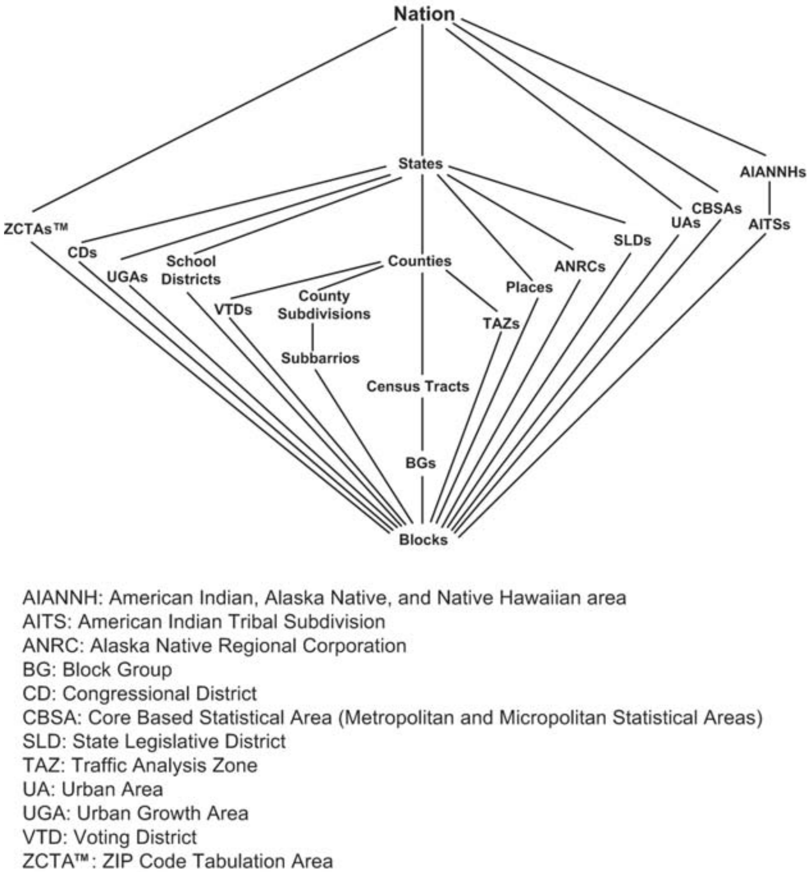


FIGURE 3.8. Hierarchical relationships of census and political or administrative areas in the United States for the 2010 census. Census block group areas and census tracts nest within counties, but their boundaries overlap the boundaries of many other political and administrative entities.

tract areas, however, do not always nest perfectly within places or minor civil divisions. Places may cut across county boundaries.

In some cases, census tracts may coincide with areas where many residents live in group quarters like prisons, military bases, or colleges and universities. Because these residents often differ from the general population in terms of age and sex and residential mobility, it is important in health applications of GIS to make explicit decisions about how to include group quarters populations. The population of a college town such as Mansfield, Connecticut, where the main campus of the University of Connecticut is located, is very different during the academic year than during the summer months (Figure 3.9).

In the United States, a complete enumeration of the population is conducted every 10 years, as mandated by the Constitution for the purposes of apportioning seats in Congress. Beginning with the 2010 census, the American Community Survey, a program initiated after the 2000 census for providing more up-to-date census data during the intercensal period, will be fully operational. The Census Bureau provides information on the dates of censuses conducted or scheduled in other countries from 1945 to 2014 (U.S. Census Bureau, 2008b) and links to statistical agencies in other countries responsible for population data (U.S. Census Bureau, 2010a).

Health Data

This section describes some of the major types of health information that can be incorporated in GIS for health planning, evaluation, and research. Our aim is to introduce these data sets and highlight geographical issues that affect data use and integration in GIS. Detailed discussion of the content of these data sources is available elsewhere (Halperin & Baker, 1992; Parrish & McDonnell, 2000; Huber, Boorkman, & Blackwell, 2008).

Registration System Data

VITAL STATISTICS

Local governments in the United States and other countries routinely collect information on all births and deaths that occur in their jurisdictions. These ***vital records*** are an important source of spatial data for public health GIS. Birth records document a wide range of conditions that affect newborn infants, including birthweight, gestational age, congenital malformations and obstetric procedures, along with the mother's demographic and social characteristics and her use of prenatal services (Friis & Sellers, 2009). Information about the infant and the birth process is generally accurate, but data for the mother, especially data based on recall of timing and events during pregnancy, can have errors and inconsistencies. Still, birth data offer a nearly complete summary of basic maternal and infant health indicators for the population.

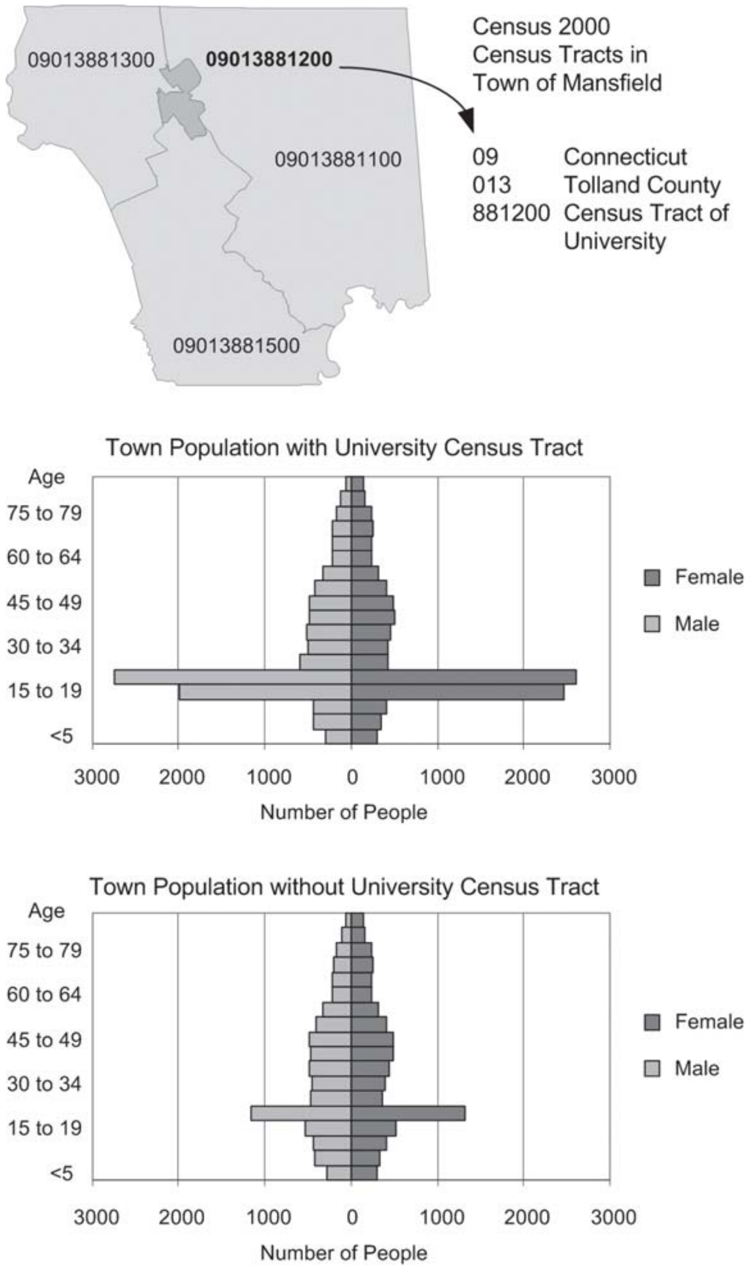


FIGURE 3.9. One census tract area from the 2000 census in the town of Mansfield has a large group quarters population because the University of Connecticut's main campus is located there. The size and age–sex structure of the town's population differs significantly depending on whether the tract where the university is located is included.

Birth records include the mother's residential address, a geographical identifier for GIS mapping and analysis. This information has been used to study environmental and neighborhood influences on maternal and infant health, for example, the effects of proximity to prenatal care services on prenatal care use and birth outcomes (McLafferty & Grady, 2004) or the clustering of birth defects in relation to hazardous waste sites (Rushton & Lolonis, 1996).

Health departments also collect and report data on deaths in ***mortality records***. Generated from death certificates, these data include demographic characteristics of the decedent and information about the cause of death, including the immediate cause and contributing factors (Friis & Sellers, 2009). Although demographic information is typically accurate, there are well-known problems with cause-of-death information stemming from errors and inconsistencies in diagnosis and difficulties in assigning causes when multiple causes are present (Garbe & Blount, 1992). Death certificates include two types of address-based geographic data: the place of death and the usual residence of the decedent. The place of death is often a hospital, nursing home, or other health care facility. This information can be used in analyzing health outcomes and service utilization by facility. In contrast, residential addresses provide a means for linking the residential environment to mortality outcomes.

Address-based vital statistics information presents several challenges to the GIS researcher. Addresses may be incorrectly coded, making it impossible to identify geographic locations. In a study of birth defects in Des Moines, Iowa, 8% of birth records could not be geocoded because of errors in the addresses and "P.O. Box" and "rural route" style addresses (Rushton & Lolonis, 1996), although geocoding using parcel or E911 data may increase the match rate (Mazumdar, Rushton, Smith, Zimmerman, & Donham, 2008). Also, because of privacy and confidentiality concerns, many health departments do not release address information (Istre, 1992). They only provide data in aggregate form, by ZIP Code, district, or census tract, making it impossible to analyze point locations. Finally, even if current residential address information is correct, it may not accurately represent the environment of the person before and during pregnancy or prior to death because the relevant exposure may have occurred someplace other than the residence (see Chapter 6). This is particularly problematic for mortality data, given that the conditions that lead to death can result from lifelong exposures and behaviors.

MORBIDITY DATA FROM SURVEILLANCE SYSTEMS AND DISEASE REGISTRIES

Looking beyond life's vital events, morbidity data are an essential source of information for public health GIS. ***Disease surveillance*** involves monitoring distributions and trends in morbidity and mortality data collected for a specified population and geographical area. There are many kinds of morbidity data, ranging from information gathered by government agencies and health care providers to information from survey research projects. These data differ greatly in content, coverage of the population, and geographic scale at which they are normally available.

Reportable disease data provide information on morbidity and mortality for certain "reportable" health conditions. Infectious disease has always been an important focus of public health surveillance in the United States (Centers for Disease Control and Prevention, 2008b). Authority to require notification of cases of disease resides in state legislatures, and there is considerable variation in state provisions. All 50 states require physicians to report cases of specified notifiable diseases to state or local health departments. Notifiable disease reports and vital records are the two health data sources available at the local level in all states.

The ***National Notifiable Diseases Surveillance System*** is operated by the Centers for Disease Control and Prevention (CDC) in collaboration with the Council of State and Territorial Epidemiologists (CSTE). Reporting by the states to the national system is voluntary. States generally also report internationally quarantinable diseases (cholera, plague, yellow fever) in compliance with World Health Organization (WHO) International Regulations. There are approximately 50 infectious diseases designated as notifiable at the national level (Council of State and Territorial Epidemiologists, 2009). The list of nationally reportable infectious diseases and other conditions changes periodically, and reporting practices may differ from state to state (Roush, Birkhead, Koo, Cobb, & Fleming, 1999).

In addition to notifiable disease reports by providers such as physicians, hospitals, and laboratories, key data sources for infectious disease reporting in the United States include sentinel systems, hospital surveillance, school surveillance, special surveys at the state and local level, vital records, and vector/host surveillance for zoonotic diseases. ***Sentinel health events*** are cases of illness that signal a need for immediate public health intervention or serve as a warning of hazardous conditions or poor quality medical care. A case of polio, for example, might signal a breakdown in the quality of immunization programs. A number of limitations of the current surveillance system have been described (Birkhead & Maylahn, 2000). The fragmentation of the system and voluntary reporting requirements affect the completeness of surveillance data. Generally the reporting system is thought to work well for diseases that are serious, have clear symptoms, and require medical attention. However, coverage is incomplete for conditions that can be asymptotic (tuberculosis), that do not necessarily compel medical treatment (animal bite, gastroenteritis), or that carry social stigma (HIV/AIDS) (Friis & Sellers, 2009).

Underreporting of infectious disease conditions may be explained by a number of factors, including provider lack of awareness of reporting requirements. The level of public concern also affects disease reporting. Infectious diseases that carry some social stigma may be concealed. For many infectious diseases, symptoms are either too mild to prompt a person to seek medical care or mimic flu-like symptoms associated with other common illnesses. For others, particularly emerging infectious diseases, the etiological definition may be incomplete or the case definition for surveillance purposes may be inadequate. "What is a case?" is not a trivial question. There may be differences of opinion about the criteria for defining a case of a disease. Sometimes, case identification requires laboratory confirmation. In addition, case definitions change with changes in scientific knowledge. Changes in case definitions over time have an impact on what is included in the surveillance database, as discussed in Chapters 7 and 8.

Active surveillance systems obtain data by searching and periodic contact with providers. ***Passive surveillance*** systems rely on reports by providers. Because of the costs associated with active surveillance, this type of system is often used strategically in limited areas or for limited time periods. Evaluation of active surveillance systems indicates two- to fivefold increases in reporting of specified diseases and other conditions not subject to active surveillance (Vogt, LaRue, Klaucke, & Jillson, 1983; Thacker et al., 1986). Surveillance method, therefore, has implications for completeness of the data, an important dimension of spatial database quality. Active surveillance systems offer a mechanism for completing and correcting information from the reportable disease record including address data used as geographical identifiers.

To protect privacy and confidentiality, federal agencies only release reportable disease statistics at the county level. Different policies exist in lower levels of government: some state or local health agencies will make information available for smaller geographic areas, or even by address, as long as the analyst agrees to maintain privacy and confidentiality. When address information exists, its accuracy can be problematic. Addresses may be missing or inaccurately coded. In an epidemiological study of reported rat bites in New York City, for example, almost 40% of bite reports had missing or incorrect address information and could not be geocoded (Childs et al., 1998).

Disease registries are centralized databases for the collection of information on specific diseases, the best examples being the cancer registries managed by state and local health authorities (Friis & Sellers, 2009). Disease registries use a reporting system similar to that for reportable diseases, with health providers reporting occurrences to the appropriate state or local registry. Some disease registries actively seek out case information, while others simply gather reports. Furthermore, some registries keep longitudinal information, following patients after diagnosis in order to track changes in health status and treatment regimes.

Cancer registries, the most extensive disease registries in the United States, offer a potentially valuable source of information for GIS analysis. Currently, all 50 states and a number of localities in the United States maintain cancer registries, some of which have existed for decades, funded through the CDC's National Program of Cancer Registries (NPCR) or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program, or both (Centers for Disease Control and Prevention, 2011a). At the national level, the SEER program is an umbrella organization for a network of cancer registries that covers about 26% of the U.S. population. SEER includes active follow-up of living patients and is used to generate national estimates of overall cancer incidence and breakdowns by gender, race, age, and geographic location (National Cancer

Institute, 2011a). The North American Association of Cancer Registries promotes uniform data standards for cancer registration and the use of cancer surveillance data (North American Association of Central Cancer Registries, 2011). As with the other types of health data, registries include residential address information, and organizations like the North American Central Cancer Registries (NAACR) have developed valuable guidelines for geocoding health data (Goldberg, 2008).

The information in health databases and disease registries is protected by laws governing privacy and confidentiality. Some states will release addresses for research studies as long as appropriate measures are taken to ensure confidentiality; however, once again, policies differ among states. Other problems with address information arise from changes and errors in the coding and formatting of addresses.

Surveillance systems and disease registries have been sources of data for many GIS case studies but very few statewide surveillance systems or registries have been fully linked to GIS (Devasundaram, Rohn, Dwyer, & Israel, 1998; Cromley, 2000; South Carolina Vital Record and Statistics Integrated Information Systems Project Team, 2005). Implementation of a statewide or national surveillance system in GIS increases the likelihood that the case database will include cases identified using different case definitions and surveillance methods. To address this problem, case definition and surveillance method should be included as fields in a surveillance database.

Survey Data

To address a broader range of health issues than covered in standard vital statistics and morbidity data sets, public health researchers often turn to ***health survey data***. Surveys deal with a diverse array of health-related topics, topics that are beyond the scope of disease reporting systems and transcend biomedical concerns. Health surveys investigate health-related behaviors, psychosocial well-being, nutritional status, stress, and individual, family, and neighborhood circumstances that affect health. The major national surveys in the United States include the National Health and Nutrition Examination Survey (NHANES) and the National Health Interview Survey (NHIS). These surveys ask a detailed set of questions to a small, representative sample of the U.S. population. NHANES focuses on physiologic measures, measures of body weight and stature, and nutritional assessments. It has been conducted in several cycles since the early 1970s (Centers for Disease Control and Prevention, 2009a). NHIS, administered annually since 1957 with the U.S. Census Bureau serving as the data collection agent, collects information on health risk factors, chronic conditions, injuries, impairments and health service utilization, based on household interviews (Centers for Disease Control and Prevention, 2011b).

The purpose of these surveys is to develop a national picture of the health status of the population. Not every place is sampled. NHANES uses a four-stage sampling procedure. In stage 1, ***primary sampling units*** or PSUs are selected.

These are usually single counties but may be groups of contiguous counties. Samples are selected with probability proportional to size. In stage 2, the PSUs are divided into smaller areas called segments generally equivalent to city blocks, and segments are selected with probability proportional to size. In stage 3, within each segment, households are listed and selected by random sample. In geographic areas with high proportions of adolescents and elderly, minorities, and low-income whites, households are oversampled. In stage 4, individuals are chosen from a list of all persons in selected households. The NHIS uses a similar multistage design to sample individuals in all 50 states and the District of Columbia. Given the purpose and sample design of these surveys, they provide data primarily at the national level.

By the early 1980s, the need for more data on health risk behaviors at the state level led the Centers for Disease Control and Prevention to develop the ***Behavioral Risk Factor Surveillance System (BRFSS)*** (Centers for Disease Control and Prevention, 2009b). Initially, 29 states participated in the program and conducted telephone surveys of the adult population. By 1994, all states, the District of Columbia, and several territories were participating. In addition to a core set of questions, BRFSS includes a set of modules addressing specific health risk behaviors and provides the opportunity for individual states and territories to add state-specific questions.

Although some states from the outset developed telephone sampling designs that would make it possible to report results for selected areas below the state level, BRFSS provides primarily state-level data. In response to demand for more local-level data, the BRFSS SMART program offers data for selected metropolitan areas and small cities with 500 or more BRFSS respondents. Through the BRFSS Maps site, users can download shapefiles to which BRFSS data for individual survey years have been joined (Centers for Disease Control and Prevention, 2009b).

Other countries have developed and implemented similar health surveys. The Health Survey for England, for example, is a series of annual surveys conducted since 1991 (U.K. Department of Health, 2011). Studies comparing these surveys have highlighted differences in methodological approaches (Aromaa, Koponen, Tafforeau, Vermeire, & the HIS/HES Core Group, 2003). Despite these differences, the surveys provide data for international comparative research.

Surveys are also used to screen for problems like lead poisoning, PKU, and hypertension. ***Screening surveys*** are proactive public health activities that attempt to uncover health problems before symptoms appear, when the problems are difficult and expensive to treat. Screening surveys differ in the range and nature of population covered. Some cover the full population, as in screening of newborns for PKU, and thus can be used to estimate incidence rates and create maps of geographic variation in incidence. By contrast, many screening surveys only target high-risk populations and people likely not to have been screened as part of their regular health care. Estimates and maps prepared from such surveys only pertain to the screened population. Reported incidence will naturally be higher in areas where more people were screened, and GIS can be used to explore geographic variation in ***screening penetration***, the percent of risk population screened.

Health Care and Health Care Utilization Data

Medical care provision generates large quantities of information on patients and the treatment they receive, and secondary use of administrative data is made in many health studies. Most medical care providers and insurers maintain data on residential addresses of patients. The geographical organization of health care affects health care utilization, as discussed in Chapters 9 and 10. Nevertheless, data on the locations of medical care providers and the health problems of patients they treat are important sources of data for GIS applications.

HEALTH PROVIDER DATA

Health service information forms another valuable spatial data layer for public health GIS. Most health care providers—hospitals, physicians, clinics—offer their services from fixed locations and can be represented as point spatial data. A few health services, such as emergency medical services and mobile clinics, move from place to place and thus can be modeled as arc or network information. Beyond location, many other dimensions differentiate health services, including price, capacity, utilization, range of services provided, and the elusive quality of care.

Information about the locations and characteristics of health care providers is widely available. ***Gazetteers*** include geographical coordinates for major health facilities such as hospitals. These coordinates can be imported into GIS for mapping and display; however, one must be careful that the location coordinates use the same scale and projection system as the foundation data layer to which they will be linked. One shortcoming of gazetteers is that they do not include data on the characteristics of health service facilities. Such information must be brought in from other sources and linked to the facility sites.

Detailed information on health care providers comes from professional organizations like the American Hospitals Association (AHA) and the American Medical Association (AMA) and, increasingly, from commercial marketing database providers like InfoUSA. The AHA publishes an annual directory of hospitals that includes statistics on utilization, personnel, services, and finances for hospitals in the United States (American Hospital Association, 2009). Included in the directory is each facility's street address, which can be geocoded to a point location. Similar kinds of directories exist for nursing homes and mental health facilities.

For physicians, the AMA's Physician Masterfile offers analogous information and includes addresses that can be geocoded to identify point locations (American Medical Association, 2011). A directory of physicians based on data from the Masterfile is also available (American Medical Association, 2009). The Masterfile covers the vast majority of physicians, but certain important subgroups may be missing, for example, doctors who earned medical degrees outside the United States whose practices are often clustered in immigrant neighborhoods. As with other types of health data, the release of data on providers raises privacy and confidentiality concerns. Physicians have protested the sale of their data to businesses, including pharmaceutical companies (Saul, 2006). State laws prohibiting the sale of doctor-specific prescription drug data are being tested in the federal courts (Saul, 2008).

Data for other types of health care providers are often harder to come by. Health clinics, for example, are operated by federal, state, and local governments as well as voluntary organizations. Each type of agency maintains a list of its own clinics, but there may be no composite listing of facilities in an area. It may be necessary to piece together information from multiple sources or conduct fieldwork to uncover all health service locations. Despite these challenges, creating spatial data layers for health care providers is generally easier than preparing health and foundation data layers. Health services are limited in number, exist at discrete locations, and change relatively slowly over time, making them more manageable to deal with in a GIS context.

HEALTH CARE UTILIZATION DATA

Hospitals generate large quantities of spatial information on patients treated in their inpatient and outpatient facilities. These ***hospital discharge data*** provide an important base for examining hospital utilization and treatment patterns, though they are generally inadequate for population-based studies of morbidity because they are restricted to patients treated in hospitals. The large literature on small-area variations in the rates of medical and surgical procedures relies primarily on hospital discharge data (Wennberg & Gittelsohn, 1982), and the data sets are widely used in health policy analysis and planning. Included in the data sets are demographic information about the patient, primary and secondary diagnoses, diagnostic procedures, treatment procedures, length of stay, and insurance status. Hospital discharge data contain the patient's residential address, but that information is rarely released due to privacy considerations. Instead, hospital data can usually be obtained at the ZIP Code level, because ZIP Codes are part of the address and thus convenient geographical identifiers for the release of hospital information.

This section has described several important, widely available health data sets that can be incorporated in public health GIS. The data sets address a cross section of public health issues and offer a framework for diverse geographical investigations. Increasingly these information resources are available on electronic media, including Internet, and are readily accessible to users in a wide variety of settings (Lacroix & Backus, 2006). Many other health data sets exist. We have not even mentioned the vast proprietary databases held by health insurance companies or the specialized data sets in areas such as occupational, veterinary, and environmental health (Weise, 1997).

Spatial Resolution of Health Data

Regardless of which data sets are used, the spatial resolution of the data is crucial for GIS applications. Although all health data sets deal fundamentally with individuals and usually include address information, none routinely release those detailed geographical identifiers because of important privacy and confidentiality considerations. Thus, the analyst is typically faced with using health data that are aggregated to predefined geographical units, such as counties, ZIP Codes, or census tracts. This raises important substantive issues, as well as significant methodological issues as discussed in Chapter 5. Substantive issues concern the validity and usefulness of particular areal units for public health planning and analysis.

Most data from federal health agencies are available at the county level. Although counties are generally good geographical units for displaying health data at the national scale, they have important limitations (Croner, Pickle, Wolf, & White, 1992). Counties are administrative, political units that bear little relationship to areas defined according to socioeconomic, demographic, or environmental criteria. Counties often encompass diverse physical environments and heterogeneous populations. Moreover, the areas differ greatly in population size and areal extent. Counties large in area visually dominate the national map, despite the fact that they may have tiny populations. Small urban counties can hardly be seen on a national map, though they have huge populations. Thus, counties are not comparable to one another, and they have little basis in population and environmental factors relevant to public health. By comparison, census tracts, defined by the Census Bureau for tabulation purposes, are more similar than counties in population size and follow moderately well the fuzzy boundaries of social, economic, and ethnic areas.

ZIP Codes, commonly used for the tabulation of health data, have problems analogous to those for counties. Originally, ZIP Codes were devised by the U.S. Postal Service to facilitate mail delivery, each ZIP Code representing a collection of mail distribution points. The areas have little correlation with socially and environmentally defined areas. In cities, some ZIP Codes encompass neighborhoods with highly divergent economic and social characteristics. For instance, one ZIP Code in New York City includes census tracts whose 1990 median incomes ranged from $15,000 to $42,000, a threefold difference. A health statistic for such a ZIP Code would represent an "average" of statistics for two very different population groups. Another problem is that ZIP Codes occasionally cut across political and census boundaries and change over time, making it difficult to overlay and integrate ZIP Code data with other sociopolitical data (Kirby, 1996; Krieger et al., 2002). Despite these problems, ZIP Codes in the United States and postal codes in other countries are often used as convenient and practical reporting units for health data in small areas. Analysts should be aware of the strengths and limitations of using ZIP Codes for GIS-based health analysis.

Making Population and Health Data Mappable

In order to use population, health, and health care data sets in GIS, the data sets must first be captured and linked to a foundation spatial database. Data capture is a complex process that draws on an ever-increasing array of tools including scanning, digitizing, downloading from the Internet, and entering data directly from the field via the Global Positioning System. This section focuses on the two procedures typically used for capturing health information—address matching and joining.

Address Matching to Locate Health Events as Points

Health information is often georeferenced by street address. For example, we might have information on the residential addresses of people who died of breast cancer, or the addresses of hospitals, health clinics, schools, or workplaces. Using the process of ***address-match geocoding***, we can convert each address to a point on a map. The point is recorded as a pair of geographical coordinates that connect to the foundation database. At its simplest, address matching involves comparison of two data sets: one containing the addresses of health events and the other a foundation database with its own address information. An address (street name, number, and city, ZIP Code, or other zone) from the first database is compared against the full array of addresses in the second, and a "match" occurs when the two agree.

Address-match geocoding procedures differ depending on the type of foundation spatial database used in matching. Street centerline, address point, and cadastral or property databases have all been used in geocoding (Zandbergen, 2008). Address point databases are commonly developed as part of E911 systems in North America. E911 or Enhanced 911 is a telecommunications system that associates a calling party's telephone number with an address. Address point databases can also be created from parcel data. The point may be located at the centroid of the parcel or at a location where the driveway serving the property intersects with the road. When a match occurs, the health event is assigned the geographical coordinates of the corresponding property. Up-to-date property databases form an accurate platform for address matching because each address is associated with a unique property on the ground. The downside is that such databases are typically very large and cumbersome to work with and, like address point databases, compiled and maintained at the local level.

Although some GIS analyses have used more than one database to geocode addresses (Lovasi et al., 2007), street centerline databases, like TIGER/Line, are widely used as a foundation for address matching. Because street centerline databases do not include unique street addresses for specific structures, but only address ranges along street segments, address matching relies on interpolation. We match the particular address to a street segment (street name and address range), and we estimate the location of the address along the segment by interpolating within the corresponding address range. For example, the street address, 107 Oak Avenue, is assigned to the segment of Oak Avenue with address range 101–119 (Figure 3.10). By interpolation, the location of 107 Oak Street is estimated to be about one third of the way along the street segment. GIS users can specify an offset to take into account the setback of the structure from the street centerline.

This form of address matching does not place points at the exact locations of the structures, but rather at estimated locations along street segments. In urban and suburban areas, where properties are spaced fairly evenly along segments, spatial accuracy is generally quite good. In rural areas, the uneven spatial distribution of properties can cause significant spatial error from interpolation.

Address matching is an iterative procedure in which we first attempt to match all addresses and then correct those that fail to match. Typically one half to two thirds of addresses match in the first attempt. We then examine the unmatched addresses for obvious errors or inconsistencies. Often there are simple errors in spelling or abbreviation that can easily be corrected. After correcting obvious errors, it is typical to achieve a "match rate" of over 90% in most parts of the United States. Anything less calls for an assessment of the quality of both the address list and the spatial database used in matching.

Addresses can fail to match because of errors in the address list, errors in the street or property database, or inconsistencies between them. Errors in the address list are common, and they include misspellings and typographical errors

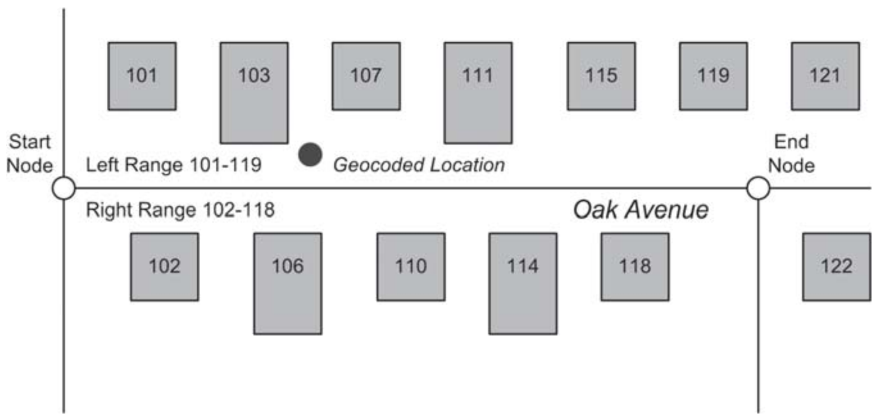


FIGURE 3.10. The TIGER/Line files include street centerline and address-range information that can be used in geocoding. This segment of Oak Avenue is represented by a start node and an end node, each with geographic coordinates. The address range for the left side of the segment contains odd-number addresses, while that for the right side contains even-number addresses. By interpolation, the location of 107 Oak Avenue is estimated to be approximately one third of the distance along the corresponding street segment of Oak Avenue. An offset was applied so that the geocoded point would not lie on the street centerline.

in the street number, street name, or zone (Table 3.2). These errors can easily be corrected by carefully inspecting and editing the address list. Furthermore, most GIS provide the option of automatically correcting the most common types of errors using simple rules and conventions. One should approach these automatic correction algorithms with caution, however, because they may falsely change the original address data and generate a false sense of accuracy.

Errors in the street or parcel database, including missing street segments and incorrect address range information, also create problems for address match-

TABLE 3.2. Sources of Error Affecting Address Match Outcomes

Record content	Street numbers	Street name	Street type	Zone (ZIP Code example)	Address match outcome for perfect match
Correct address	16	Main	St.	13501	Match at correct location
Correct street segment	Left 2–20 Right 1–19	Main	St.	13501	
Error in address record					
Incomplete address		Main	St.	13501	No match
Error in street number	166	Main	St.	13501	No match
Error in street name	16	Nain	St.	13501	No match
Error in street type	16	Main	Rd.	13501	No match
Error in zone	16	Main	St.	113501	No match
Address does not correspond to a real structure	16	Main	St.	13501	Match represents a structure that does not exist
Error in street segment record					
Missing range	Left Right	Main	St.	13501	No match
Error in range	Left 2–14 Right 1–19	Main	St.	13501	No match
Range applied to wrong side of street	Left 1–19 Right 2–20	Main	St.	13501	Match represents incorrect location
Error in street name	Left 2–20 Right 1–19	Nain	St.	13501	No match
Error in street type	Left 2–20 Right 1–19	Main	Rd.	13501	No match
Error in zone	Left 2–20 Right 1–19	Main	St.	113501	No match
Incomplete street network database					No match

ing. As the accuracy of spatial databases improves, it is less common than in the past to find true errors in such databases. Rather, most errors result from the time lag between new residential development and database update. Addresses fail to match because they are located in newly developed areas that have not been mapped or entered into a spatial database. Since these addresses must then be geocoded or digitized by hand, it is well worth the investment to use the most accurate and up-to-date street or parcel database.

Finally, addresses can fail to match because of inconsistencies between the address list and the foundation database. These include differences in street naming convention—for example, "6th Avenue" *versus* "Avenue of the Americas"—or in abbreviation—"St." *versus* "Str." Most GIS automatically correct obvious differences in abbreviation.

Although most analysts emphasize the "match rate," it is important to remember that a successful address match does not guarantee accuracy. Even if an address is successfully matched, it may not be assigned to the correct location. A field check of over 500 geocoded residential addresses to assess spatial accuracy uncovered a variety of errors (Cromley, Archambault, Aye, & McGee, 1997). The relative locations of 7% of the cases were incorrect. A few cases (less than 1%) had been geocoded to locations more than 500 feet away from the correct location. This type of error would be of particular concern in any study measuring distances from the geocoded location to another location because the true distance would be over- or underestimated. The remaining cases were estimated to be out of position by less than 500 feet. About half of these cases were on the wrong side of the street or on the wrong corner of an intersection. This type of error would be of particular concern in any study aggregating cases to an area like a census block or block group because census area boundaries often coincide with street centerlines, so cases on the wrong side of the street would be aggregated to incorrect spatial units. For 1% of the addresses, no residential structure could be found. Either the structure had been removed or the street number was incorrect but fell within a valid address range. The type of error—successfully geocoding an address that does not exist—may account for the higher match rate for street centerline geocoding versus parcel geocoding. Such errors can have significant impacts on spatial analyses based on geocoded data (Griffith, Millones, Vincent, Johnson, & Hunt, 2007).

These findings emphasize the importance of obtaining accurate address information and the need to look beyond the match rate in geocoding. Typically, the collection of addresses is decentralized. Addresses are entered at the source institution, for example, a hospital, doctor's office, or health clinic. From there, the institution transmits the information to a public health agency for mapping. Unless the addresses are used for billing or follow-up, the institution will have little stake in their accuracy and completeness. Errors emerge much later during address matching, and data editing and cleaning are performed by GIS personnel far removed from the source of data collection. Improving accuracy in geocoded address information requires not just better address-matching algorithms, but institutional arrangements that foster accuracy at the source.

The findings also emphasize the need for field checking of data, particularly when research findings are sensitive to the locations of cases in a few places. Researchers involved in GIS studies at a community scale can benefit from field trips to the study area before data collection and analysis to familiarize themselves with residential patterns and other landscape features of relevance to the particular study.

The widespread use of geocoded health data has led many health analysts to investigate methodological issues in geocoding (Rushton et al., 2008). A relatively neglected issue in the study of geocoding methods is the spatial distribution of errors. Many analysts have noted that geocoding rates and accuracy are lower in rural areas, but more research is needed to model other forms of error and their implications for spatial data analysis.

Joining Health Data to Geographical Areas

Many population and health databases only present information for geographical areas like counties, ZIP Codes, or census tracts. They include the area name and/or identifier and a set of variables that describe the health events, population, or other attributes of the area—for example, the census tract number and number of diagnosed cases of AIDS by tract. Capturing area data in a GIS involves ***joining***. We join the tabular data to a foundation spatial data set of area boundaries based on a common field like the census tract identifier. The data for each tract are attached to the corresponding tract in the foundation database.

Joining requires that each geographic area have a unique identifier, either a name or number. In the United States, state names are unique, as are ZIP Code numbers. However, many widely used areal units such as census tracts or blocks have identifier numbers that are unique only within larger units of geography. Census tract numbers are unique only within counties, and block numbers are unique only within tracts. A project that cuts across these larger units must create a new field that uniquely identifies each small area. In a tract-level study that encompasses many counties in several states, for example, the state number and county number must be included along with the tract number to define each tract (Figure 3.11).

For more than 30 years, the Census Bureau used the Federal Information Processing Standard (FIPS) codes for states, named populated places, primary county divisions, and other entities, issued by the National Institute of Standards and Technology. At the time of preparation for the 2010 census, the NIST standards were withdrawn and the Bureau was transitioning to a set of codes issued by the American National Standards Institute (ANSI). Many of the codes adopted by ANSI are unchanged, but users need to familiarize themselves with the new standards (U.S. Census Bureau, 2010b). The InterNational Committee for Information Technology Standards 446-2008 standard is tied to the Geographic Names Information System (GNIS) managed by the U.S. Geological Survey. Familiarity with the ANSI and Census Bureau identifiers for geographical units within GIS databases is important for accurately joining and manipulating geographic databases produced by the

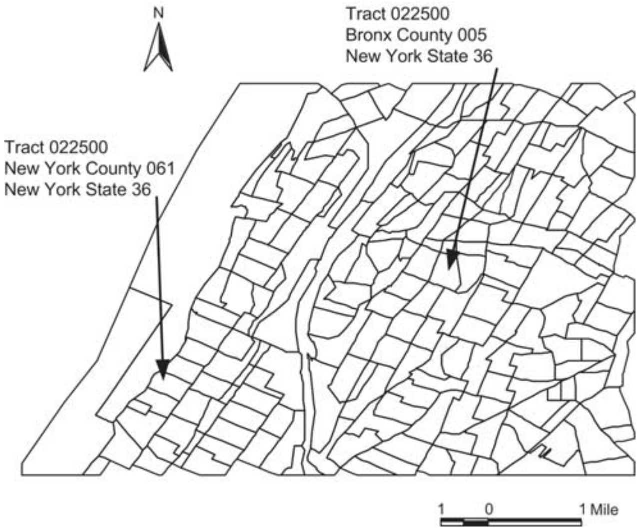


FIGURE 3.11. Census identifiers for tracts, block groups, and blocks are unique only in the context of the hierarchy of census units. Two tracts in New York City have the same tract identifier number, 022500. Use should be made of the full codes, 36061022500 for the tract in New York County and 36005022500 for the tract in Bronx County, to avoid confusion and errors in joining tables of data.

federal government (Table 3.3). State governments may have developed additional numeric identifiers for geographical units within their states.

Typically, joining links area-based health information to the corresponding geographic areas in a foundation spatial database. However, the procedure can also be used with address-based health data to find the area in which a health event is located. We match the address field in the health data set directly to the corresponding address field (street name and address range) in a foundation database table, like TIGER/Line, that contains area identifiers. Two tables, one for the health data set and the other for the foundation data set, are joined based on common address information. If this approach is used, the ability to map address-based health data as points is lost. As a consequence, the methods described in Chapters 4 through 10 for analyzing patterns of health data represented as points cannot be applied. Spatial information is lost when address-based health data are joined to areas rather than geocoded as points.

Database Integration

The power of a GIS lies in its ability to link, integrate, and manipulate the diverse types of spatial data described in this chapter. Integrating such data sets

TABLE 3.3. Comparison of ANSI, Census, and State Identifiers for an Area

	Hartford (Populated place)	Hartford (City)	Hartford (Town of)
ANSI/Census state code for Connecticut	09	09	09
ANSI/Census county code for Hartford County	003	003	003
ANSI GNIS feature identifier	213160	2378277	213442
Census code	37000	37000	37070
State of Connecticut town identifier	—	—	064

Note. These entries show that there are six different numerical codes for the same geographic area called Hartford. The ANSI and Census codes for the state of Connecticut and for Hartford County are the same. The ANSI standard using GNIS (Geographic Names Information System) codes has three different feature identifiers for Hartford. The census has two different codes for Hartford, one as an incorporated place and one as a minor civil division. Under a numerical coding system used by the state, the town of Hartford is 064. GIS users need to be aware of the coding systems that have been used to assign identifiers to geographic areas when joining and linking databases.

can be challenging, especially when the data sets differ in scale, resolution, and geographic extent. Most GIS packages include a series of cartographic and geographic procedures for linking and integrating spatial data sets (Table 3.4).

A common data integration problem arises when data layers that will be overlayed or linked in a GIS rely on different coordinate systems or different map projections. Typically this occurs when a health or an environmental data layer is being integrated with a designated foundation data layer. Common in all GIS are procedures for transforming coordinates so that they are consistent with those of the foundation data layer. ***Coordinate translation*** involves computing new coordinates as a mathematical function of the original set. Linear transformations, for example, can be used to move, stretch, or twist the coordinate axes (Figure 3.12). These simple linear transformations are often necessary when integrating spatial data from a digitizing tablet or scanner with existing geospatial foundation data sets like DOQQs or DLGs.

Sometimes geographical errors in overlaying data layers stem from positional inaccuracies that are unevenly or unpredictably distributed across the map. In this case, matching data layers requires nonlinear coordinate transformations that stretch or shrink different parts of a map until features align correctly with those on the foundation data layer. ***Rubbersheeting*** is the process of geometrically adjusting features to force a digital map to fit the designated foundation data layer (Antenucci et al., 1991). Rubbersheeting changes the relative locations of features, thus distorting the original map. Therefore the process should be used judiciously to make relatively small changes in map coordinates.

TABLE 3.4. Spatial Database Collection and Preprocessing Operations

Function class	Function
Data collection	Scanning Digitizing Address-match geocoding
Data conversion	Importing/exporting Edgematching Clipping Raster/vector conversion
Geometric transformation	Translation Rotation Map projection Rubbersheeting
Generalization	Line thinning Line smoothing

Of course, when the map is being linked to an up-to-date, planimetrically correct foundation data set, rubbersheeting can partially compensate for positional errors in the source map. Distorting an inaccurate source map may be a good thing. Rubbersheeting is often required in order to integrate data with low or unknown positional accuracy with more accurate foundation data layers, for example, in linking the TIGER/Line files, with their variable positional accuracy, to a DOQQ base.

Another kind of coordinate transformation is needed when data layers are based on different map projections. ***Map projection transformation*** is the change in coordinates from one map projection to another. Data that come from different sources often utilize different map projections, so that coordinates must be reprojected for the data to overlay properly. All GIS have built-in functions for converting among common map projections.

Another common problem in creating and linking spatial data sets involves changing the geographic extent of the data set. The analyst may want to focus on one portion of the mapped area—for example, one municipality within a county—or to join maps together to create a map layer that covers a larger geographic area. In GIS, one can extract a portion of a mapped data set by cutting out the portion from surrounding areas and saving it in a separate file (Figure 3.13). These ***clipping*** or ***windowing*** procedures are easy and efficient in GIS where they can be done by using the cursor to define a rectangle or irregular shape.

Edgematching is a procedure for joining maps together by matching common features along the shared map boundary. For example, a particular road

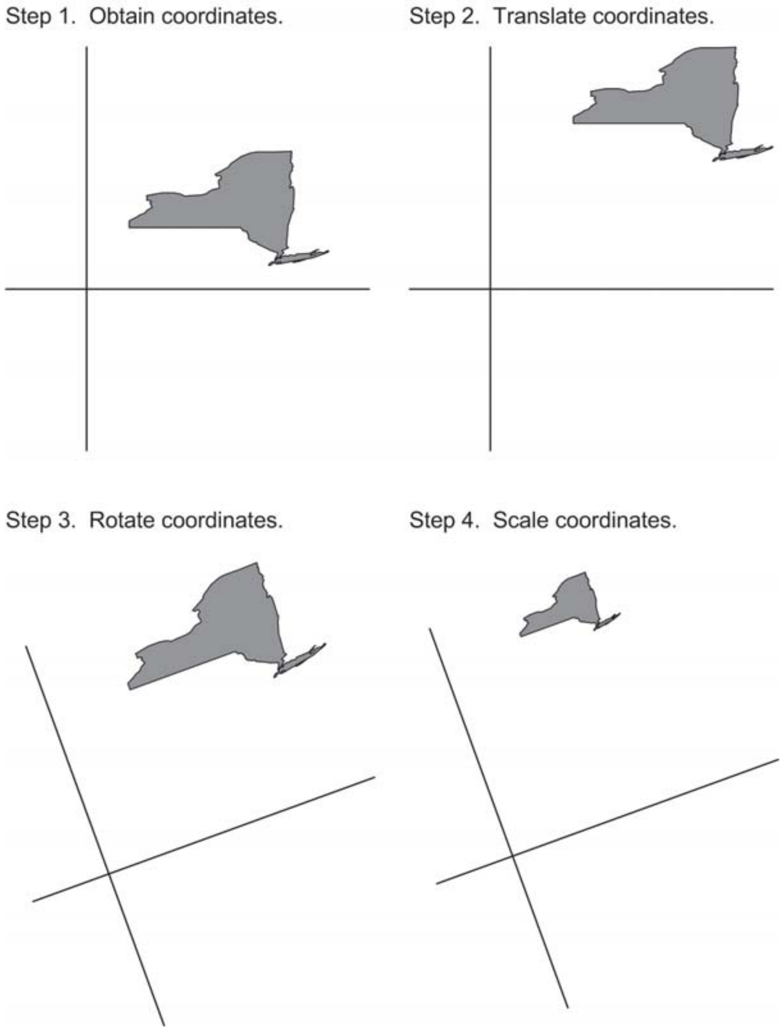


FIGURE 3.12. Coordinate translation of a spatial database of New York State.

appearing at the edge of one map is matched to its counterpart at the edge of the adjacent map (Figure 3.14). Most edgematching procedures also adjust features located in the area of overlap between the two adjacent maps to create a seamless map. For the procedure to work correctly, the maps being joined must have the same scale or resolution and contain comparable features. It is important to realize that edgematching creates a new topology for the database as, for example, two line segments are joined into one. The new topology enables new geographical analyses that address issues within the larger area.

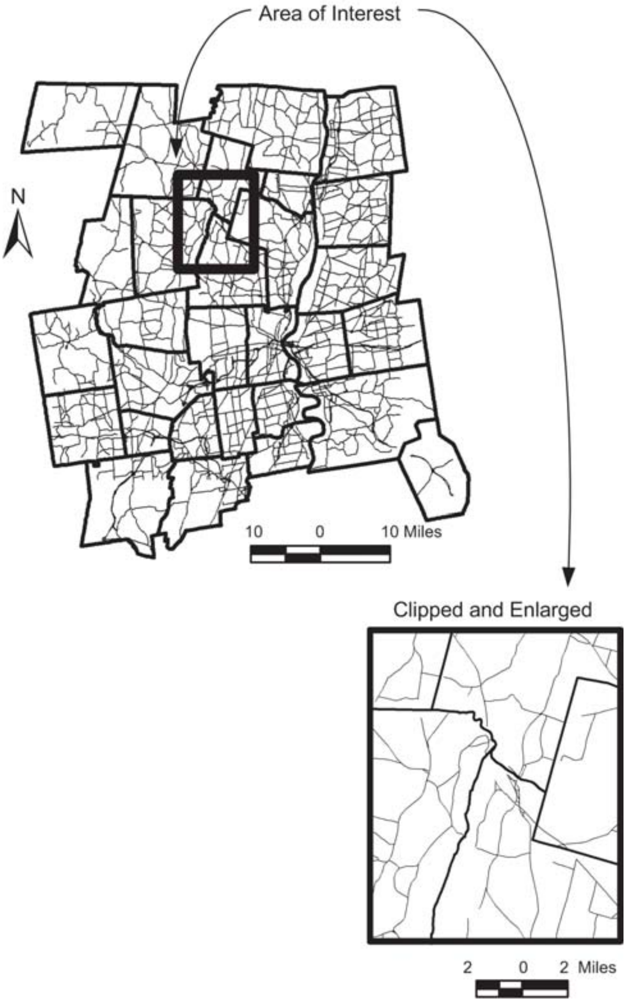


FIGURE 3.13. A window created around an area of interest can be used to "clip" the features of interest for viewing and analysis and for creating new databases containing just the clipped features.

Data Sharing

Assembling diverse spatial data sets and linking them with foundation spatial data is a time consuming, labor-intensive, and expensive process. The final product—an integrated ensemble of health, environmental, social, and foundation data—represents not only a major investment, but also a major resource, with value to other users analyzing issues in the same geographic area. ***Data sharing***, or the transfer of data between two or more organizations, offers many impor-

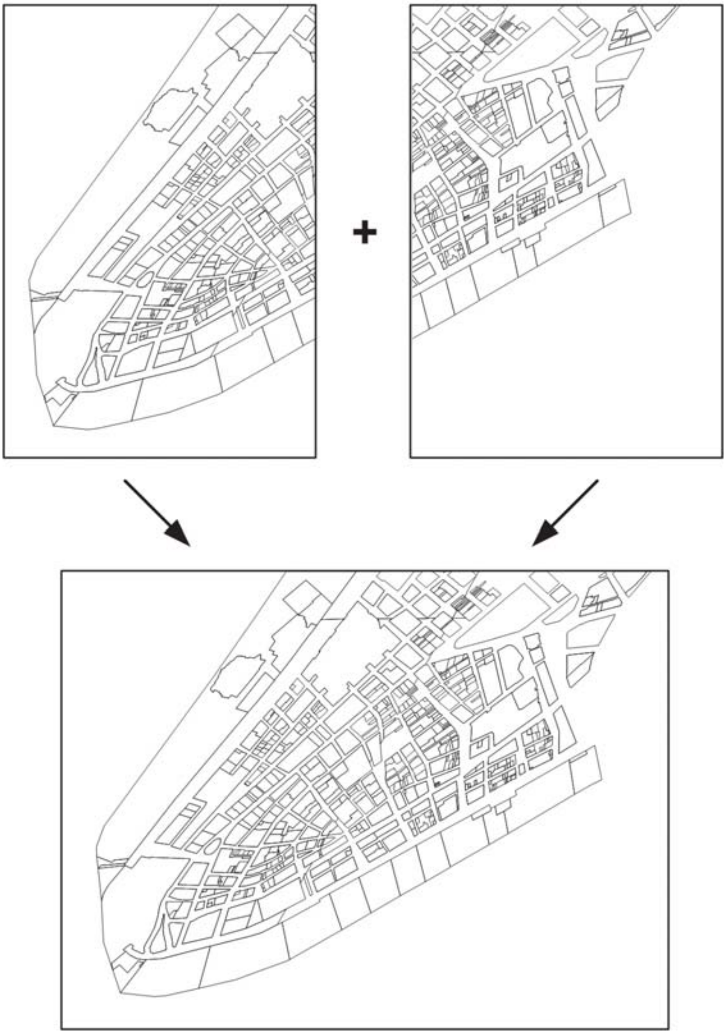


FIGURE 3.14. Two spatial databases of information for adjoining areas are joined by "matching" common features along the boundary to create a single seamless database.

tant benefits to the developers and users of geographic information (Onsrud & Rushton, 1995). The value of spatial data derives from its use, so enabling diverse groups to draw on the same data creates value by stimulating use. Data sharing is also a means for spreading the costs of database creation among multiple users and avoiding needless duplication of effort. Finally, there are often synergies in multiple use and analysis of a common spatial database. One group's insights spark another's, resulting in greater value overall.

Organizations at the regional, state, and national levels have increasingly recognized these benefits and taken steps to promote spatial data sharing. States have taken the lead by creating spatial data clearinghouses or unified state-level spatial databases. Most of these efforts involve extensive participation by local governments, which provide spatial data and draw upon it for local and regional planning purposes. At the federal level, the challenges of developing a ***national spatial data infrastructure*** in the United States over the last two decades have been acknowledged (Goodchild, Fu, & Rich, 2007; Craig, 2009).

Despite the many advantages of data sharing, technical and institutional barriers often get in the way. Sharing requires networked systems and agreements and common data formats that permit electronic exchange of information among users. Differences in hardware, software, and metadata standards impede spatial data sharing. As the volume of spatial data produced and used has grown, producers and users of data have needed to confront the legal implications relating to the dissemination and use of data. Issues of intellectual property rights, contract law, and liability affect data sharing in many countries (Cho, 2005).

More fundamentally, sharing requires cooperation among diverse institutions and branches of government and a shared sense of purpose. Differences in organizational needs, cultures, and interests make cooperation among organizations challenging at best (Obermeyer, 1995). Organizations often operate autonomously, emphasizing their particular needs and missions. In many cases, according to Craig (1995, p. 108), "agencies could share data, but they choose not to do so." Thus, data sharing is an inherently political process reflecting power, inertia, and access to resources. These political and institutional factors far outweigh the technical barriers to data sharing (Onsrud & Rushton, 1995).

An important barrier to sharing health data is the need to protect the privacy and confidentiality of health information. Many state and federal agencies and health care providers gather health data on individuals and are involved in data sharing. The development of health informatics, including electronic and personal health records, has raised additional privacy, confidentiality, and security concerns (O'Carroll, Yasnoff, Ward, Ripp, & Martin, 2003). A survey of public health professionals in Canada and the United Kingdom revealed that 71% identified privacy issues as an obstacle to public health practice (AbdelMalik, Boulos, & Jones, 2008). In the United States, the ***Health Insurance Portability and Accountability Act (HIPAA)*** of 1996 established standards for privacy of individually identifiable health information (U.S. Department of Health and Human Services, 2003). Geographic identifiers are included in the list of identifiers that must be removed to de-identify data so that an entity that passes the data to a third party can be held harmless under the law (Table 3.5). State departments of public health in the United States that use GIS are able to use individually identifiable health data in their work. Entities covered by HIPAA may also create limited data sets for use by other parties who enter into a data use agreement prior to the disclosure and use of the limited data. Methods for mapping individual geographic data to protect privacy are discussed in Chap-

TABLE 3.5. HIPAA Identifiers

Name
All geographic subdivisions smaller than a state, including: Street address and equivalent geocodes City and equivalent geocodes County and equivalent geocodes Precinct and equivalent geocodes ZIP Code and equivalent geocodes Except for: The initial three digits of a ZIP Code if, according to publicly available data from the Census Bureau; The geographic unit formed by combining all ZIP Codes with the same three initial digits contains more than 20,000 people; and The initial three digits of a ZIP Code for all such geographic units containing 20,000 or fewer people is changed to 000.
All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older.
Telephone numbers
Fax numbers
Electronic mail numbers
Social Security numbers
Medical record numbers
Health plan beneficiary numbers
Account numbers
Certificate/license numbers
Vehicle identifiers and serial numbers, including license plate numbers
Device identifiers and serial numbers
Web Universal Resource Locators (URLs)
Internet Protocol (IP) address numbers
Biometric identifiers, including finger and voice prints
Full-face photographic images and any comparable images
Any other unique identifying number, characteristic, or code, except as permitted

ter 7. There is growing awareness that advances in technology have made re-identifying health data easier and that alternatives to standard de-identification practices are needed.

Conclusion

This chapter has examined spatial data resources for public health GIS in the United States and geographical, technical, and institutional concerns in data integration. Investing in a GIS means investing in spatial data. Given the wide array of data sets available and the high costs of new database development, organizations need to assess carefully their spatial data needs and view development as a long-term investment rather than a short-term expense. As developers and users of spatial data, it is essential that public health organizations participate in the emerging efforts to create open, accessible, and integrated spatial data resources. Agencies need to plan how spatial data will be used internally, how to make it accessible to others, and how to promote spatial data sharing in partnership with other organizations.

CHAPTER 4

Mapping Health Information

Preparing and displaying maps of health information are among the most important functions of public health GIS. GIS offer a flexible, computerized environment that facilitates new forms of data exploration and analysis. One can easily pan across a map, zoom in on areas of interest, or query a database to examine areas or events of special concern. Health information can be linked with social and environmental features to examine geographical associations. The map, then, is just one product of a process of exploring, viewing, and analyzing spatial information. There is no perfect map; rather, each map is one of an almost infinite array of possible representations of spatial information. This chapter describes the procedures for displaying spatially referenced health information and preparing maps in GIS. After a general introduction to the mapping process, we discuss strategies for mapping health information. The next section considers how we can move beyond the map to view and explore health information. The final section addresses map design—creating a map for presentation or publication.

The Mapping Process

Advances in computer technology and GIS have fundamentally changed the process of mapmaking. Traditionally, maps were viewed primarily as tools for communication (DeMers, 2000). The main goal was to communicate information most effectively by carefully preparing a "finished" map. Issues of design and composition were paramount. Today's GIS and computer technology have turned this perspective on its head. The mapping process now emphasizes data visualization, exploration, and analysis, rather than the preparation of a finished map (Schuurman, 2004).

The mapping process brings together four key elements: the *spatial data* that are stored or entered into the GIS, as discussed in Chapters 2 and 3; the *representations* of that information on maps or computer screens; socially defined *queries* about the spatial data; and the *analysts* who create maps and respond to queries (Figure 4.1). The activities associated with mapping exist in and are shaped by the social and economic contexts in which maps are made (Elwood, 2006; Kitchen & Dodge, 2007). People and organizations involved in the mapping process decide what information will be analyzed, the types of queries to be made, and how the information is represented on a map. This politics of maps and mapping is discussed in Chapter 12 in relation to community-based health GIS.

The elements in the mapping process are interconnected. Queries—the questions asked about geographical issues and associations—play a central role. They define the kinds of data and information that are collected and analyzed, and how the analyst represents that information on a map or computer screen. Viewing the data spatially often leads to new and revised queries, and, thus, to new maps. Viewing data spatially may even provide an impetus for collecting new spatial data. Thus, the mapping process links data, representations, queries, and people in an iterative and fluid way. It has no clear starting point, no necessary "finished" product, and can lead to many different representations of the same information.

The use of mapping in examining the high rates of breast cancer in Long Island, New York, illustrates well this new mapping process (National Cancer Institute, 2011b). In the 1980s, many women in Long Island expressed concern about the high rates of breast cancer in their communities. Taking matters into their own hands, the women formed breast cancer coalitions, conducted surveys of breast cancer prevalence in their communities, and created simple "pin maps" of breast cancer prevalence. Their maps generated a host of hypotheses about the

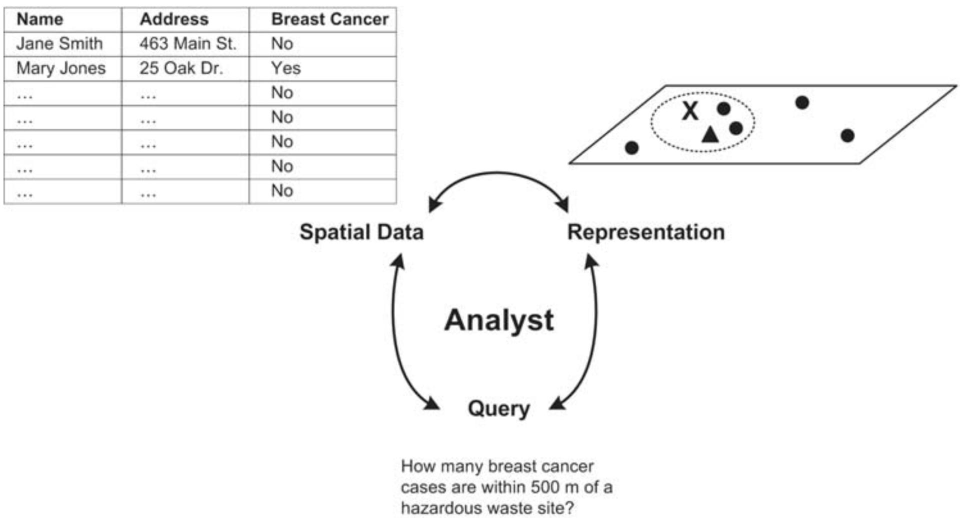


FIGURE 4.1. The mapping process.

links between environmental and social factors and breast cancer (Carlin, 2001). Some community groups plotted the locations of environmental features on their pin maps of breast cancer cases. At the same time, the New York State Health Department conducted a GIS-based investigation into the connections between breast cancer risk and exposure to hazardous industrial facilities and high traffic density (Lewis-Michl et al., 1996). Data from a case–control interview study of breast cancer were overlaid on maps showing the locations of industrial facilities and major traffic corridors, and the relationships between health and environmental factors were analyzed.

These investigations did not produce conclusive findings, but they laid the groundwork for a much larger GIS and epidemiological study of breast cancer and environment on Long Island, sponsored by the National Cancer Institute. An Internet-based health GIS was developed to give community groups and researchers access to a diverse array of health, social, and environmental data sets for the region (National Cancer Institute, 2011b). In addition, a wide variety of epidemiological studies were conducted to investigate the links between environmental exposures and breast cancer risk. Although these studies provided little to no evidence that environmental factors were responsible for the high rates of breast cancer (Winn, 2005), researchers continue to explore new possibilities.

This example is instructive because it illustrates the close ties between mapping, analysis, and data exploration. Many different groups, including grassroots community groups, researchers, local and state health departments, and federal agencies, have viewed and analyzed spatial data for Long Island. They have explored hypotheses, prepared maps, and created queries to represent their diverse perspectives and concerns about the links between environmental contamination and breast cancer risk. The interplay between people, social institutions, and mapping reveals maps as "emergent" rather than fixed (Kitchen & Dodge, 2007). Maps are constructed to solve problems in particular contexts. They are viewed and interpreted differently by different audiences, and they in turn shape audience understandings. Thus maps both reflect and reproduce wider social and political relations. In this dynamic mapping process, how people and organizations use GIS to represent and interact with data is critically important.

Representing Health Information

A key component of the mapping process is the representation of spatial information. ***Representation*** is the process of creating symbols to portray objects, quantities, or events. Maps are representations, as are the views on a GIS computer screen (MacEachren, 1995). Representations both illuminate and conceal information. The symbols on a map or display reveal features and associations, while features not represented by symbols are hidden from view. No map could possibly represent the richness of the earth's surface and its inhabitants, so, as Monmonier (1996, p. 1) writes, "the map must offer a selective, incomplete view of reality."

Representing health information on maps requires the intelligent use of symbols. Symbols convey information. They reveal what is important, show contrast, and identify patterns. As discussed in Chapter 2, the six ***visual variables***—size, shape, orientation, spacing, color hue, and color value—differentiate symbols on maps (Bertin, 1979; Slocum et al., 2009). The symbols and mapping strategies used in representing health information vary according to the type of spatial information that is to be displayed. Point data are often shown on dot or point symbol maps, area data on choropleth or dot density maps, and linear data on network or flow maps. These are not rigid choices, however. A single map or view can combine all three types of information to show complex features and patterns; point symbols can effectively show certain kinds of area data. These different mapping approaches are discussed in the sections that follow.

Representing Point Information

Much health information consists of point locations—hospitals, residences of people who experience particular health problems, accident locations, and haz-

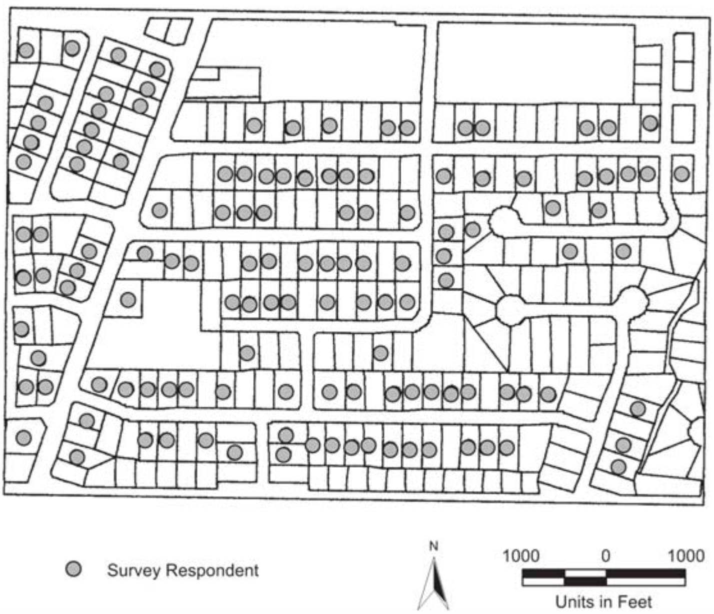


FIGURE 4.2. A point symbol map showing residential locations of survey respondents.

ardous facilities—that can be represented effectively on dot or point symbol maps. On a ***point symbol map***, point symbols correspond to one or more events, and concentrations of point symbols reveal clusters of events (Figure 4.2). As the number of points varies across the map, the viewer senses the general pattern of density change. Point maps are useful devices for viewing health information. As noted earlier, during the 1990s, many community-based breast cancer groups in Long Island gathered data on the residential locations of women diagnosed with breast cancer (Carlin, 2001). Point symbol maps of geocoded locations were used to search for clusters of breast cancer and to generate hypotheses about links to environmental hazards.

A challenge in using point symbol maps to represent spatial patterns of disease is that differences in dot density may simply reflect differences in risk population. A cluster of breast cancer cases may coincide with a cluster of women whose age and sociodemographic characteristics place them at risk of breast cancer. One way to address this issue is to display both the locations of people diagnosed with a disease and those at risk by using contrasting point symbols. Differences in density reveal clusters of one group relative to the other (Figure 4.3).

Creating point maps involves the careful choice of symbols. Symbols differ in size, shape, hue, and the other visual variables. Varying symbol size is one way to create contrast and show quantity. In a ***proportional symbol map***, symbol size is proportional to the number of events at a place. Large dots are highly vis-

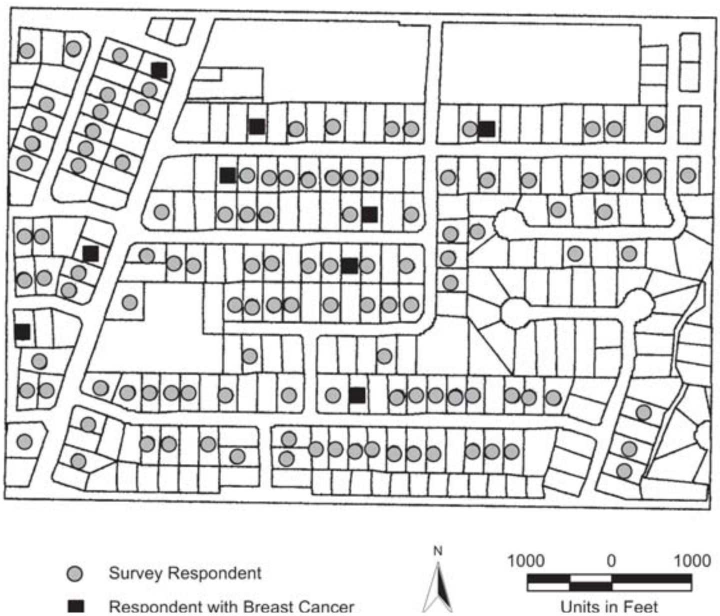


FIGURE 4.3. Use of contrasting point symbols to differentiate respondents.

ible, representing large concentrations of events. Shape also distinguishes symbols, making them easier to perceive and understand. Symbols can be ***geometric***, involving simple geometric shapes like circles or squares, or ***pictorial***, involving simple pictures such as houses or churches (Fry, 1988). Geometric symbols are generally easier to read and distinguish on maps, but are not as legible as picture symbols.

Color hue and value are also important devices for differentiating symbols and other map features. A collaborative research project between the National Center for Health Statistics and geographers at Pennsylvania State University examined the effectiveness of color for representing public health data on maps (Brewer, MacEachren, Pickle, & Herrman, 1997). They concluded that "color is clearly worth the extra effort and expense it adds to map making because it permits greater accuracy in map reading" (p. 434). The use of color is discussed in more detail in the section on choropleth mapping.

The map in Figure 4.3 combines shape and color to create contrast. Small, circular gray dots represent women over age 25, and black squares identify women diagnosed with breast cancer. The darker shade and contrasting shape draw the eye to the places where women diagnosed with breast cancer live, while the smaller circles form an almost continuous distribution in the background.

When the density of dots is high, dot maps can become cluttered and difficult to interpret. Dots hide other dots, obscuring differences in density. In these situations, the analyst can use proportional symbols to distinguish areas of high and low concentration. Proportional symbols maps can even become cluttered, however, and large symbols are not geographically precise. In these cases, the analyst can use a density estimation method, like kernel estimation, discussed in the next chapter, to create a continuous representation of point density. Another strategy is ***area conversion*** in which points are grouped into geographic areas, and the areas are shaded according to the number of events within them.

Some kinds of health events, such as motor vehicle collisions, are clustered along roads or other line segments. In displaying this information, policymakers often want to know the density of events by line segment—for example, which roads or street segments have the most collisions. To represent such information, we can shade line segments based on numbers of events. Varying the color or width of the line segment highlights differences in number or density of events. Figure 4.4 shows a conventional point symbol map of motor vehicle collisions in part of Connecticut and a corresponding shaded line segment map of the same information. The second map clearly reveals roads and intersections with high collision frequencies.

The GIS analyst must experiment with combinations of symbol size, shape, and color, as well as the overall mapping strategy, to create an effective map representation of point information. Most GIS systems make this experimentation relatively easy, offering an array of options for symbol design. One can create alternative layouts in the GIS and view them on the computer screen, making adjustments as needed. The map or map series is published in its final version only after choosing the most effective design.

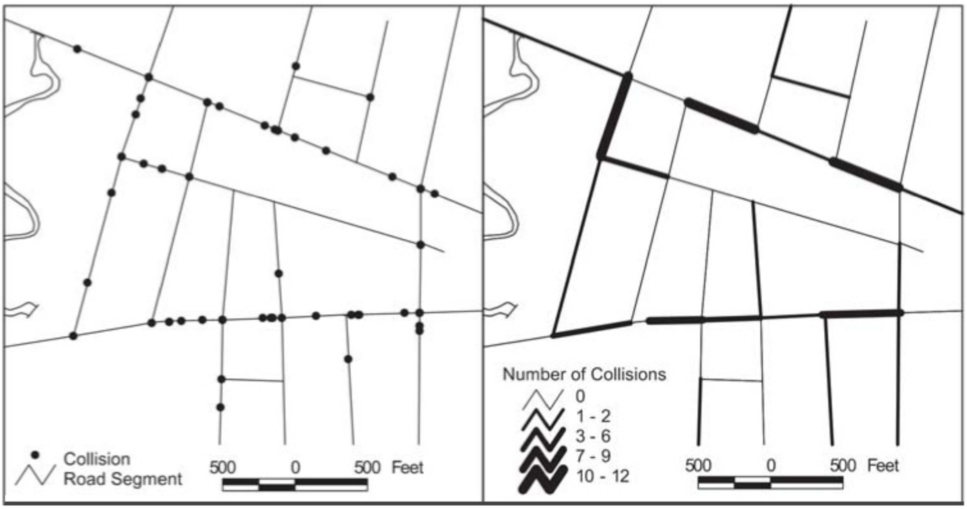


FIGURE 4.4. A point symbol map of motor vehicle collision locations on the left was converted to a line symbol map as shown on the right.

Representing Area Data

Health information is often available for areas—ZIP Codes, counties, states, or countries—that form a template for representation. Area health data are spatially "filtered" with respect to predefined zones and thus are dependent on the zoning system used. In representing area information, we fill areas with symbols, colors, or patterns to show the intensity or number of events within the area. ***Dot density maps*** use point symbols within areas to depict numbers of events in the corresponding areas. In ***choropleth maps***, areas are shaded with different colors, patterns, or intensities to display place-to-place variation. These kinds of maps are widely used in health mapping for several reasons. First, much health and demographic data are only released by area, and, second, by not showing precise locations, such maps avoid concerns about privacy and confidentiality. However, area-based maps cannot display the detailed geographic patterns that emerge from point symbol maps, or maps that show the underlying spatial organization of street networks, residences, and other landscape features. They give the impression of uniformity *within* areas and show sharp changes *between* areas, when in fact the underlying distribution may change gradually or continuously. They often, therefore, arbitrarily partition the underlying distributions of cases and population, affecting the validity of rate calculations for the mapped areas.

The appropriate mapping approach depends on the type of area data being represented. If the data refer to counts of events, people, or facilities by geographic area, dot density mapping is generally the preferred approach. Dots are arranged randomly within areas, and the viewer perceives differences in numbers in the changing patterns of dot density (Figure 4.5). The number of dots

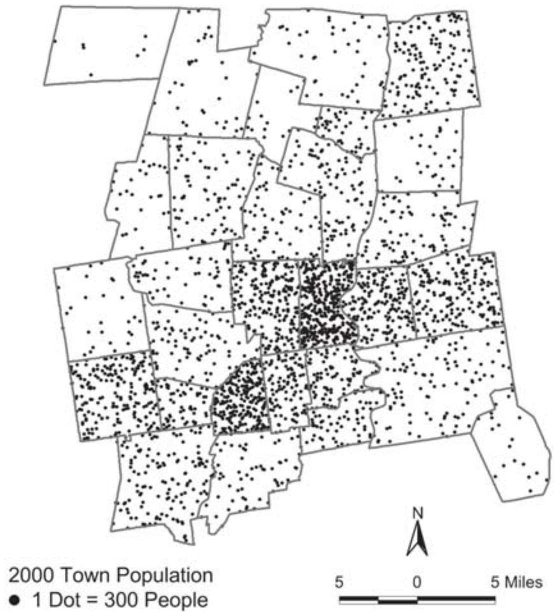


FIGURE 4.5. A dot density map of population distribution by town in Hartford County.

is proportional to the number of events, say, one dot representing 300 people. Typically dots are arranged randomly within areas to avoid sharp breaks in the dot pattern at area boundaries.

CHOROPLETH MAPPING

More commonly, area health data refer to rates or ratios or other statistics that apply to areas. In these situations, choropleth mapping is the preferred approach. In a choropleth map, the data values that fall within a specific class interval are assigned a unique color, shade, or pattern. Differences in intensity are visible in the varying colors or patterns across the map.

Class Interval Selection. A key issue in choropleth mapping is the choice of class intervals. Changing the class interval scheme can fundamentally change how the map looks and the message it sends. Most GIS offer the mapmaker a range of options for defining class intervals. A common one is ***equal interval classification***, in which the range of the data values (maximum value – minimum value) is divided into a fixed number of classes (Figure 4.6). Each class represents an equal interval of possible data values. Although this method works well for some data distributions, it performs poorly if there are extreme values in a highly skewed data distribution. In this case, the vast majority of areas will fall

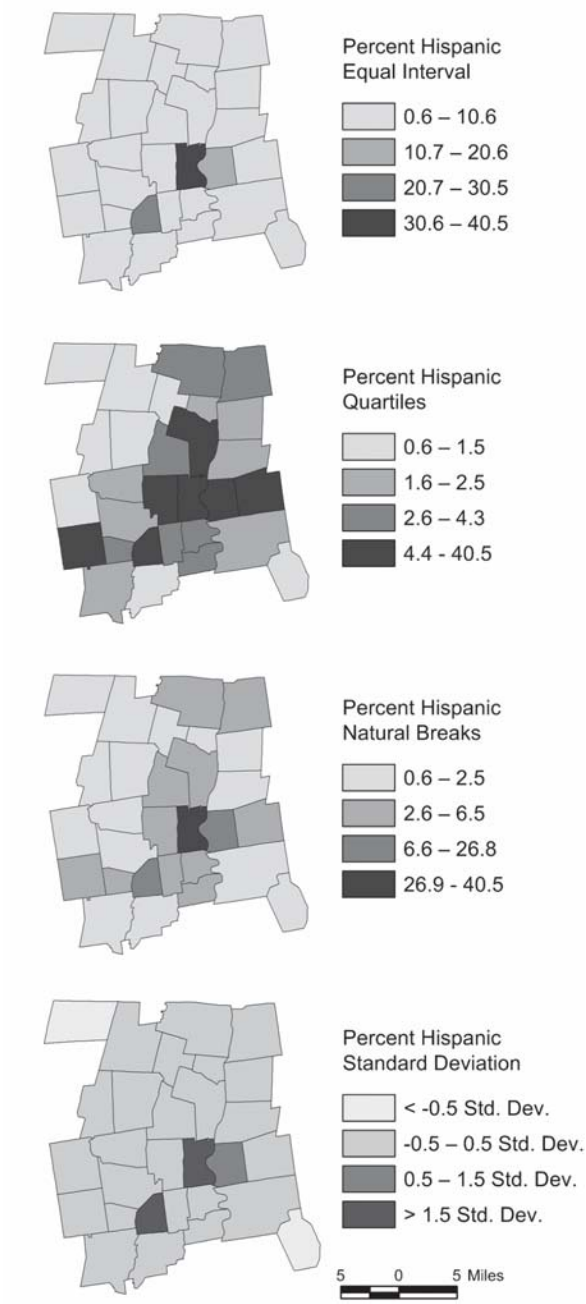


FIGURE 4.6. Choropleth maps of the same data created using different methods for determining class intervals.

into one or a few classes, while other classes remain empty. The map will show little spatial variation in such instances, because the majority of observed values are very similar. The fact that there are empty classes may be important in understanding the data distribution. The equal interval approach was suggested by Becker (1994) as a suitable approach for developing classifications to facilitate comparison of maps.

Alternatively, the ***quantile classification*** method (more generally known as the n-tile method) creates an equal *frequency* of values in each of n classes. Values are ranked from low to high, and the ranked values are divided into n classes each containing an equal number of values. For example, in the quantile method using quartiles (where $n = 4$), a data set of 100 values will be divided into four classes with 25 values each. The lowest class will contain the 25 smallest values—the first quartile—and so on. This method ensures that each class is equally represented on the map, but the results can be misleading because areas that have very similar data values may be assigned to different classes. These areas will appear quite different on the map, even though their actual data values are quite close. In addition, areas that have very dissimilar values may be assigned to the same class. The quantile method tends to perform better than the equal interval method for highly skewed data distributions, because it differentiates values in the bulge of the data distribution. However, the method typically does not produce intervals that are similar in size. The first class might include data values ranging from 0–2, the second from 3–24, and the third from 25–110. Logically one would infer that the classes represent similar data ranges when in fact they do not. Quantile maps can be highly misleading unless the viewer carefully consults the map legend. If the data distribution is uniform, there will be no difference between the equal interval and quantile classifications.

The ***natural breaks*** method searches for ***breaks*** in the data distribution, that is, natural divisions among groups of data values. Classes represent real clusters of data values, and class breaks separate the clusters. The advantage of this method over the previous one is that it does not arbitrarily divide observations that have similar values into different classes. However, it produces class intervals that are neither of equal width nor with equal numbers of observations, so the results are unpredictable and totally dependent on the data. In some GIS software, natural breaks are determined statistically as breaks that minimize within-class variation (Jenks & Caspall, 1971; Jenks, 1977).

The **z*-score classification method*** is often used when choropleth maps are compared. For a particular variable, each data value is transformed to a z-score that represents the value's deviation from the mean measured in units of standard deviation. A z-score of -1.2 indicates that the data value is 1.2 standard deviations below the mean. In classifying z-score maps, we typically create an odd number of class intervals, with the middle interval centered on zero. Intervals of equal size can be constructed, or one can define intervals based on z-score statistical significance levels. Z-score maps are good for visual comparison because the maps display each variable in a common metric.

For a fuller discussion of the strengths and limitations of the many methods that exist for defining class intervals, we urge the reader to consult a good cartography reference such as Brewer and Pickle (2002) or Slocum et al. (2009), or the ChoroWare software toolkit for choropleth map classification, which is available online (Xiao, Armstrong, & Bennett, 2005).

Why does the approach to data classification taken matter to the public health analyst? Monmonier (1996) points out that many choropleth maps can be made from the same data simply by changing the class intervals. It is possible to manipulate map readers' impressions of spatial patterns of health events, simply by changing class intervals (Figure 4.6). Viewers' perceptions and understandings of the patterns and data displayed on choropleth maps differ depending on the map classification method used (Brewer & Pickle, 2002).

An innovative method for choropleth mapping that addresses the problem of defining class intervals is the ***classless choropleth map*** (Tobler, 1973). Instead of defining a fixed number of class intervals, areas are shaded on a continuous value scale with shading intensity or value proportional to the actual data value observed for the geographical unit (Figure 4.7). A detailed legend is required to show the actual data values associated with each shade. These maps are visually effective, and they avoid somewhat arbitrary decisions about class interval selec-

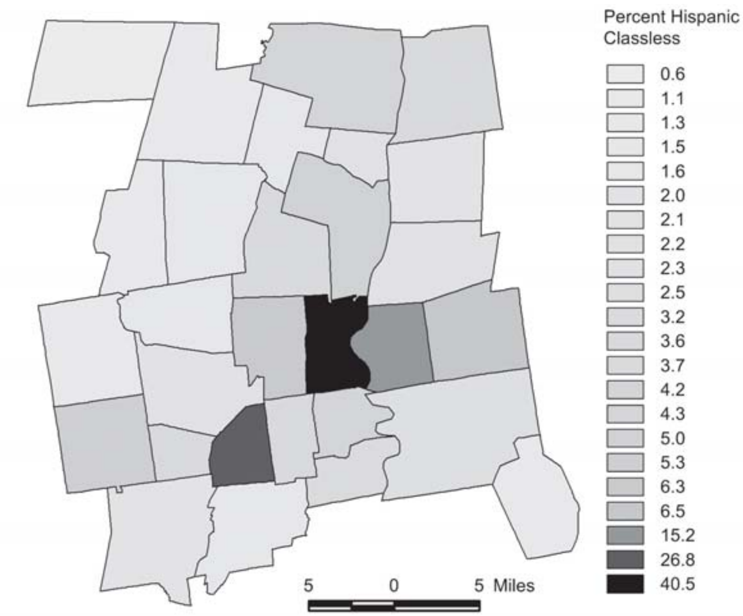


FIGURE 4.7. In a classless choropleth map, continuous shade tones of a single hue correspond monotonically to unique data values within the distribution being mapped.

tion. Classless choropleth maps can now be easily produced using downloadable software such as MAPresso (MAPresso, 2011).

The classless map attempts to address one of the major drawbacks of choropleth map classification schemes like the equal interval, quantile, and natural breaks methods. The drawback is that these classifications are derived from the univariate distribution of values and ignore spatial relationships among the units for which data are being mapped. As a result, statistical "cliffs" in the data distribution may not match with visual "cliffs" or differences between places on the map.

Methods for defining optimal class intervals that take into account statistical distribution and spatial relationships have been developed for interval/ratio and for ordinal level data (Cromley, 1996; Cromley & Mrozinski, 1999). Cromley and Cromley (1996) developed a classification method that maximizes spatial similarity among contiguous units in the same class interval and applied the method to data published in the *German Cancer Atlas*. Although the ability to generate a matrix that describes spatial contiguity of the mapping area is important in implementing these methods and GIS systems can produce this type of matrix, most GIS systems, as yet, have not incorporated these approaches to map classification in their standard options for classifying data for choropleth mapping.

Sometimes a series of choropleth maps is constructed to compare the spatial distribution of a health indicator over time or among population groups. When a series of maps is being prepared, the choice of class intervals depends on the purpose of the map comparison. Using the same class intervals for each map in the sequence facilitates direct comparison of maps for different points in time or for different groups. Comparing maps is straightforward, because color tones on each map represent equivalent data values. It is easy to see whether the health outcome of interest has increased or decreased over time. Grauman, Tarone, Devesa, and Fraumeni (2000) created a sequence of maps showing cervical cancer mortality in the United States for different time periods. The map sequence shows the dramatic decline in mortality during the 1960s and 1970s, a result of increased cancer screening that led to increased detection of the disease at an early stage (Grauman et al., 2000).

If the purpose of the map series is to compare geographic inequality among maps, using different class intervals for each map makes sense. The class intervals adjust for differences in the average value of the phenomenon among time periods, revealing geographic variation or inequality within each map. Maps of cervical cancer mortality over time using different class intervals indicate shifts in the geographic location of high and low mortality regions in each time period (Grauman et al., 2000).

However, when different class intervals are used in a series of maps, care must be taken in map interpretation. For example, maps of male and female lung cancer that use different class intervals may give the impression that the disease rate associated with the "red" area on the map displaying female lung cancer rates is as high as the rate associated with the "red" area on the map displaying male lung cancer rates when, in fact, the highest rates among females are much lower than the higher rates observed among males (Becker, 1994).

Color and Choropleth Mapping. Another important issue in choropleth mapping is the choice of colors used in area shading. Colors have three dimensions: hue, saturation, and value. ***Hue*** is described by the words we commonly associate with color tones—for example, red, blue, purple. ***Saturation*** refers to the dominance of hue in the color, its vividness, and ***value*** refers to its lightness or darkness. Generally speaking, the best practice is to use a single hue and vary its saturation. Less saturated hues typically represent areas with low data values, and more saturated hues, higher data values. However, when high data values predominate, one may want to reverse these gradients: a general principle is to avoid maps that show large areas of highly saturated color (Tufte, 1990). Most GIS include standard color ramps for choropleth mapping, but other schemes are possible, generating a rich set of options for color selection (Zeileis, Hornik, & Murrell, 2009).

In some health-mapping projects, the goal is to show data values in relation to some average or population value—for example, mortality rates above or below the national norm or *z*-score values. For these types of maps, a ***diverging color scheme*** is often highly effective. One hue is used to represent data values above the norm and another hue represents data values below the norm. For each hue, differences in saturation or value reveal the difference from the norm. The ColorBrewer 2.0 website is a great tool for exploring and mapping both single hue and diverging color schemes (Harrower & Brewer, 2003). Shifting from one color scheme to another, the viewer can choose the scheme that represents the data clearly and effectively. Guidelines for selecting an effective scheme are included.

Choice of color also depends on how and where the map will be viewed and by whom. Many colors wash out when viewed via LCD projector; others are difficult to reproduce in print; some color schemes are confusing for people with color-blindness. The ColorBrewer 2.0 website rates the usability of different color schemes in relation to these concerns.

Constructing Legends for Choropleth Maps. The ***legend*** on a choropleth map links the graphic symbols (e.g., colors) on the map to the corresponding data values. Traditional choropleth map legends show the range of data values associated with a particular graphic value (Cromley & Cromley, 2009). Standard cartography texts provide guidance on designing and labeling legends for choropleth maps using different map classification methods (Slocum et al., 2009). The traditional legend provides essential information for map interpretation, but it does not convey the distribution of data values within class intervals. To address this issue, Kumar (2004) proposed the use of a ***frequency histogram legend*** that shows the frequency distribution of values in relation to class intervals (Figure 4.8a). Although the frequency histogram legend is a clear improvement

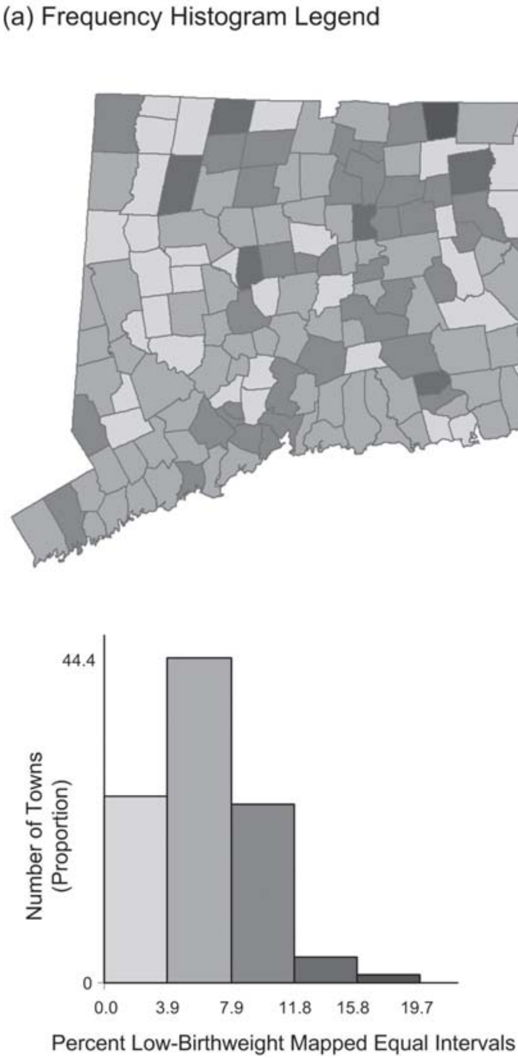
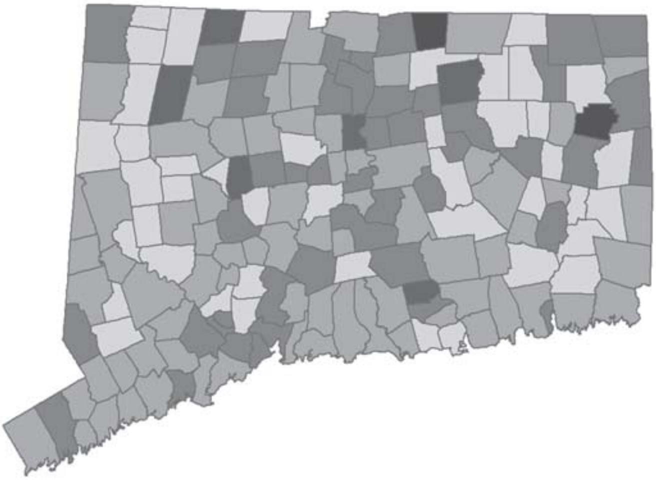


FIGURE 4.8. A choropleth map of low-birthweight rates for towns in Connecticut using equal interval classification and five classes. Figure 4.8a shows the map with a frequency histogram legend. The legend shows that almost half of the towns fell in the 4.0 to 7.9% low-birthweight class interval. Figure 4.8b shows the same map with an ogive or cumulative frequency legend. The cumulative frequency curve shown on this legend would remain the same, regardless of the classification method or number of class intervals used in the map of these rates for the 169 towns of Connecticut.

(b) Cumulative Frequency Legend



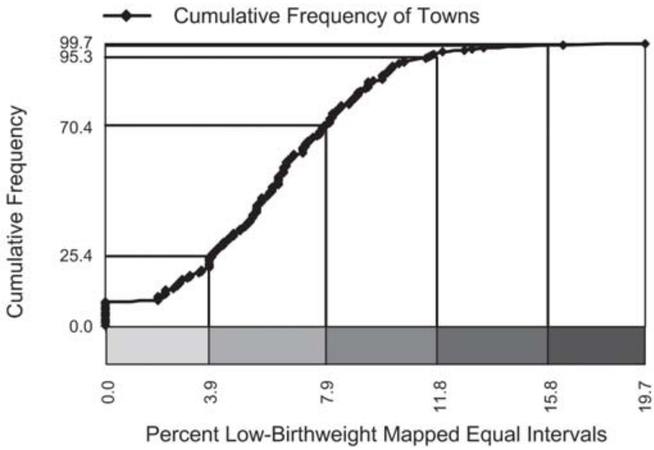


FIGURE 4.8. *(cont.)*

over the traditional choropleth map legend, it still involves grouping data values into intervals so that the appearance of the legend will differ based on the number and widths of intervals used. In response, Cromley and Ye (2006) developed a ***cumulative frequency (ogive) map legend***. This legend is constructed by first ranking the data values in ascending order and plotting the corresponding cumulative percentage of total value along the *y*-axis. Color tones associated with each value are displayed on the *x*-axis (Figure 4.8b).

For health data, an advantage of cumulative frequency legends is their ability to display both numerator and denominator values associated with health/disease rates. The ***numerator/denominator cumulative frequency legend*** (Cromley & Cromley, 2009) shows both sets of values in a single legend. It starts by placing areas in rank order according to incidence rate and plotting the cumulative percentages for the respective numerator and denominator values. Thus, the cumulative frequency distributions for the numerator and denominator are superimposed. A wider gap between the two cumulative distributions represents greater spatial inequality in health/disease rates. Figure 4.9 presents a choropleth map of low birthweight by town in Connecticut with a numerator/denominator cumulative frequency legend.

To summarize, health data analysts using GIS to prepare choropleth maps need to investigate how data are distributed through their range, how data are distributed spatially across the choropleth units, and which class breaks have particular substantive value for the analyst and the viewer. Choosing an appropriate color scheme and map legend helps improve the map's legibility. Analysts can take advantage of the relative ease of developing a series of maps using different classification, legend, and symbolization approaches to evaluate the impacts of these choices on the message the map conveys. GIS enables public health analysts to produce multiple views of the data in the form of different types of maps and charts to support more effective analysis of data and communication of results.

THE MODIFIABLE AREA UNIT PROBLEM

In mapping and analyzing area data, analysts also need to be aware of the ***modifiable area unit problem***—the impact of the location and configuration of areal units (zones) on the outcomes of the analysis (Openshaw, 1984). In a sense, there is no "true" choropleth map. The map's appearance and the message it conveys vary depending on the size, number, and configuration of area units. For example, a map showing cancer cases in Illinois by county will look different from a map showing the same data by census tract.

The modifiable area unit problem has two dimensions: the zoning (boundary) effect and the scale effect (Wong, 2009). The ***zoning effect*** refers to the impact of the locations of area boundaries on the appearance of a choropleth map and on the analysis of area data. For health data, the location of area boundaries in relation to the underlying distribution of health events and population is important. Depending on where boundaries are located, zones can divide clusters of health events or concentrate them in a single zone. For centuries, politicians have exploited the boundary effect to "gerrymander" electoral districts, purposely drawing boundaries to achieve a desired electoral outcome. Monmonier (1996, p. 158) demonstrates how this works with an example based on John Snow's famous map of cholera in London (Figure 4.10). When Snow's point data are aggregated to zones, the geographical pattern of cholera varies greatly depending on how zone boundaries are drawn. On some maps, the cluster of

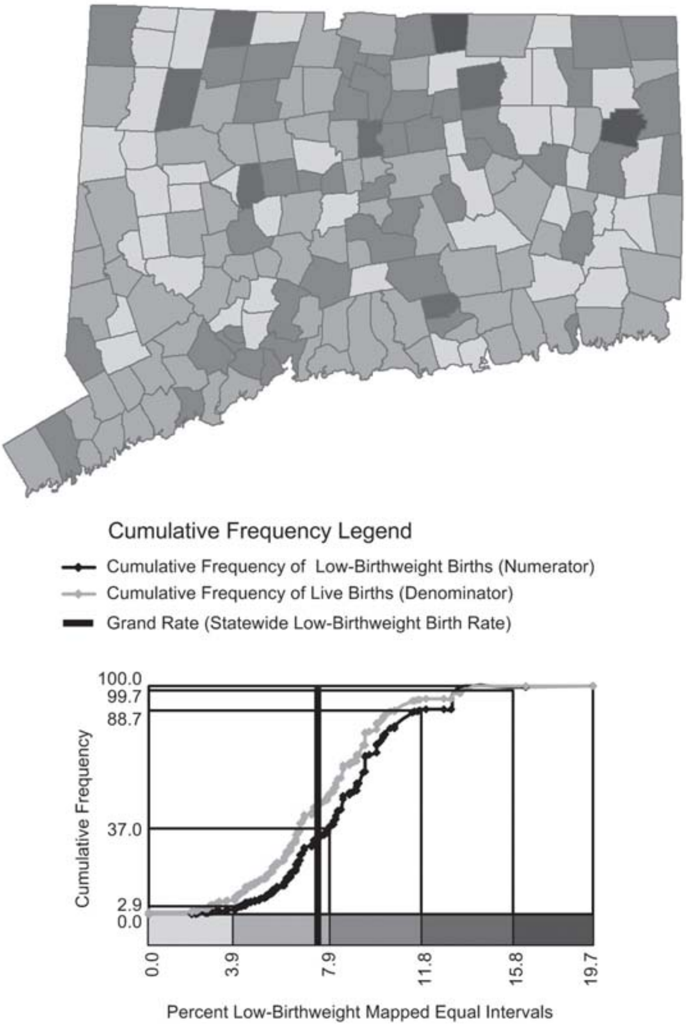


FIGURE 4.9. A choropleth map of low-birthweight rates for towns in Connecticut with a numerator/denominator cumulative frequency legend. The towns in the second and third class intervals have low-birthweight rates from 3.9 to 11.8 %. The cumulative frequency curve of low-birthweight births (the numerator curve) shows that 85.8% of all low-birthweight births occurred in towns in the second and third intervals. From Cromley and Cromley (2009). Originally published by BioMed Central in the *International Journal of Health Geographics*, Open Access.

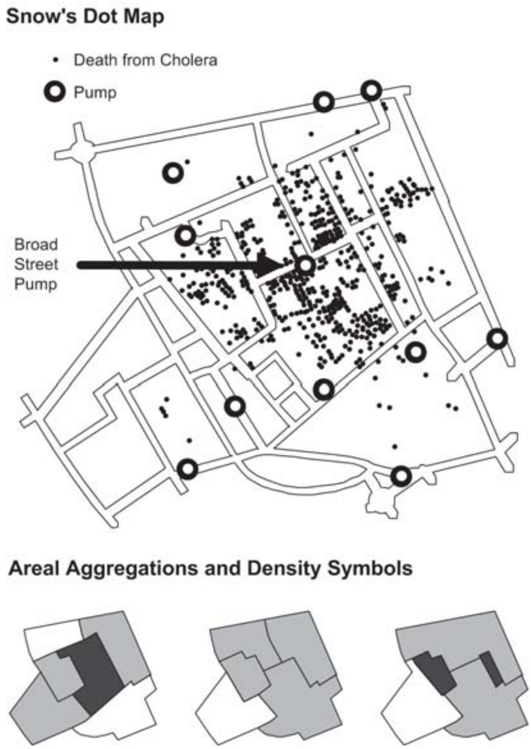


FIGURE 4.10. A reconstruction of John Snow's map of cholera cases in London and three choropleth maps produced by different areal aggregations. From Monomonier (1996). Copyright 1996 by University of Chicago Press. Reprinted by permission.

cases around the Broad Street pump disappears. This issue is discussed in relation to vector-borne disease in Chapter 8.

Choropleth maps and analytical results also vary with the number and sizes of areal units—the spatial scale of data. This is known as the ***scale effect***. Small areas capture the underlying pattern of health events, showing fine-grained variation over space. In contrast, large areas conceal local differences, reducing the variation in values over space. A county-scale map cannot show differences among towns and neighborhoods, for example, and state-level data hide disparities across counties. The scale of areal units affects our perceptions and understanding of health inequalities while viewing choropleth maps (Figure 4.11).

Choropleth maps also can mislead the viewer by giving a false impression of equivalence among areas. The fact that each area has a data value associated with it implies that the areas are comparable in size and significance. A choropleth map by state gives equal significance to the health statistics for Wyoming and California despite the vast difference in state populations. Furthermore, less populated areas are often larger in size, attracting attention when shaded on the

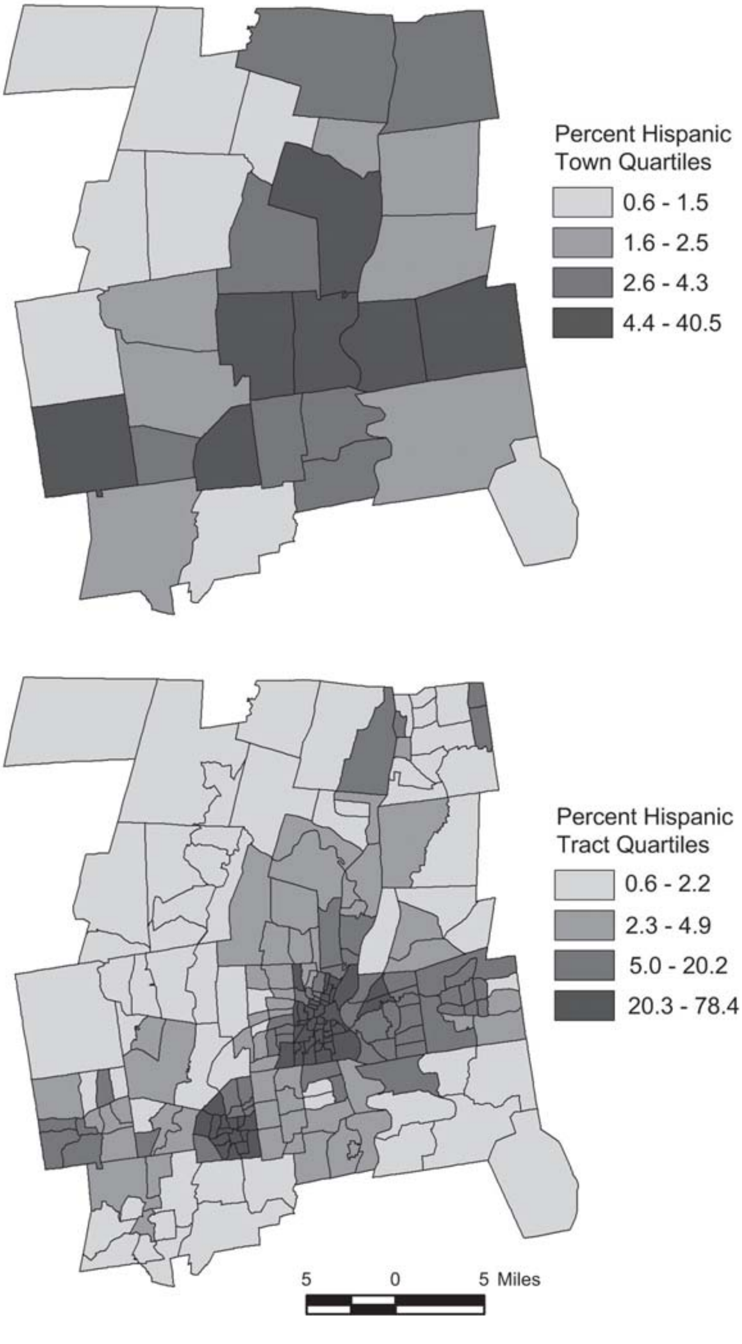


FIGURE 4.11. Maps of Hispanic population by town and census tract show different patterns of spatial variation using the same quantile classification method. Spatial variation is greater in the census tract map.

map. On a state map, Montana looks much more prominent than Rhode Island, though its population is less.

To tackle this problem, one can create a ***cartogram*** in which the sizes of areas are proportional to their populations. Places with large populations are expanded in size and appear large on the map. Danny Dorling and colleagues have used cartograms very effectively to depict global health disparities. Figure 4.12 shows a cartogram of infant mortality rate by country, and the size of the country on the cartogram is proportional to the number of infant deaths. The stark concentration of infant deaths in Africa and Asia literally jumps out from the map. Cartograms effectively convey the magnitude of health disparities, but they can be difficult for viewers to interpret. By distorting traditional geographic space, cartograms remove the familiar geographic reference that frames viewers' understandings; however, standard choropleth maps also distort when areas differ greatly in population size.

When mapping area data, it is essential to analyze the effects of the modifiable area unit problem on maps and results. By using small-area data, one can always show finer-grained and more detailed spatial patterns than can be shown with large-area data. It is often the case, however, that the analyst only has access to large-area information. How can this problem be addressed? One approach is to use ***ancillary data*** to estimate variation within large areas. Ancillary data describe geographic features that constrain the distribution of risk population or health events. As discussed in Chapter 6, ancillary data can be used to allocate events within large areas to subareas in which the events are most likely to occur.

Another approach is to apportion data for large areas based on related data for smaller areas. This is the problem of ***areal interpolation***, of estimating values

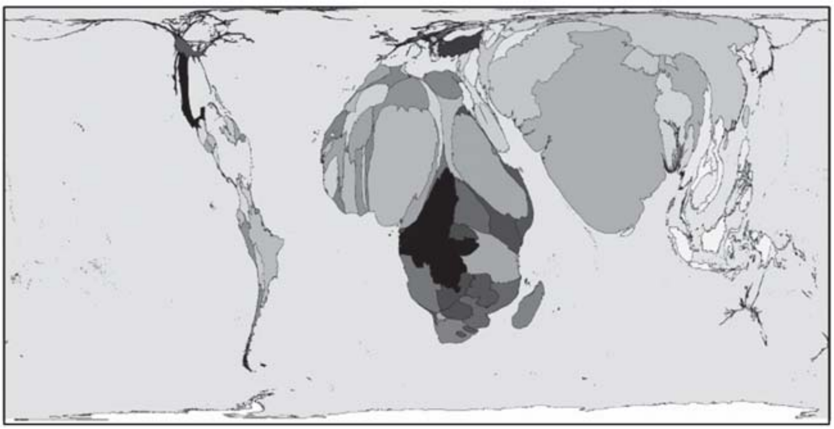


FIGURE 4.12. A cartogram of infant deaths. The sizes of the countries show the proportion of infant deaths worldwide that occurred in the country. From Dorling and Barford (2007). Copyright 2007 by the World Health Organization. Reprinted by permission.

for areas (Goodchild & Lam, 1980). For example, assume that we know from hospital discharge data, the number of children by ZIP Code who were hospitalized for asthma. Data on risk population, in this case the number of children, are available for much smaller geographic units, census blocks. If the blocks nest perfectly in ZIP Codes, we can apportion the asthma data for a given ZIP Code to blocks within that ZIP Code based on the proportion of children residing in each block. This method assumes that the risk of hospitalization for asthma is constant within a given ZIP Code. If the small areas do not nest perfectly, the most common approach is to allocate the population of split blocks to corresponding ZIP Codes based on the area of overlap—***the area weighting method*** (Haining, 2003). The area weighting method is discussed in detail in Chapter 6 and displayed in Figure 6.13. More complex methods that rely on additional ancillary information are discussed in Flowerdew and Green (1994) and Hawley and Moellering (2005).

By using ancillary data, we can move from small-scale data to a larger-scale, more detailed representation of spatial variation. Shifting from one geographic scale to another is important in analyzing many kinds of public health issues, for example, environmental health and access to health care, as discussed in Chapters 6 and 9. GIS greatly facilitate the use of ancillary data and the matching of data across geographical scales. Defining areas of overlap and allocating data from one layer to another are standard operations in most GIS.

Despite their advantages for representation, small-area health data pose "large statistical problems" (Diehr, 1984). Often there are few health events in each small area, making maps and estimates unreliable. This is especially true if the data cover a short time period or a rare disease. This small numbers problem is discussed further in Chapter 5.

Viewing Health Information

The ability to visualize and explore health data interactively is a main advantage of GIS in public health analysis. A ***view*** is a graphic representation of data. It is the part of the computer display board that one can see on the computer screen. The extent of the view is always less than or equal to the addressable space in a data set. Simply put, one cannot display a larger geographic area than one has data for. In GIS, the spatial objects in the view and the tables of attributes describing them can be directly linked. The analyst can access the two together and explore the relationships among attribute data in the table and the spatial representation of that data in the view.

Public health analysts will typically approach a database with two types of questions. What are the health problems of interest, and where do they occur? Where are the places of interest, and what kinds of health problems occur there? Many public health organizations are organized functionally to address specific kinds of health problems: maternal and child health, infectious disease, or injury. GIS users in these settings will have already established what kinds of health problems are of interest and may want to use GIS to gain a better understanding of the geographical distribution of these health problems. For example, we might ask, "Where are the residential locations of all children born with birth defects in Arizona?" Here the attribute (children born with birth defects) is given, and we want to know where the events that have these attributes occur. We call this "viewing by attribute."

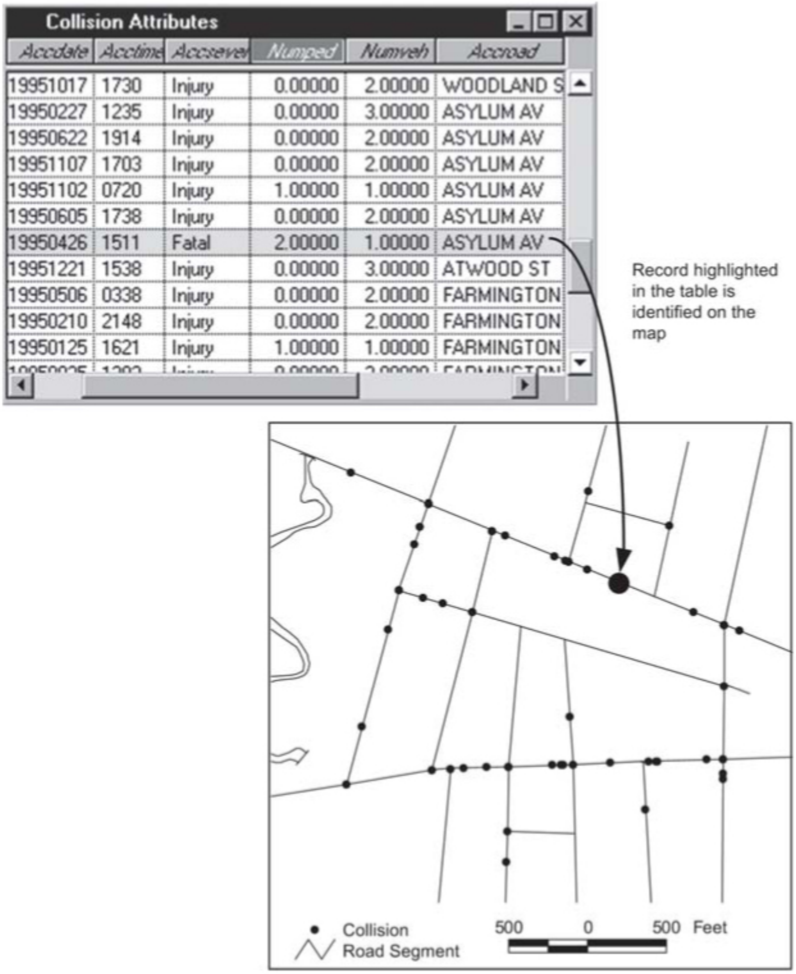
On the other hand, we might be concerned about health events at a particular place. The environmental health analyst might wish to study a variety of health outcomes in an area affected by a contamination event. Individual citizens and community groups are often interested in health problems experienced in their own neighborhoods or communities. In these cases, the GIS users have already established a place of interest and want to understand the health status of the population in that place. We call this "geographical viewing." For example, we might ask, "How many cases of Lyme disease and associated tick-borne disease occurred in Fairfield County, Connecticut, last year?" Interacting with spatial databases displayed in a GIS view enables public health analysts to answer the basic questions of what and where.

Viewing by Attribute

Viewing by attribute starts with the characteristics of events, as described in the table, and identifies those events on a map. In the simplest case, we identify the location of a single event. For example, assume that we have a database showing the locations of all motor vehicle collisions in Connecticut. One particular collision that resulted in a fatality is of interest. We can select that collision record in the table, and its corresponding location will be highlighted in the display (Figure 4.13). Some systems enable users to pan and zoom to the selected collision even if it is not currently in the view at the time the collision is selected.

A more complex operation is to select by attribute, identifying multiple records based on their common attributes. We select events that have particular characteristics and display their locations on the view. For example, policymakers may want to know the locations of all motor vehicle collisions that involve pedestrians. We query the table for all collisions involving pedestrians, select those collisions, and their locations will be identified on the view (Figure 4.14).

It is also possible to select events by attribute and location simultaneously. For example, we might want to identify all motor vehicle collisions that occurred in Hartford, Connecticut, and involved pedestrians. The geographical query is to select the town of Hartford, and the attribute query is to select only collisions involving pedestrians. We need to find the ***intersection*** of these two queries—events that satisfy both conditions. Depending on how the data set is structured, there are several ways to perform these queries. One is first to select Hartford in the view, then query the data set for collisions involving pedestrians. Alternatively, if the name of the city where the collision occurred is included in the data set, it would be possible to query the database to select events where the accident town is Hartford and the collision type is pedestrian.



Accdate	Acctime	Accsever	Numped	Numveh	Accroad
19951017	1730	Injury	0.00000	2.00000	WOODLAND S
19950227	1235	Injury	0.00000	3.00000	ASYLUM AV
19950622	1914	Injury	0.00000	2.00000	ASYLUM AV
19951107	1703	Injury	0.00000	2.00000	ASYLUM AV
19951102	0720	Injury	1.00000	1.00000	ASYLUM AV
19950605	1738	Injury	0.00000	2.00000	ASYLUM AV
19950426	1511	Fatal	2.00000	1.00000	ASYLUM AV
19951221	1538	Injury	0.00000	3.00000	ATWOOD ST
19950506	0338	Injury	0.00000	2.00000	FARMINGTON
19950210	2148	Injury	0.00000	2.00000	FARMINGTON
19950125	1621	Injury	1.00000	1.00000	FARMINGTON

FIGURE 4.13. Viewing by attribute allows the user to highlight a motor vehicle collision with the attribute "Fatal" from a table of data and find the location of the collision in the GIS display.

Most complex queries involve ***Boolean operations*** (Table 4.1). Boolean operations are operations that are applied to sets and logical propositions. In GIS queries, these operations define how the attributes of events are to be combined. Two common Boolean operations are AND (intersection) and OR (union). The AND operator finds the intersection of two attributes. It identifies events that satisfy two conditions simultaneously.

In the above example, we selected collisions based on an AND query. A sequence of AND operations finds events that satisfy more than two conditions simultaneously. To identify collisions that occurred in Hartford, involved pedes-

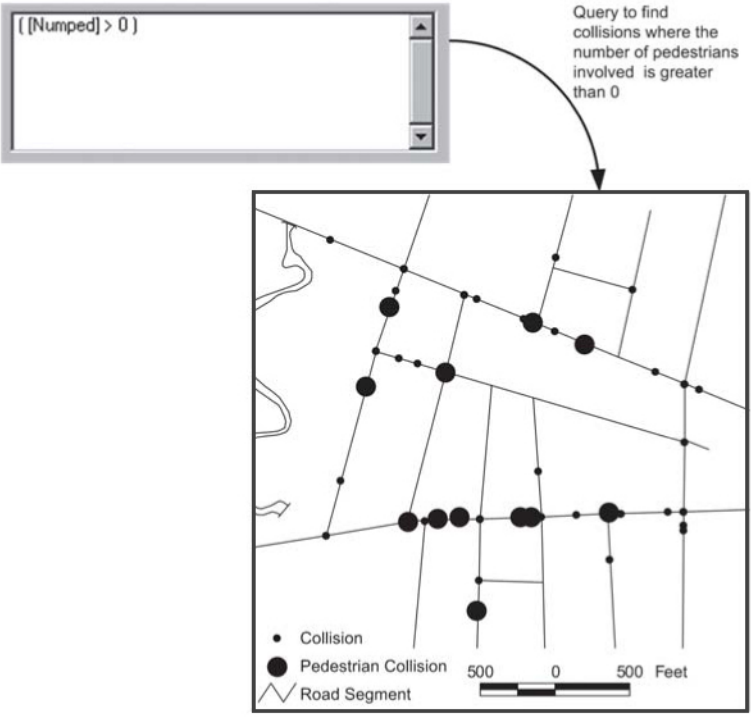


FIGURE 4.14. Viewing by attribute allows queries of a database table to identify all collisions that satisfy a set of criteria and highlight the locations of these collisions in the GIS display.

trians, and a school bus, we would create the following query to find every record where the accident town is “HARTFORD”, the collision type is “Pedestrian,” and the vehicle type is “School Bus.”

The OR operator finds the ***union*** of two attributes, that is, events that satisfy at least one of the two attributes. To select collisions that involved a pedestrian or a school bus, we would create the query to find every record where the collision type is “Pedestrian” OR the vehicle type is “School Bus.” The collisions identified in this query would include any collision involving a pedestrian, any collision involving a school bus, and any collision involving both. Queries can combine Boolean operators to identify detailed subsets of events for display, as shown in Table 4.1.

The ability to perform complex queries, both geographical and attribute-based, is an important feature of GIS. In a real-world setting, the challenge is to move from verbal queries like “Show me the fatal collisions that occurred on two-lane roads and involved people who were driving while intoxicated (DWI)” to the precise logical and geographical statements that are needed to perform these queries in GIS. This query looks for events that satisfy three conditions:

TABLE 4.1. Boolean Operators

Boolean operator	Notation	Definition	Example
Equality	S = T	A relationship between two sets when the sets contain precisely the same elements.	If S is the set of collisions obtained by using the GIS to find all collision points within the polygon of the Hartford town boundary and T is the set of collisions obtained by using the GIS to query the collision attribute table to find all collisions with AccTown=HARTFORD and there are no errors in the spatial or attribute information, then the two sets would have the same collisions as members and the sets would be equal.
Subset	$T \subseteq C$	A relationship between two sets where every element of S is an element of the second set T.	If C is the set of collisions occurring in the state of Connecticut and T is the set of collisions occurring in Hartford, then T would be a subset of C. C would not be a subset of T.
Intersection	$P \cap T$	A binary operation that takes two sets and returns the set of elements that are members of both the original sets.	If P is the set of all pedestrian collisions in the state of Connecticut and T is the set of all collisions in Hartford, then the intersection of P and T would be the set of all pedestrian collisions in Hartford.
Union	$B \cup H$	A binary operation that takes two sets and returns the set of elements that are members of at least one of the original sets.	If H is the set of all pedestrian collisions in Hartford and B is the set of all collisions in Hartford involving a school bus, then the union of B and H would be the set of all collisions in Hartford involving *either* a school bus *or* a pedestrian, or both.
Empty or Null	$\varnothing$	The set contains no elements.	No motorcycles were involved in collisions in Hartford that also involved a pedestrian. A query to find the intersection of pedestrian and motorcycle collisions in Hartford would return an empty set.

(cont.)

TABLE 4.1. *(cont.)*

Boolean operator	Notation	Definition	Example
Difference	B\H	A binary operation that takes two sets and returns the set of elements that are members of the first set but *not* the second set.	The difference of B and H would be the collisions in Hartford that involved school buses not including any school bus collisions in Hartford that involved pedestrians.
Complement	P′	A unary operation applied to a set that returns the set of elements *not* in the set. The complement is always taken with reference to a universal set.	The complement of P, the set of all pedestrian collisions in the state of Connecticut, would be the set of all other collisions in Connecticut. The complement of T, the set of all collisions occurring in the town of Hartford, would be the set of all other collisions in Connecticut.

fatal collisions, location on two-lane roads, and DWI. In addressing the query, we need to satisfy all three conditions and find the intersection among them. Creating queries calls for a skilled GIS analyst who can translate verbal statements into logical ones and perform the GIS operations needed to satisfy them.

Geographical Viewing

Geographical viewing starts with a geographic area(s) of interest and asks about the attributes of events located within those areas. In a GIS view, the analyst can select locations, or pan and zoom to particular locations, and then examine the attribute data in the table for those selected events. The map is the starting point, and the analyst links back to attribute information in the table.

Using standard query tools in GIS, one can select features according to their point or area locations. Figure 4.15 contains a view of point data showing the locations where people were bitten by rats in New York City in a certain year. We can click on one of those points to view the attribute data associated with the particular cases at that point—age of person, location of the bite on the person, activity, and so on. GIS users can also select by area. To determine which rat bites occurred in the borough of Staten Island, we select that borough by “turning on” the borough data layer in the display and clicking on the borough of Staten Island, then selecting cases occurring within the selected borough. This operation gives access to the attribute information for cases in that borough. Tables, charts, and statistics can be generated to describe the aggregate characteristics of those selected cases, such as median age, gender, and location (home, work, park) where the bites took place.

FIGURE 4.15. The geographical viewing capabilities of a GIS enable users to access the attributes of an event like a rat bite that occurred at a particular location.

Another kind of geographical viewing is a type of windowing in which we select events within a user-defined area. Most GIS include a tool that allows users quickly to draw a rectangle, circle, or irregularly shaped polygon in the view. Events located inside the polygon are selected so that one can explore and summarize their characteristics (Figure 4.16). By drawing a window around those cases and selecting them, the analyst can generate summary statistics that compare events inside and outside the window.

Geographical viewing offers a powerful means of exploring data based on geographical location. The map, combined with the viewer's knowledge and perceptions of places on the map, stimulates queries about what types of health problems exist in particular areas and why the problems cluster geographically.

Changing the View

Views are not static. Within the limits defined by the scale and extent of the data set, one can change the view, moving across the map or focusing on areas of interest. ***Pan*** refers to movement across a map, bringing new areas into the view. Often the view includes just part of the geographic extent of a data set. Using the pan function, we move across that geographic extent to bring another part of the map into the view. We can also change the view by zooming in or out. When we ***zoom in*** to an area, we move toward it, keeping it in focus in the view. Zooming in is useful for getting a closer look at areas of special interest. By zooming in to a cluster of health events, the detailed geography of the disease cluster becomes apparent. Drawing upon other data layers, we can observe the concentration of events along roads, or in relation to parks, landfills, and other features. To

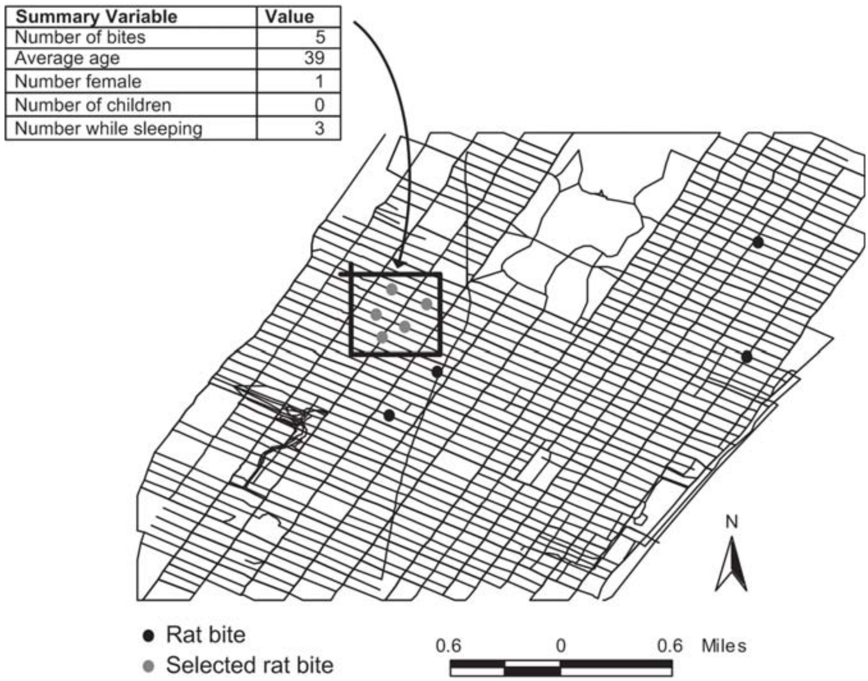


FIGURE 4.16. Selecting rat bites within a user-defined rectangular window. Attributes of bites inside the window are summarized in tabular form.

zoom out is to move away from the map, bringing a larger geographic area into the view. Zooming out is useful for examining regional differences, for getting a "wide-angle" view of broad spatial patterns.

An important point to remember about changing the view is that the scale and extent of data in the GIS limit the view. If our foundation data encompass only the state of Iowa, we cannot pan beyond the state to see neighboring areas of Minnesota or Nebraska. Similarly, if our original foundation data is at a scale of 1:24,000, by zooming in to the map we will not see more detail than exists at the 1:24,000 scale. This is one reason why it is so important to think carefully at the beginning of a GIS mapping project about the scale and extent of the foundation data to be used.

Viewing and Analyzing Geographical Associations

Another way to change the view is to add or remove features visible on the view. Known as ***cartographic overlay***, such procedures involve the overlay of data layers to show geographical associations. In most GIS, cartographic overlay requires a simple click of the mouse to make data layers visible or not visible in the display. For example, we might begin with a crime map that shows locations where assaults occurred. To identify places where school children might be at risk of assault, we activate the schools' data layer, overlaying the school locations on the view of assaults (Figure 4.17). The overlay reveals that school *A* is in an area with a greater number of assaults than school *B*, which may mean that students who attend school *A* are at greater risk of assault. By adding and removing data layers in this way, we can begin to visualize and explore geographical associations.

Spatial queries are procedures for creating new information about the geographical relationships among data layers that are visible in cartographic overlay. For example, we might want to know how many homes are located in a protected watershed region. To answer this question, we need to overlay the point data layer showing the locations of homes with the area (polygon) data layer depicting the watershed. The spatial query counts the number of homes that fall within the watershed. Spatial queries are similar to the queries described earlier, except that they (spatial queries) are based on the relative locations of features, rather than on their attributes. While cartographic overlay offers a picture of the relationships among layers, spatial queries provide a quantitative or qualitative summary of those relationships.

Spatial queries may differ on the type of spatial data—raster or vector—that one is working with. Raster layers that are registered together and have the

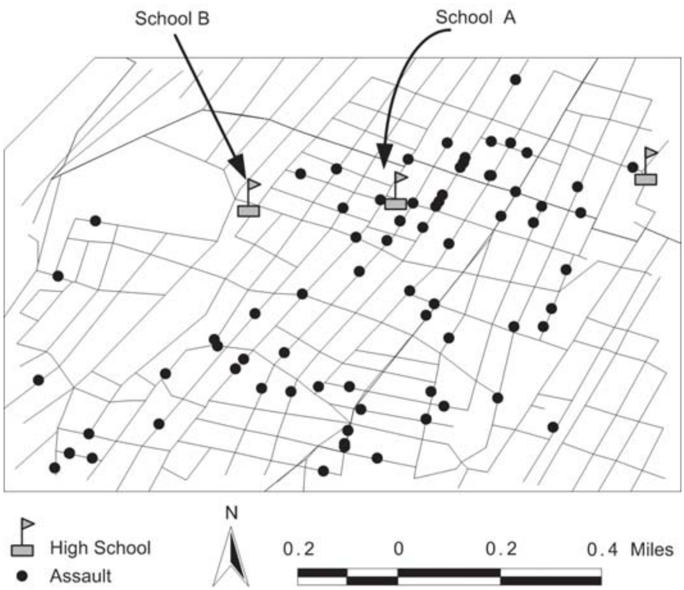


FIGURE 4.17. Cartographic overlay of a data layer showing high school locations with another data layer showing the locations of assaults. The overlay reveals that high school A is located in an area with a higher number of assaults than high school B.

same pixel size will overlay perfectly. Under these circumstances, spatial queries can be performed on a pixel-by-pixel basis using ***map algebra*** (Tomlin, 1990). Map algebra consists of Boolean and algebraic operations to combine and relate pixel values. All of the operations described in Table 4.1 can be performed to create new raster images based on combinations of data layers. If the pixel values are measured on an interval or a ratio scale, they can be combined via algebraic operations. For vector-borne diseases, like Lyme disease and hanta virus, such operations have been used in estimating vector density as a function of elevation, land cover, and vegetation density, as discussed in Chapter 8.

When using vector data, spatial queries are based on comparisons of relative location as defined by geographical coordinates. The types of spatial queries depend on the types of features—point, line, or polygon—being compared. Queries that compare two point data layers typically involve calculating distances between points. One can determine both average distances between points on different layers and numbers of points in one data layer within a given distance of points in the other data layer. These types of queries have many potential applications in public health GIS, but they are especially important in analyzing locations of health care services and geographical accessibility to those services, as discussed in Chapters 9 and 10. For example, health service planners often want to know: How far is each town from its nearest hospital? This spatial query involves comparing two point data layers, one showing the locations of towns and the other the locations of hospitals.

Spatial queries involving point and polygon vector data are also common in public health GIS. In this case, the queries involve identifying the areas where points are located, or the number of points within areas, or the number of points within a distance radius of an area. One of the most common types of queries is to determine the area in which a point is located, for example, the political district or block in which a residence is located (Figure 4.18). The point data layer is superimposed on the polygon data layer that shows area boundaries. Then, a ***point-in-polygon*** operation is performed to identify the particular area that contains the point. Although the human eye can easily visualize and respond to a point-in-polygon query, such queries are complex from a technical and computational standpoint and are based on polygon topology. After completing the query, the polygon name or number can become a new attribute in the point data layer. Alternatively, one can use point-in-polygon operations to identify and characterize the points that fall inside particular polygons. The watershed example, mentioned above, exemplifies this type of point-in-polygon query. Note that the number of homes located within the watershed becomes a new attribute of the watershed polygon.

Another class of spatial queries addresses the associations between linear features and point and area features. Line-point queries involve distance relationships; for example, we might ask, how many homes are located within 500 meters of a highway? To respond to this query, the GIS computes distances from each home to the highway and counts the number of homes that have a distance

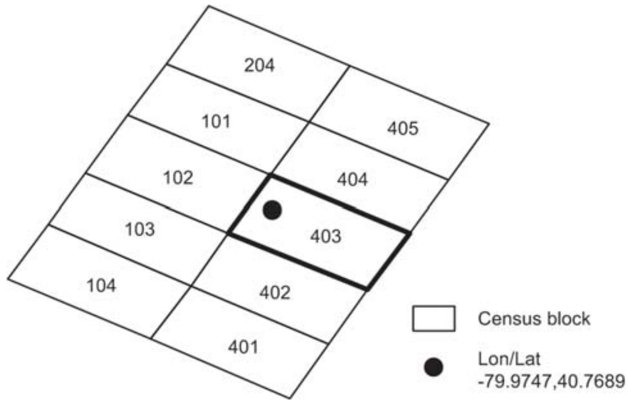


FIGURE 4.18. Point-in-polygon operation to find the block in which a point is located. The operation involves comparing the geographical coordinates of the point, in this case (lon,lat) coordinates, to the coordinates of the vertices of the polygon to determine if the point is "inside" the polygon.

less than 500 meters. Line-in-polygon queries ask which linear features fall within a particular polygon: that is, which highways pass through the Williamsburg neighborhood? A ***line-in-polygon*** operation, similar to the point-in-polygon operation noted above, is used in solving these types of spatial queries.

The final class of vector spatial queries involves comparing two or more polygon data layers. One can overlay the polygon layers and use Boolean or algebraic operations to create new information. For instance, we might want to find all census tracts that are located in the service area for Saint Francis Hospital and have a median household income under $25,000. We overlay the hospital service area data layer with the layer containing census tracts with incomes below $25,000 and use the Boolean AND operator to find tracts, or parts of tracts, that satisfy both criteria. Polygon overlay can be a time-consuming process, especially when it involves large numbers of polygons, because it works from the detailed topological relationships among the polygons (DeMers, 2000). To save time in polygon overlay, we can often preprocess the polygon data layers to remove polygons that are not of interest. In this example, we could eliminate census tracts with incomes above $25,000, using only the remaining tracts in performing the spatial query.

Spatial queries are an essential component of all GIS systems, enabling the user to create new information based on geographical relationships between various layers of spatial information. Chapters 5 through 11 consider the use of different types of spatial queries in diverse public health GIS applications, including environmental health, communicable and vector-borne disease, and health services access and location.

GIS and Map Publication

The mapping process as implemented in a GIS emphasizes the representation and analysis of information. Nevertheless, finished maps remain an important product of GIS to support the various stages of compilation, exploration, and analysis of data and to present results. A single map or a series of maps may be prepared for publication, either in traditional printed format or in digital form on the Internet. These maps are finished products that may be viewed and referenced by a diverse audience, and there are a number of design elements worth considering in preparing maps for publication.

Key Elements of Thematic Maps

Thematic maps typically contain certain key elements (Table 4.2). These elements define the nature and source of a map's contents. They also assist the map reader.

The ***title*** of the published map describes that major theme of the map. Carefully written titles are important for communicating that theme. The ***legend*** of a thematic map identifies and defines the symbols used on the map. ***Neatlines*** define the borders of the map and areas within it, including insets and legends. They can be used effectively to partition the document and draw the eye to different map elements.

Published maps should also include a north arrow and scale. The ***north arrow*** indicates map orientation. As discussed in Chapter 2, map ***scale*** is important because it affects the detail that can be portrayed.

Some cartographers have suggested that these are not essential elements for thematic map design (Slocum et al., 2009). Given a map of the United States, for

TABLE 4.2. Elements of Thematic Maps

Element	Description
Title	Describes major theme of map.
Legend	Defines map symbols.
Neatlines	Define borders of the map sheet and areas within it, including insets and legends.
North arrow	Describes map orientation.
Scale	Describes map distance in relation to earth distance.
Source	Describes source, date, and reliability of mapped data.
Agency	Identifies agency responsible for preparing and/or publishing the map.

example, readers will recognize the map domain and have a sense of the orientation and scale even if they are not explicit. Many public health GIS applications, however, deal with localities or regions that may be less familiar to map readers. A north arrow should be included, particularly if the orientation of the map on the page is different from north and if direction is important in interpreting the map content. This might be the case in mapping the plume from a point source of emissions into the atmosphere.

The scale of the map is information that should also be clearly presented. Maps have traditionally been compiled at a series of standard scales (1:24,000, 1:50,000, 1:200,000, and so on), and the scale element assists the reader in placing the published map along this series. Furthermore, published maps are documents that are acquired by libraries and research institutes, catalogued, and made accessible to a potentially large audience. Inclusion of a map scale assists in cataloguing and helps individuals find maps that might reveal the spatial patterns of interest because the maps were prepared at a scale that can display the necessary detail.

Published maps may also include the ***source*** of the data that are the theme of the map. The information might include the date of the data and pertinent information about the reliability of the data. Finally, the ***agency*** responsible for preparing and/or publishing the map should be included.

GIS software packages include a range of functions that help users take the information from a cartographic display of data and compose a map document for printing or storing in common graphics formats. Text, charts, tables, photographs, and other elements like the date the map is printed can often be incorporated into the map document. Layouts can be stored as templates to preserve the standard elements in the design.

Maps on the Internet

Maps, like other documents that traditionally have been printed and distributed on paper, are now primarily being published in digital form on the Internet (Kraak & Brown, 2000; Kraak, 2004). In the criminal justice field, police departments, such as the Chicago Police Department, are posting maps of crime occurrence on the Internet to give citizens access to the information for their communities (Chicago Police Department, 2011). Public health agencies and researchers are increasingly making maps of health events and health care services available on the web. A good example is the National Cancer Institute's Cancer Maps and Graphs website that allows users to create their own customizable maps of cancer mortality (National Cancer Institute, 2011c). Many other examples of Internet-based mapping of health and health care services are discussed in Chapters 6 through 12.

Internet mapping has changed greatly in the past decade. Early efforts focused primarily on creating ***static maps***, maps viewed as documents on the Internet (Peterson, 2005). These maps are created by scanning existing paper maps or by saving map layouts in one of a variety of common graphics formats like ***JPEG*** (Joint Photographic Experts Group) or ***GIF*** (Graphics Interchange Format). Static maps can basically be located and viewed. Viewers do not interact with these maps and cannot modify them.

The next generation of web mapping emphasized ***dynamic or interactive maps***—maps that allow the user to interact with the map in some way (Plewe, 2007). These maps may allow viewers to display the map in a different projection, pan or zoon, or separate map data layers (Peterson, 2005). Viewers may be able to click on an element in the view or pass the mouse over an element to obtain further information (Kraak & Brown, 2000). Many interactive maps are also ***on-demand maps*** that are generated from databases according to requests made by the viewer. In the cancer mapping site mentioned above, users can choose the type of cancer, time period, age group, and scale of data (state, county) for mapping. Dynamic maps typically rely on web-based GIS to "serve" maps over the Internet.

A good example of an interactive mapping website is the U.S. Centers for Disease Control and Prevention's site for interactive mapping of heart disease and stroke (Centers for Disease Control and Prevention, 2011c). For each disease, users can create customized maps of mortality or hospitalizations and show the locations of hospital facilities. Maps can be constructed for different time periods and gender and race groups, at county and state geographical scales. Sites like this one promote the sharing of health information with researchers, community organizations, and the public.

As discussed in Chapter 1, the latest generation of web mapping, focused on Google Earth®, has fundamentally changed how maps are viewed and created on the Internet (Plewe, 2007). Satellite image data and basic mapping tools are freely available to anyone with access to the Internet. Users construct their own maps by panning and zooming to areas of interest and controlling the types of data to be displayed. Users can contribute their own content to the map and share content with others in an open, distributed, and potentially collaborative, mapping process. Boulos and Burden (2007) describe this process as the "democratization of GIS"—a shift from proprietary mapping systems to user-driven, collaborative map production.

Government agencies are also creating similar kinds of mapping systems that enable users to construct their own maps and download customized data sets. The U.S. Geological Survey's Digital Map-Beta (U.S. Geological Survey, 2009a) lets users turn data layers on and off, pan and zoom, and create mashups for use with Google Maps®. In the United Kingdom, users can create customized maps of specific areas at a range of geographical scales using the Ordnance Survey's Get-a-Map system (Ordnance Survey, 2011c). Natural Resources Canada provides a wealth of spatial data and maps online (Natural Resources Canada, 2009). Although most of these systems lack the collaborative feature discussed earlier, they typically provide standardized and high-quality data and maps on a wide range of environmental and social features.

Map production on the Internet raises many new issues for cartographic design and map distribution. New types of symbology and visualization options are possible on the web. At the same time, map production and distribution may be limited by issues such as processing speed and the computer screen as a medium. In addition, online map publication involves much more than just design questions because the map image has a ***front end***—the user interface for the mapping application displayed within the browser—and a ***back end***—the digital databases accessed to respond to queries and requests and to produce the resulting map.

Web mapping greatly expands the range of cartographic variables and map symbols beyond the original set of visual variables proposed by Bertin (see Figure 2.17). Attributes like ***transparency***, ***fog***, and ***blur*** can be used to differentiate features in ways not possible on paper maps (Kraak, 2004). Blur is often used to indicate data uncertainty; fog obscures parts of the map, removing them from view. Internet mapping also permits use of dynamic symbols such as blinking symbols to show stability over time or draw attention to important features. Animated map sequences showing spatial and temporal change (discussed in Chapter 7) are easily created in a web environment. The web also provides an ideal platform for three-dimensional mapping of health data. Boulos and Burden (2007) describe how virtual reality software linked with Google Earth can be used to map the movement of diseases over space and time and to depict changes in local environments that promote or disrupt disease transmission.

Although web mapping has many advantages, map production and distribution may be limited by technological barriers. Despite increases in the size and resolution of computer screens, the screen still limits the portion of map that can be viewed and its level of detail. That portion will be determined by the size of the screen and the area inside the web browser. If a map is larger than the available space, viewers will have to scroll to see every part of the map. Moreover, computer screens are raster devices. Screen resolution is generally much lower than the resolution of many desktop printers, which are capable of 1,200 ***dots per inch*** or ***dpi*** as a minimum. Viewers' monitors may have resolutions ranging from 60 to 100 dpi. The lower resolution also limits geographic detail, text size, and shade patterns.

It is less expensive to publish maps in color online than in printed form. However, not all colors are browser independent. Viewers who do not have color printers will only be able to print black and white versions of the maps.

In addition to the design of the map, publication of maps online requires design of a user interface. The user interface provides the viewer with access to tools for map navigation, map display, and map querying and analysis. In intranet applications that provide access to maps within organizations, interface design can be very simple. For Internet applications where maps will be accessed by a potentially large and unknown group of viewers, interface design may be more challenging.

A general principle is that the interface design should feel natural and intuitive to users (Harrower & Sheesley, 2005). Users are becoming much more accustomed to navigating web maps, having interacted with online mapping systems in their everyday lives. For basic navigation tasks like pan and zoom, they may expect to find specific symbols in specific positions on the user interface. Harrower and Sheesley (2005) outline detailed criteria to guide the design of user-friendly map browsing tools: the interface should facilitate sequential and nonsequential browsing, allow precise control over the extent of browsing, include local and global orientation cues, and tightly link user commands and the map response.

The greater the number of functions available to the user for interacting with the map, the greater the number of controls there will be in the interface and the more complex it will become. Map images, graphical icons and buttons, text with and without hyperlinks, and forms supporting various types of input can all be used in interfaces supporting mapping applications (Plewe, 1997). The arrangement of elements in the interface is also important. Controls that will be used more frequently may be placed in different areas from those that will be used only rarely. Map navigation controls for panning and zooming are generally placed close to the map image.

As the number of functions increases, processing requirements increase. There is a tension between how much of the processing occurs on the server versus the client machine. Generally, more functions result in more processing on the client. These functions can usually only be supported if the client has or can obtain the necessary plug-in or software to support the function.

KML (Keyhole Markup Language) is the current standard for expressing annotations, features, and visualizations on web-based two-dimensional maps and three-dimensional earth browsers such as Google Earth. It is used to place vector objects including points, lines, and polygons on a map or browser and to annotate those features with descriptive information such as text or images. KML has a tag-based structure in which the tags are used to determine how and where features appear (Google, 2011b). Locational information is expressed in lat/lon coordinates based on the standard latitude, longitude reference system with the WGS84 datum. Altitude can also be included. The format for specifying a point location in KML is: longitude, latitude, altitude. Table 4.3 shows an example of a KML script to create a point placemark. Note that the ***placemark***

TABLE 4.3. Sample KML Script to Create a Placemark

```
<?xml version="1.0" encoding="UTF-8">
  <kml xmlns="http://www.www.opengis.net/kml/2.2">
    <Placemark>
      <name>UIUC, Geography</name>
      <description>Department of Geography, University of
        Illinois</description>
      <Point>
        <coordinates>-88.2264606,40.107335,0</coordinates>
      </Point>
    </Placemark>
  </kml>
```

has three elements—a name, a description, and a set of lat/lon coordinates—that identify where the point is located and the feature it represents. When KML files reference complex features like paths and polygons, and include detailed descriptions, the files can be very large. These large files are typically compressed into the ***KMZ*** (zipped) format. An advantage of the zipped format is that it allows custom placemarks or images to be included in feature descriptions. Also, one can incorporate network links to data and descriptions on the web in order to access data from remote locations and distribute it to a large number of users.

What are the implications of these developments for GIS applications in public health? The growing trend toward distributing information on the Internet means that public health organizations need to plan for data distribution. Will maps or data, or both, be provided? If publication on the Internet is important, the institutional requirements for a successful GIS application increase. In addition to staff who can maintain the databases used in-house and design GIS applications, staff are needed who have expertise in web and interface design. Staff are also needed who can administer and maintain the servers and software that support the Internet mapping and data distribution applications. Even if the GIS is operated in a self-contained environment, public health analysts will need to become familiar with online mapping because much of the foundation and other data they will need to access to develop their GIS applications will be distributed over the Internet.

Conclusion

GIS has revolutionized the process of mapmaking. From obtaining data to developing maps, to creating on-demand images on the Internet, the process can be accomplished in a fully automated, digital environment. Technological developments are stimulating a new mapping process that emphasizes exploration and visualization of health information in new and innovative ways, rather than construction of finished maps. But the ease of operating GIS makes it even more imperative that users have a firm understanding of the basic principles of geography and cartography. Maps can lie. They can mislead just as easily as they can lead (Monmonier, 1996). Successful mapping depends on the knowledge and skills of analysts who use the systems and the integration of those systems in decision-making processes at the state, local, and community levels. As the links between the GIS and public health communities expand, mapping, viewing, and analyzing geographically based health information will occupy an even more central position in efforts to improve the performance of essential public health activities and promote community health and well-being.

CHAPTER 5

Analyzing Spatial Clustering of Health Events

Public health professionals are often faced with the task of investigating disease ***clusters***, unusual concentrations of health events in space and time. Clusters can come to the attention of public health departments when concerned citizens perceive an excess of ill health in their communities or through surveillance systems that detect an unusual concentration of health events by searching for patterns in routinely collected data (Neutra, Swan, & Mack, 1992). Whether the analysis is ***confirmatory***, verifying that a perceived cluster exists, or ***exploratory***, searching for patterns, GIS can play a crucial role in analyzing spatial clusters. As GIS technology develops, innovative spatial statistical methods are being linked with GIS to analyze the spatial clustering of disease in populations and to assess changes in health status and disease prevalence over time.

This chapter discusses methods for analyzing spatial clustering of health events and the use of GIS to implement these methods. We cover a representative set of methods and explain the procedures and concepts that underpin the methods. The emphasis is on GIS operations and applications rather than on statistical issues. We divide the methods into two broad categories: field-based and object-based approaches.

Spatial clustering methods can help provide answers to an array of fundamental public health questions. Do any unusual clusters of health events exist in an area? What places have unusually high or low prevalences of disease? Where are the risks of ill health highest or lowest? Spatial analysis methods offer a means of filtering health information in order to describe geographical patterns and identify unusual occurrences of health events.

Figure 5.1a shows the residential locations of children who have leukemia in a hypothetical city. A geographical cluster of cases appears in the northwest section of the city. Does this area have an unusually high rate of childhood leukemia? Answering this question requires several important bits of information. First, we need to examine the number of cases of leukemia in relation to the population at risk. The ***population at risk*** is the set of people who, because of their age or gender, can contract the health problem of interest. As discussed in the Introduction, the larger the risk population, the larger the number of cases one would expect to find. By definition, childhood leukemia occurs in children, so the risk population comprises all children living in the city or some part of the city. Because population is distributed unevenly over space, the density of health events will vary even when the underlying rate of ill health is uniform. Figures 5.1b and 5.1c depict two alternative spatial distributions of risk population for the leukemia example. In Figure 5.1c, the leukemia cluster occurs in an area of high-risk population density; thus population accounts for the disease "cluster." In Figure 5.1b this is not the case.

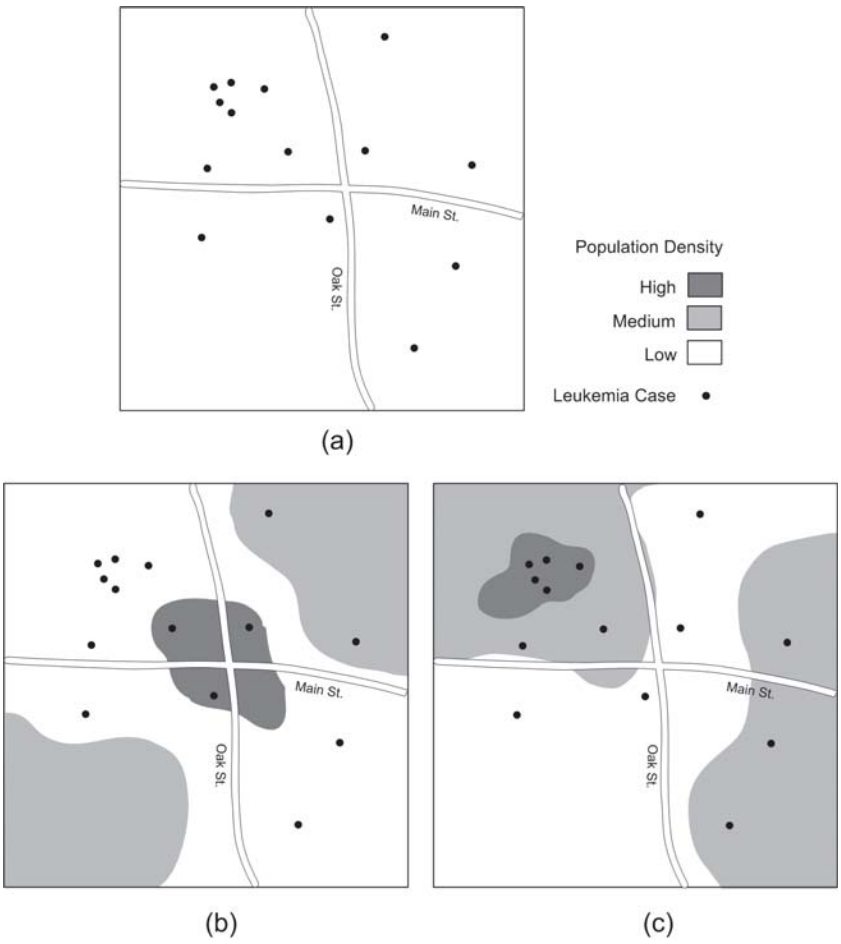


FIGURE 5.1. A hypothetical distribution of leukemia cases in relation to two different spatial patterns of risk population.

A second issue in analyzing clustering is to define the geographic extent, or ***scale***, at which clustering occurs. A cluster of cases within a 5-square-mile area has a very different meaning than a cluster of cases within a 2,500-square-mile area. The first indicates a highly localized cluster of disease, whereas the second identifies a large region with an elevated disease rate. All clustering methods focus on one, or, in some cases, a few, spatial scales. Sometimes the analyst can control the choice of spatial scale; however, often the analyst will find scale dictated by the scale of the underlying population or health data. The leukemia example addresses clustering at the intraurban scale; that is, it searches for clustering within small neighborhood areas rather than across cities or regions.

Scale critically affects the kinds of inferences that can be drawn from cluster studies. Clustering within cities or communities reflects localized factors such as point sources of environmental contamination. In contrast, elevated disease rates for states or regions result from regionwide factors like climate, culture, politics, and legislation, along with more localized factors. The scale at which a health problem is studied should reflect an understanding of the disease process and likely causative factors. Furthermore, the pattern at one geographical scale can be associated with patterns at other scales. A state's "average" rate of disease may result from having some communities with unusually high rates and others with unusually low rates, or from having many communities with rates close to the state average. These disparate rates are concealed in the statewide average. Analyses below the state scale are needed to reveal such high-rate communities.

Third, analyzing clustering requires a set of ***criteria*** for judging how much clustering exists. Does an excess of one or two cases constitute a cluster? Where do we draw the line in defining "significant" clusters? There is no perfect answer to these questions, but geographical and statistical methods can help analysts and policymakers make scientifically informed decisions. Many clustering procedures rely on statistical criteria that describe the likelihood that clusters could have arisen by chance in a given population. Such criteria may utilize a known probability distribution such as the Poisson distribution, or they may utilize Monte Carlo simulation methods, which involve generating a large number of random possible outcomes, discussed later in this chapter. Some procedures also emphasize the arrangement of health events, not just in relation to population at risk, but also in relation to potential sources of contamination or environmental hazard—for instance, do the events cluster near a toxic waste facility, or are they arranged along roadways or power lines? These methods are discussed in Chapters 6 and 8.

This chapter is divided into three sections. The first looks at mapping health outcomes within fixed geographic areas, and it raises an important issue that underpins all spatial cluster investigations: the small numbers problem. The second section highlights methods for cluster detection that are explicitly spatial. These methods involve aggregating health information over geographic space using field- or object-based approaches. The final section considers the rapidly expanding set of methods for examining clustering in space and time.

Mapping Disease Rates: The Small Numbers Problem

Often cluster investigations begin with data on published rates of disease for geographical areas such as counties or ZIP Codes. Such data can be displayed on a choropleth map, as discussed in Chapter 4. When the areas differ in risk population size, however, as is typically the case, the calculated rates of disease for those areas have different degrees of variability. Rates for areas with small populations vary more and are less reliable than those for more highly populated areas. For small-population areas, a difference of one or two cases can make a huge difference in incidence or prevalence rates. This is known as the ***small numbers problem***.

Figure 5.2 illustrates the small numbers problem with data on low birthweight (percent of babies born weighing less than 2,500 grams) for two areas that differ in population size. Area 1 averages under 200 births per year, while Area 2 averages over 1,600 births per year. Note the large variability in low birthweight in Area 1 from year to year. The low-birthweight rates fluctuate from 5 to 16% and appear unpredictable. In Area 2 the rates are more stable, ranging from 12 to 16%. A choropleth map of low-birthweight rates by area for a single year does not represent their varying degrees of reliability. For small areas like Area 1, the map can give a false picture of the level of health depending on which year's data happen to be selected for mapping. Area 2's mapped value is likely to be closer to its "true" underlying value.

Probability Mapping

Probability mapping is a well-established statistical method for addressing the small numbers problem (Choynowski, 1959). In probability mapping we map

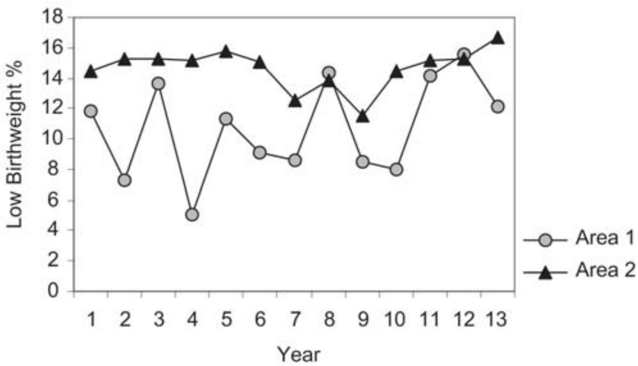


FIGURE 5.2. The small numbers problem illustrated with data on low birthweight over time for two health areas in New York City. Area 1 is a "small" health area, averaging 200 births per year, and its low-birthweight rates fluctuate greatly from year to year. Rates are much more stable in Area 2, a "large" health area with 1,600 births per year on average.

the statistical significance of rates rather than the rates themselves. Statistical significance is measured by probability values that show the likelihood of a rate occurring given the normal rate of disease in the corresponding national or regional population. We refer to this rate as the ***population rate***, *p*. The probability value for an area indicates the likelihood that the rate observed in that area would occur by chance if the underlying risk of disease was equal to *p*. Probability values close to 0 or 1 indicate rates that are significantly different from the population rate.

There are many statistical methods for computing probability values. One of the most common relies on the ***Poisson distribution***, used for modeling the probability of rare binary (present/absent) events in large populations. Many health problems (e.g., cancers, birth defects) fit this definition because they are rare, occurring in only a small fraction of the population, and binary, either present or absent in an individual. Consider a small area containing a population, *n*, and *k* cases of disease. We want to find out whether the presence of those *k* cases in a population of size *n* is unusual. In other words, is the actual number of cases significantly higher than expected based on the national or regional prevalence rate?

If the national or regional rate is *p*, the expected number of cases in the study area, lambda, is the study area population, *n*, multiplied by the national or regional rate:

$$\lambda = np$$

For example, if the study area contains 40,000 people and the national prevalence rate is 1 per 10,000, we would expect four cases in the study area because $\lambda = 0.0001(40{,}000) = 4$.

If we know the number of cases, *k*, occurring in a study region population, we can use the Poisson distribution to determine the probability, *P(k)*, that the observed number of cases would occur in a population of the study region's size. The Poisson distribution states that in a population of size *n*, the probability of *x* cases occurring is $P(x) = (e^{-\lambda}\lambda^{x})/x!$. In this example, the probability of one case would be 0.073 (Table 5.1). From this calculation, we can determine the probability of *k* or more cases occurring by chance, $P(x \geq k)$, if the true rate of disease in the population were *p*. That probability is calculated as

$$P(x \geq k) = 1 - \sum_{x=0}^{k-1} P(x)$$

For example, if there are six cases of disease in the study region where only four cases were expected based on national or regional rates, the corresponding probability value would be 1 – 0.785, or 0.215. Note that the value 0.785 comes from summing the Poisson probability values for $x = 0$ through $x = 5$. This

TABLE 5.1. Poisson Probabilities, λ = 4.0

Number of cases (x)	Probability $P(x)$	Cumulative probability $P(\leq x)$
0	.0183	.0183
1	.0733	.0916
2	.1465	.2381
3	.1954	.4335
4	.1954	.6289
5	.1563	.7852
6	.1042	.8894
7	.0595	.9489
8	.0298	.9787
9	.0134	.9919
10	.0053	.9972

means that there is a 21% chance of 6 or more cases occurring by chance if the underlying prevalence is 1 per 10,000. The closer this probability value is to zero, the smaller the likelihood that it would arise by chance alone. In this case, since the probability is not particularly small, we infer that the rate of disease is not unusually high. In general, probabilities less than 0.05 or 0.01 are considered to indicate significantly high prevalence rates.

Comparing choropleth and probability maps of low-birthweight rates for Manhattan in New York City that were produced with a GIS illustrates the differences in these approaches (Figure 5.3). The map of actual rates shows considerable variation among neighborhoods, with an area of high rates in northern Manhattan. In the probability map, some of the areas with exceptionally high or low rates disappear. These are low-population areas whose rates are unstable owing to the small numbers problem.

Probability mapping is a useful way of addressing the small numbers problem when mapping area health data, but it has two limitations. First, it does not preserve the content of the original data. Instead of mapping health incidence rates, it shows probability levels whose only connection to the rates themselves is through a statistical computation. Second, probability mapping tends to overemphasize the significance of rates in areas with large populations, because statistical significance is directly related to sample size. For an area with a large population, a rate that is slightly higher than the expected rate will often be statistically significant because the size of the population increases statistical power. This means it is easier to reject the null hypothesis that there is no difference in rates. Thus, a statistically significant difference may not be substantively meaningful. Analysts need to look beyond statistical significance by examining the raw diseases rates, the locations of high-rate areas, and any additional information that might assist in interpreting high-rate areas. GIS support this more comprehensive view of statistical significance.

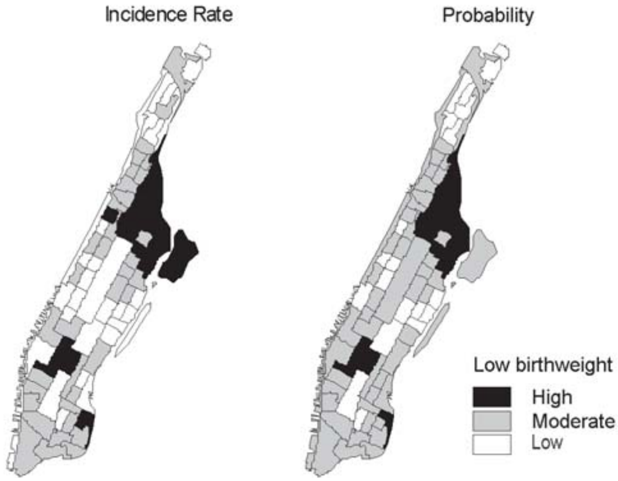


FIGURE 5.3. A map of incidence rates and a probability map of the same low-birthweight data for Manhattan show different patterns. On the probability map, areas in the high and low categories are places that have rates significantly higher or lower than the overall rate. Compared to the incidence rate map, the probability map shows fewer areas in the high and low categories. Many areas with small populations drop out of those categories on the probability map.

EMPIRICAL BAYES ESTIMATION

Empirical Bayes estimation is a method for addressing some of these issues while dealing with the small numbers problem (Clayton & Kaldor, 1987; Langford, 1994; Waller & Gotway, 2004). It represents a compromise between probability mapping and simple choropleth mapping of rates. In empirical Bayes estimation, rates are adjusted upward or downward, or ***smoothed***, according to the size of the population on which they are based. The smoothing process pulls rates toward the national or regional rate, making the rates more stable and less variable. The rates for small areas are smoothed more than those for large areas, reflecting differences in reliability linked to population size.

Three assumptions underlie empirical Bayes methods (Langford, 1994). The first is that the smoothing process should not affect the overall rate for the study area. We assume that this overall rate is reliable and unbiased. Second, as noted earlier, rates for small areas are adjusted more than rates for large areas. Finally, we assume that the incidence rates for all areas in the study region follow a known probability distribution, called the ***prior distribution***. Some common distributions are the gamma, beta, and log normal distributions.

The mathematical details of empirical Bayes estimation are well beyond the scope of this book, but by using these three assumptions we can describe the conceptual basis for the procedure. Many prior distributions used in empirical Bayes smoothing have two parameters: α, which describes the shape of the distribution, and β, which indicates the scale of the distribution. These are estimated statistically to "best fit" the distribution of actual health rates. The incidence rate across all areas is β/α.

Now consider an area i with population P_i and k_i cases of disease. The actual incidence rate for area i is k_i/P_i. Having estimated alpha and beta from the prior distribution, the smoothed incidence rate for area i is calculated as

$$(k_i + \beta)/(P_i + \alpha)$$

When area i is small in population size, k_i and P_i are small and the smoothed rate approaches the overall incidence rate. Conversely, when k_i and P_i are large, they dominate the smoothing process, and the smoothed rate for area i is very close to the actual rate, k_i/P_i.

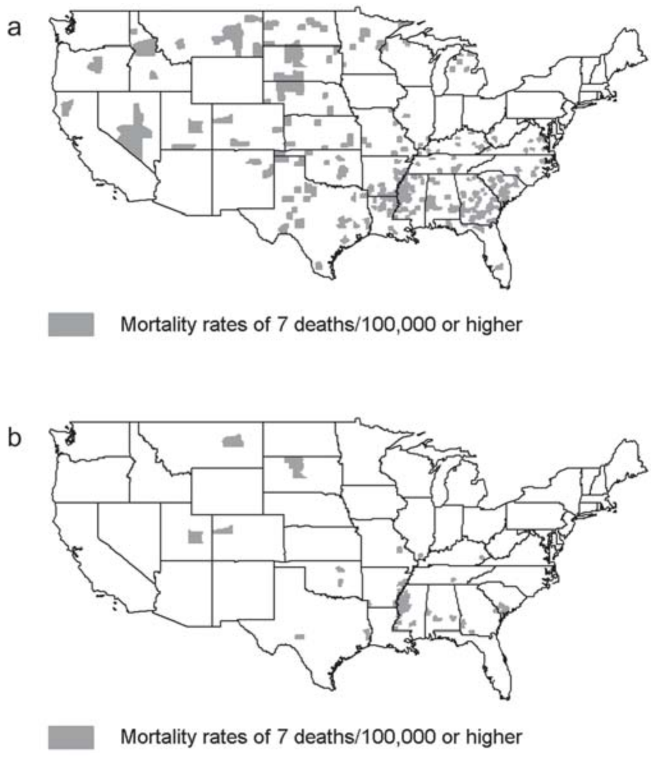


FIGURE 5.4. The choropleth map in Figure 5.4a shows U.S. counties with observed fire- and burn-related mortality rates of seven deaths per 100,000 population or higher, 1979 to 1987. The map in Figure 5.4b shows the empirical Bayes fire- and burn-related mortality rates of seven deaths per 100,000 population or higher, 1979 to 1987. The choropleth map and the smoothed map show different patterns. From Devine and Lewis (1994). Copyright 1994 by John Wiley & Sons, Ltd. Reprinted by permission.

As a consequence, empirical Bayes smoothing greatly affects rates for areas with small populations. Figure 5.4 shows a map of actual death rates from fire by county for the United States and a map of smoothed rates for the same data (Devine & Lewis, 1994). Counties with rates above 7.0 are shaded on the maps. The map of actual rates shows spatial clustering of high rates in the sparsely populated counties of the western states. More than 200 counties have rates above 7.0. In contrast, the smoothed map shows only 50 counties with high rates. Many of the less populated counties in the West moved out of the high-rate category. Because of their small populations, their rates were adjusted down toward the overall mean. Despite their high actual rates, these counties have such small populations that we cannot confidently categorize them as "high-rate" places.

An important issue in empirical Bayes estimation is to define the overall rate to which other rates are smoothed. For many kinds of health problems, using the same rate for all areas "washes out" the geographical variation and dependence that we know exists. Neighboring areas often have similar rates because of similarities in their social, economic, and environmental characteristics. A localized rate for the region in which an area is located serves as a better benchmark for smoothing.

Marshall (1991) describes procedures for performing ***localized empirical Bayes smoothing***. These procedures recognize and model the spatial dependence that characterizes virtually all health information. Generally the procedures work by computing for each area the localized rate of disease in its neighborhood. Then the disease rate for area *i* is smoothed toward this neighborhood rate rather than the national rate. Both traditional empirical Bayes rates and spatially smoothed rates can be computed using free downloadable software such as GeoDa (Anselin, 2003a).

Spatial Clustering Methods

The past two decades have seen major advances in the development of clustering methods that are explicitly spatial. These methods evaluate the occurrence of health events within neighboring geographic areas or among neighboring point locations (e.g., individual people, residences, or towns). Many draw upon GIS to provide innovative ways of viewing and analyzing health data that enable public health analysts to identify places with elevated disease rates, to gain insights into the likelihood that such rates would occur by chance, and to prioritize areas for further investigation.

This section considers a representative, but not exhaustive, set of methods. For more information on the full suite of methods, the reader can turn to several excellent reviews of the literature (Marshall, 1991; Waller & Gotway, 2004; Lawson, 2006). Cluster detection methods can be divided into three groups: those that assess overall clustering in a study area (***global methods***); those that seek to identify cluster locations (***local methods***); and those that assess clustering around a point source like a hazardous facility (***focused methods***). The latter two are of more direct interest to GIS because they can take full advantage of the display and analytical capabilities of GIS. In this chapter we discuss local methods used in identifying cluster locations. Focused tests are considered in Chapter 6, and the interested reader can also find useful information in Bithell (1995) and Puett et al. (2005).

From a GIS standpoint, the methods described in this section loosely follow the field and object data models, discussed in Chapter 2, in analyzing cluster locations. ***Field-based methods*** scan the study region, searching a geographical "field" of health data for evidence of clustering. In contrast, ***object-based methods*** search for clusters by grouping nearby cases of disease (objects) and then testing whether the grouped cases are likely to represent a statistically significant spatial cluster. Object-based methods start with case locations or, in the case of area data, with areas containing concentrations of cases, and attempt to build clusters by agglomerating nearby cases or areas.

Both field- and object-based spatial clustering methods require a procedure for defining how close places are in geographical space. Such procedures generate spatial weights—numerical values that describe closeness—which are then used in spatial clustering methods. We begin by discussing procedures for defining these spatial weights and then consider field- and object-based spatial clustering methods. The final section discusses methods for analyzing clustering in space and time.

Defining Spatial Weights

Methods for analyzing spatial clustering rely on spatial weights that describe the proximity of neighboring areas or health events. A ***spatial weight*** is a numerical value with a higher value indicating greater geographical proximity. Defining spatial weights involves operations that can easily be accomplished in GIS.

When using area health data, spatial weights are often defined based on ***contiguity*** (also known as ***adjacency***), that is, whether or not areas share a common boundary (Figure 5.5a). If areas i and j border each other, they are considered to be contiguous/adjacent and are assigned a spatial weight equal to one ($w_{ij} = 1$). Areas that are not contiguous are assigned a spatial weight of zero. GIS systems include operations for identifying contiguity among areas.

Contiguity can be directly observed for areas like census tracts or counties that are polygon features. For point features, one must associate each point with a corresponding polygon in order to define contiguity. Analysts can accomplish this using Thiessen polygons. A ***Thiessen polygon*** is a polygon whose boundaries demarcate the area that is closer to a particular point feature than to any other point (Figure 5.5b). Thiessen neighbors are points whose Thiessen polygons are contiguous (Brassel & Rief, 1979).

Another criterion for identifying spatial weights is ***proximity***, the distance between health events or areas. When using point data for health events, the

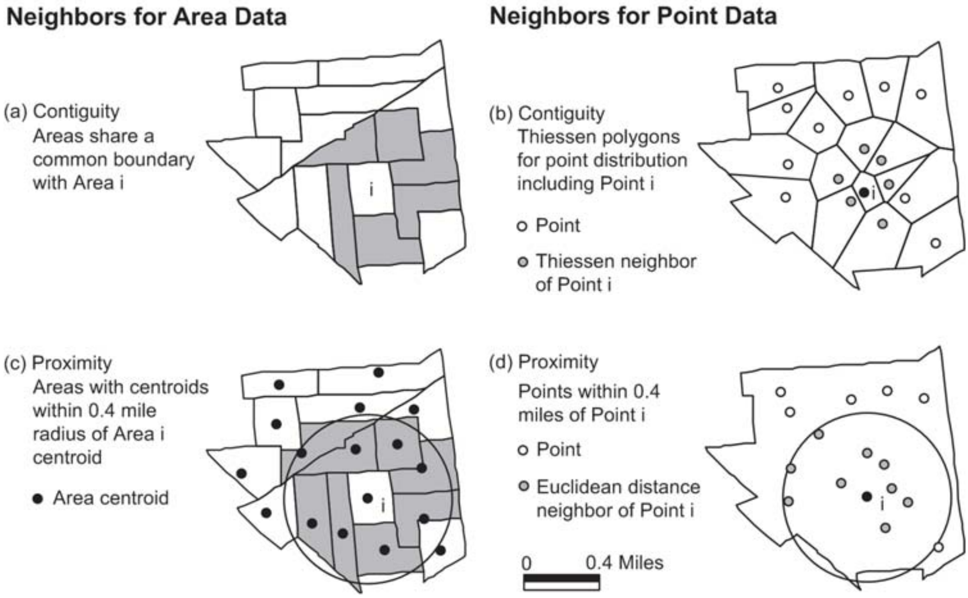


FIGURE 5.5. Defining the spatial neighbors for area and point data based on contiguity and proximity. Note how the neighbors differ depending on the criteria used.

distance between points serves as a convenient measure of proximity. Typically, Euclidean distance is used, but network and other distance measures may be more appropriate when interactions between people or places are constrained by a transportation network. After determining interpoint distances, spatial weights can be defined based on a critical cutoff distance, such that the weight equals one if two points are located within the critical distance, or zero otherwise (Figure 5.5d). Alternatively, we can compute spatial weights as an inverse function of distance—say, $w_{ij} = d_{ij}^{-b}$—where b represents the rate of decline in the weight with increasing distance.

Defining ***proximity*** is more complex for area data than for point data, and several methods have been proposed (Figure 5.5c). All involve determining the ***centroid***, or central location, for each area, and then calculating the distance between area centroids. Centroids can also be population-weighted so that they more accurately reflect the uneven spatial distribution of population within each area (Talbot, Kulldorff, Forand, & Haley, 2000). Most GIS include built-in functions for defining area centroids. Using centroids, we can define spatial weights as an inverse function of distance or based on a critical distance. Alternatively, we can define spatial weights based on the fraction of area lying within the critical distance of *i*'s centroid. Areas *j* and *i* are "neighbors" (i.e., $w_{ij} = 1$) if a large proportion of *j's* territory falls within the critical distance radius. GIS can be used to automate these spatial operations for computing spatial weights.

Field-Based Spatial Clustering Methods

Field-based spatial clustering methods search the entire area of interest for clusters of health events. The methods involve scanning the study region, moving from area to area or point to point. At each location, clustering is assessed within a local neighborhood according to statistical or mathematical criteria for cluster detection. One of the earliest and most influential field methods is the Geographical Analysis Machine developed by Stan Openshaw and colleagues in the 1980s (Openshaw, Charlton, & Craft, 1988). Although GAM is not widely used today, its technique of scanning a space within overlapping ***spatial windows*** (also known as ***spatial filters***) has been widely adopted in more recent procedures for cluster detection. This section considers a representative set of field-based cluster detection methods.

LOCAL MEASURES OF SPATIAL AUTOCORRELATION

An important class of methods for cluster detection includes localized measures of spatial dependence such as Getis and Ord's (1992) ***G_i* statistic*** and Anselin's (1995) ***local indicators of spatial autocorrelation (LISA)*** statistic. These measure the association between a value at a particular place and values for nearby or adjacent areas. The statistics are useful for finding disease clusters based on area data, although they may be applied to point data under some circumstances. A cluster is a region that has unusually high counts or rates of disease, that is, a local concentration of high values.

Getis and Ord's G_i* statistic illustrates well the structure of these localized measures of spatial dependence. Consider an area divided into m subareas. x_i refers to the value of the health indicator (e.g., incidence or prevalence rate or standardized mortality ratio) for area i. w_{ij} is a spatial weight defining the nearness of area i to area j, based either on contiguity or proximity, as discussed above. Given these definitions, the standardized G_i* statistic is

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}}{s\{[(nS_{1i}^*) - (W_i^*)^2]/(n-1)\}}$$

where

$$W_i^* = \sum_j w_{ij}(d) \quad \text{and} \quad S_{1i}^* = \sum_j w_{ij}{}^2$$

$$\bar{x} = \sum_j x_j \Big/ n$$

$$s = \{\sum (x_j - \bar{x})^2 / n\}^{0.5}$$

G_i* is positive when high rates of disease cluster in *i`s* local neighborhood. This indicates a disease cluster or geographical grouping of high prevalence rates, a ***hot spot*** (Figure 5.6). A negative value of G_i* shows a spatial cluster of low rates, known as a ***cold spot***.

Although G_i* is a statistical measure, calculating it involves several common GIS operations. GIS can be used for determining the spatial weights that indicate proximity/adjacency and for creating choropleth maps of the G_i* values that are useful in identifying cluster locations.

Another widely used measure of local spatial autocorrelation is the LISA statistic (Anselin, 1995). The LISA statistic is a local version of the well-known global indicator of spatial autocorrelation, Moran's *I*. Consider a study region that is divided into subareas. The LISA statistic for subarea *i* is

$$I_i = z_i \sum_{i \neq j} w_{ij} z_j$$

where z_i represents the standardized value of the health indicator of interest in area *i* and w_{ij} is a spatial weight measuring the nearness of subareas *i* and *j*. LISA essentially measures the statistical correlation between the value in subarea *i* and the values in nearby subareas. LISA values close to zero indicate little or no statistical association among neighboring values. A positive LISA statistic identifies a spatial concentration of similar values. These may be high values that represent high rates of disease—a high-high cluster or hot spot—or low values representing low rates

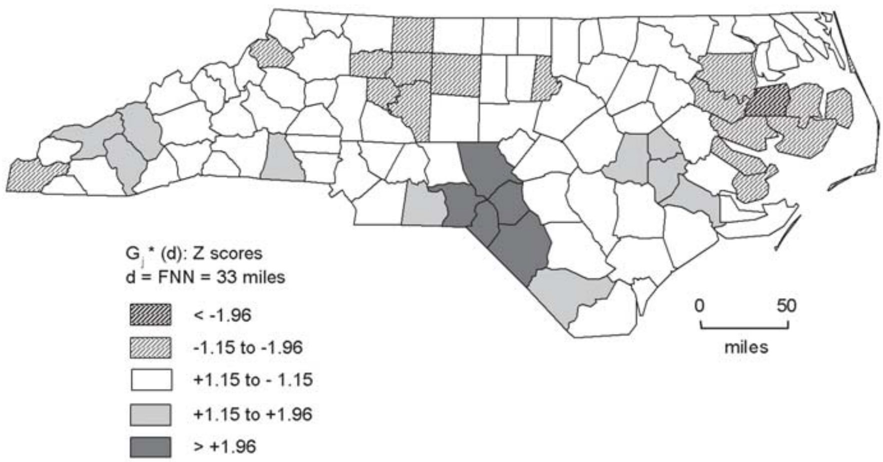


FIGURE 5.6. A map of standardized values of the G_i* statistic, a local measure of spatial autocorrelation, showing spatial patterns in SIDS rates in counties of North Carolina for 1979 through 1984. The furthest nearest neighbor distance, *d*, is 33 miles. Reprinted from Getis and Ord (1992). Copyright 1992 by John Wiley & Sons, Ltd. Reprinted by permission.

of disease—a low-low cluster or cold spot. When the LISA statistic is negative, we have a spatial cluster of *dissimilar* values, such as an area with a high disease rate surrounded by areas with low disease rates. GeoDa software can be used to calculate and map LISA statistics (Anselin, 2003a), as illustrated in Figure 5.7.

How does one determine if an observed LISA statistic is statistically significant? Significance testing is accomplished by ***Monte Carlo*** methods. Monte Carlo significance testing involves simulating the distribution of a test statistic such as LISA under a null hypothesis. Known as the ***reference distribution***, this null hypothesis distribution is generated via repeated random sampling or random reassignment of data values. To evaluate statistical significance, we compare the observed value of the test statistic to the simulated values in the reference distribution. If n is the number of simulations and m is the number of simulated values that exceed the observed test statistic value, $m/(n + 1)$ provides an estimated Monte Carlo p-value for use in significance testing (Waller & Gotway, 2004).

In Monte Carlo significance testing for LISA, the typical null hypothesis is that the health outcome of interest has *no* local spatial association. We then ask: "What LISA statistics would we expect to find if data values had no spatial association?" To determine the reference distribution of LISA statistics under this null hypothesis, the observed data values are randomly redistributed among subareas. This random reassignment of data values is performed a large number of times, for example, 999 times. For each randomization, the LISA statistic is computed for each subarea. These 999 statistics form the reference distribution, which describes the LISA values that would be obtained under the null hypothesis of no spatial association. The observed LISA statistic for subarea i is compared with its corresponding reference distribution to determine the statistical significance of local spatial clustering.

Two issues arise in the Monte Carlo randomization process when applied to local measures of spatial association. First, because the spatial neighborhoods around each subarea overlap, the significance tests associated with the LISA values for each subarea are not independent. This is the problem of ***multiple testing*** (Kulldorff, 1998), and it is important to include a correction for multiple testing in performing significance tests on LISA values (de Castro & Singer, 2006). Second, when the health indicator of interest is a disease rate (as opposed to a count of health events), the small numbers problem and the associated problem of variance instability also emerge as important in significance testing. As McLaughlin and Boscoe (2007) demonstrate, the end result is overdetection of clusters in less populated rural areas, and underdetection in densely populated areas. McLaughlin and Boscoe propose an alternative randomization procedure to overcome this problem.

KERNEL ESTIMATION

Kernel estimation is not a cluster detection method per se, but a method for exploring and displaying spatial patterns of point health data to show areas of high concentration. It is a method for generating a map that shows the density of

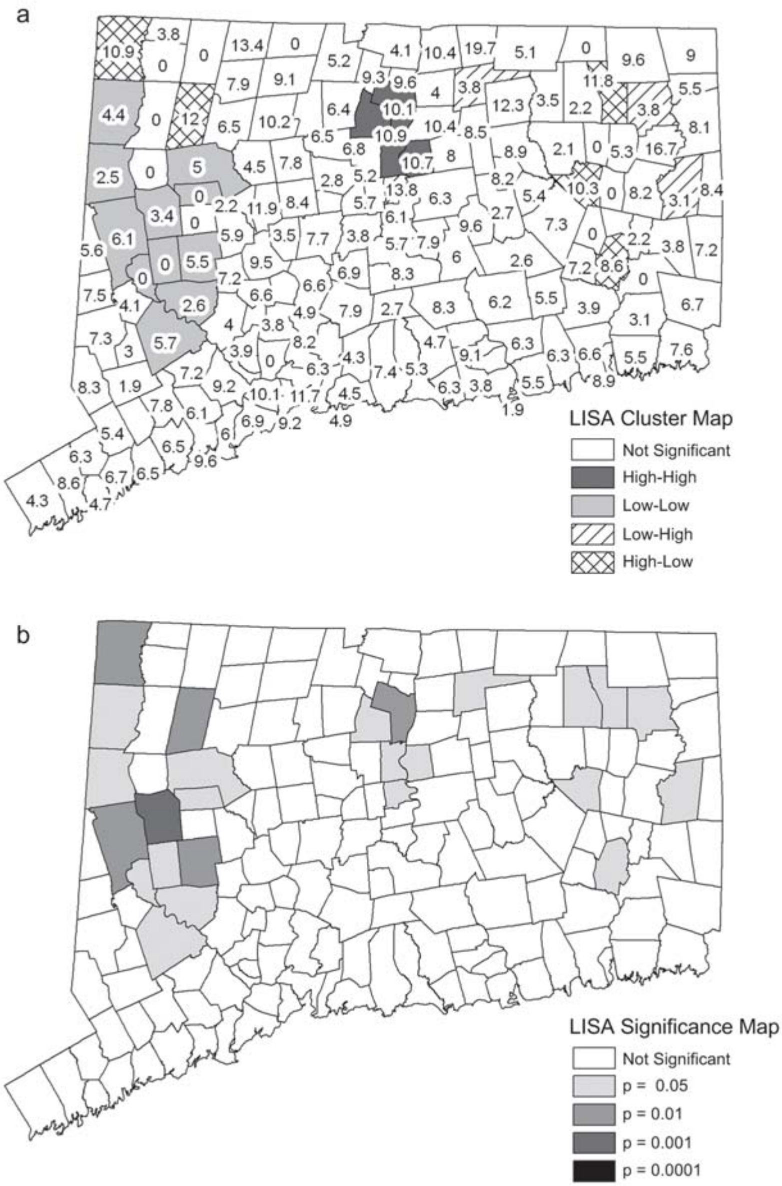


FIGURE 5.7. Low birthweight rate clusters in Connecticut based on LISA. Numbers in Figure 5.7a show the percent of low-birthweight births in each town. The cluster in the north-central area is a cluster of towns with similarly high low-birthweight rates. The cluster in the northwestern part of the state is a cluster of similarly low rates. The hatched areas show towns where the rates are significantly different from rates in surrounding towns. Significance values are shown in Figure 5.7b. Data from the Office of Policy, Planning and Evaluation (2002).

health events modeled as a continuous field (Gatrell, Bailey, Diggle, & Rowlingson, 1996). Although health events occur in particular human or animal hosts, at distinct locations, the risk of ill health exists almost everywhere. Thus, we can view health risk as being distributed continuously over space, with "peaks" representing areas of poor health and "valleys" areas of better health. Kernel estimation is useful for generating such a continuous surface from point data.

In kernel estimation, a spatial window or ***kernel*** is moved across the study area, and the density of events is computed within this window (Figure 5.8). Typically, the window is a circle with a constant radius, or ***bandwidth***. Events within the window are weighted according to their distance from the center of the window, the point at which density is being estimated (Bailey & Gatrell, 1995). The ***kernel function*** describes mathematically how those weights vary over distance. Events located near the center have a greater weight than those distant from the center. In this way, kernel estimation reflects the underlying geographic locations of events within each window.

After computing kernel estimates of the density of health events within each regularly spaced window, one can generate a map of density using standard map contouring procedures. The smoothed surface may be displayed as a contour map, a three-dimensional surface, or as a continuously shaded map with gray or color tones representing density levels (Figure 5.9).

Smoothed maps of health events are useful for showing variation in disease intensity, but they do not assess clustering in relation to risk population. However, we can use kernel estimation for cluster detection by creating a spatially smoothed map of risk population similar to the smoothed map of health events. At each grid point we compare the disease intensity to the intensity of risk population by dividing disease intensity by population intensity using grid calculation tools. Clusters exist when the disease intensity greatly exceeds the population intensity (Han et al., 2005). Recently, researchers have developed randomization procedures for testing the statistical significance of spatial clusters identified based on kernel estimation (Wheeler, 2007).

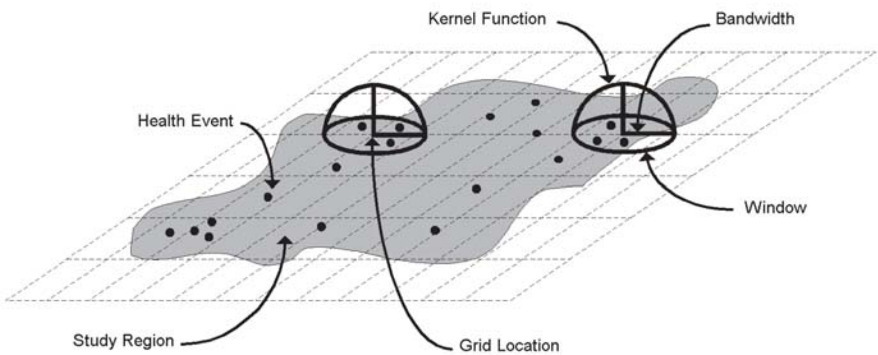


FIGURE 5.8. A schematic of the kernel estimation method.

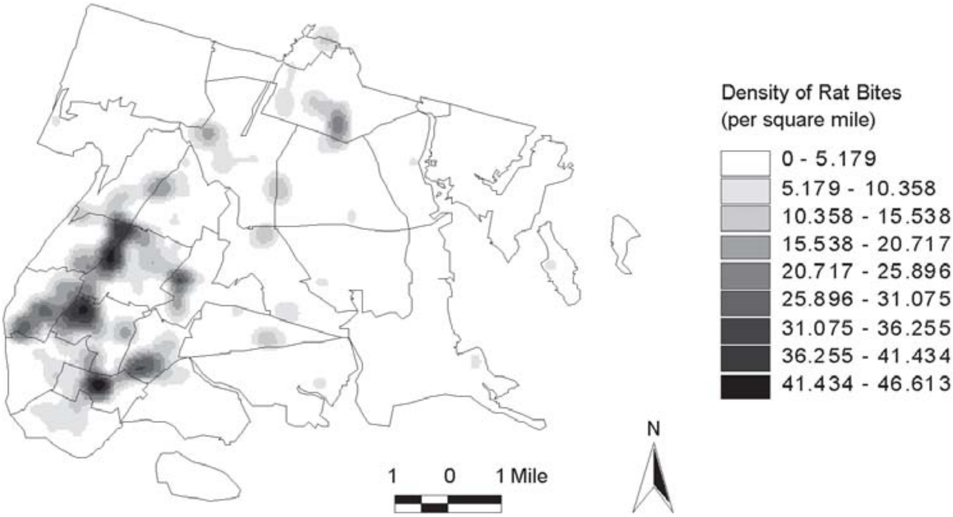


FIGURE 5.9. A contour map, generated by kernel estimation, showing the density of rat bites per square mile in the Bronx, New York. Data provided by the New York City Department of Health.

A key issue in implementing kernel estimation is selection of the bandwidth. Larger bandwidths smooth the data more, removing local variation. In contrast, small bandwidths result in very little smoothing, producing an irregular, “spiky” map. Generally we seek a compromise between these two extremes. One approach is to experiment with different values of the bandwidth and choose the one that gives the best balance between smoothing the data and depicting local variation. Algorithms for finding the optimal bandwidth are also available (Loader, 1999). Another promising approach is to use ***locally adaptive bandwidths***, different bandwidths over different parts of the study area (Brunsdon, 1995; Benschop et al., 2008). Carlos, Shi, Sargent, Tanski, and Berke (2010) describe the use of an adaptive bandwidth that reflects differences in underlying risk population. At each location, the bandwidth expands until it encompasses a threshold level of population. A similar bandwidth selection method is used in SaTScan™ and Disease Mapping and Analysis Program (DMAP) discussed in the next section.

Kernel estimation has been widely used in mapping the uneven spatial distribution of health events and, more recently, in identifying spatial clusters. The method is not only highly effective for visualizing geographic patterns of point health events, but it also offers a flexible tool for cluster detection. Comparisons reveal that kernel estimation is more likely to detect irregularly shaped clusters than are other field-based methods (Wheeler, 2007). Moreover, it can be extended for use in case–control study designs and to incorporate individual-level risk factors (Han et al., 2005).

SPATIAL SCAN STATISTIC (SATSCAN)

Another important and widely used clustering tool is the ***spatial scan statistic*** implemented in SaTScan software (Kulldorff, 1997). SaTScan software is available for free download from the National Cancer Institute (Kulldorff, 2010). Like kernel estimation, the method utilizes a field approach and involves searching the entire study region for clusters. The method begins by laying a regularly or irregularly spaced grid of points across the region. Around each grid point, circular or elliptical windows of different sizes are drawn. For circular windows, the radius of the circle varies continuously from zero to a maximum limit specified by the user, or the radius may be based on risk population size, in which case the radius varies across the map. Within each circle, the method computes the likelihood that the risk of disease is elevated inside the circle compared to outside the circle based on a likelihood ratio test. The circle with the highest likelihood value is the circle that has the highest probability of containing a disease cluster. Monte Carlo randomization is used to evaluate statistical significance of the likelihood ratio for each circle. From the significance tests, SaTScan identifies the most likely spatial cluster (circle) in a data set, as well as secondary clusters that have lower likelihood ratio values. Clusters can be displayed on a map and/or treated as spatial objects for further analysis. Innovative methods have been developed for visualizing SaTScan clusters and analyzing the sensitivity of cluster locations to the choice of radius (Chen, Roth, Naito, Lengerich, & MacEachern, 2008).

The spatial scan method is highly flexible, providing a wide range of options for spatial analysis of public health data. It incorporates alternative statistical models of disease risk, including Poisson and Bernoulli distributions (Kulldorff, 2010). It can adjust for individual-level risk factors (covariates), so that one can detect if spatial clusters remain after controlling for risk factors such as age, gender, or risk behaviors. In addition, SaTScan includes procedures for analyzing clustering in time and space–time. It also permits scanning for clusters within an elliptical window, as opposed to a circle. Clusters that are not compact in shape are more likely to be detected using the elliptical window.

Pollack et al. (2006) used the spatial scan method to identify spatial clusters of late-stage colorectal cancer in California. Their study relied on data from the California Cancer Registry from 1996 to 2000 for all persons diagnosed with colorectal cancer at age 50 years and older. Data were geocoded to point locations based on the residential address at the time of diagnosis. Two spatial clusters emerged from the analysis, one a cluster of higher-than-expected risk of late diagnosis located northeast of San Francisco and the other a cluster of lower-than-expected risk in the San Diego area. High rates of late diagnosis may result from poor access to and knowledge of colorectal cancer screening; therefore the high-risk spatial cluster is a logical priority area for targeting screening and prevention efforts.

DISEASE MAPPING AND ANALYSIS PROGRAM

The ***Disease Mapping and Analysis Program (DMAP)***, developed by Rushton and Lolonis (1996), implements a field-based method that incorporates innovative procedures for visualizing and analyzing cluster significance. Like kernel estimation, it uses a window of constant or adaptive size to scan the study area for clusters; like the spatial scan statistic, it provides information about the likelihood that a cluster might have occurred by chance. The method uses Monte Carlo procedures to simulate possible spatial patterns of health events within a geographically fixed risk population. In effect, it simulates alternative maps of health events that are consistent with a given null hypothesis about disease risk in the population. The null hypothesis is that the risk of disease is constant or that it can be estimated based on known risk factors. For example, in analyzing infant mortality the null hypothesis might be that each infant born has the same risk as any other of dying in the first year of life. The simulations provide a null hypothesis distribution of possible health outcomes in different parts of the study area for comparison with actual patterns. Clusters are places where the actual number of health events is significantly larger than the number found in the corresponding null hypothesis distribution.

DMAP typically utilizes point location data for both individuals at risk and cases of disease to generate simulated patterns of health events. Together, these cases and non-cases form the total risk population. If the incidence of ill health was the same everywhere, or if it was simply based on known risk factors, then each of these individuals would face a predictable risk of ill health. Based on this null hypothesis, the method simulates alternative spatial patterns of health events to provide a benchmark for comparison. For each individual at risk, the method randomly generates an outcome (event/non-event) based on the known incidence rate for the study area. This process is repeated for each individual at risk, creating a simulated pattern of health events within the risk population. The number of events in the simulated pattern need not equal the number of events in the actual data set, although the overall risk is constant.

Rushton and Lolonis (1996) recommend generating several thousand of these simulated patterns. As with other field methods, a regular grid of points is superimposed on the study area each time a simulated pattern is generated. Overlapping circular zones are created around each grid point (Figure 5.10). Within each of these circular zones an incidence rate is computed for each simulated pattern. The full set of 1,000 or more simulated rates serves as a benchmark distribution for that circular zone. The zone's actual rate is compared to its simulated rates, and the percentage of simulated rates that are less than the actual rate provides a measure of significance. A high percentage means that, in that zone, the vast majority of simulated rates fall below the actual rate, indicating that the actual rate is unusually high. The percentages can be displayed on an isarithmic map to show areas with significantly high rates.

This method involves an interplay of GIS operations and statistical simulation procedures. The GIS functions are relatively straightforward: first, create

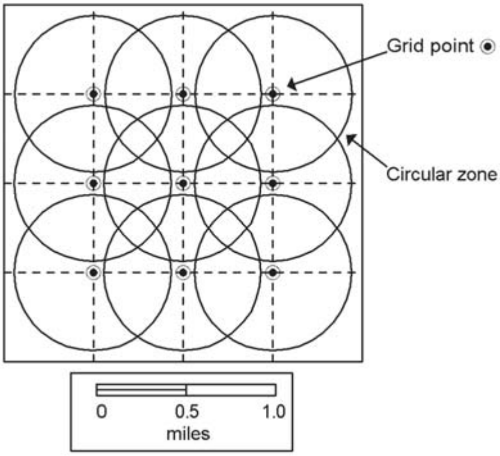


FIGURE 5.10. Overlapping circular zones generated around grid points in the Rushton and Lolonis method of analyzing clusters. From Rushton and Lolonis (1996). Copyright 1996 by John Wiley & Sons, Ltd. Reprinted by permission.

grid points and circles, and second, count the numbers of events and non-events within each circle. After that it is necessary to generate the simulated, random patterns of health events. Using a random number generator, the analyst generates a simulated outcome (either event or non-event) for each individual at risk. Then the system is used to compute the numbers of simulated events and non-events within each circle to derive a simulated rate. This is repeated thousands of times to determine the benchmark distributions for comparison with actual rates. If the GIS software cannot perform statistical simulations, it is necessary to use statistical programs or to write specialized software for this purpose.

Rushton and Lolonis (1996) used this method to analyze spatial clustering of birth defects in Des Moines, Iowa. Local health officials were concerned about poor infant health indicators in the city and wanted to be able to identify neighborhoods where birth defect rates were significantly high. From birth records, the researchers knew the residential locations of all births in the city—the risk population—as well as the locations of infants born with birth defects—the health event. One thousand simulated patterns of birth defects were generated. Figure 5.11 shows areas where the actual birth defect rate exceeded 95% of the simulated rates. The map depicts an elongated cluster of high rates in the east-central portion of the city.

The DMAP cluster analysis tool has been extended in several important ways since the mid-1990s. DMAP no longer requires a regular point grid for estimating health risk. Instead, one can choose to densify the grid—increase the density of grid points in some areas—so that highly populated areas contain a higher density of grid points than less populated areas (Cai, 2007). This makes it possible to observe finer-grained spatial clusters of ill health in densely popu-

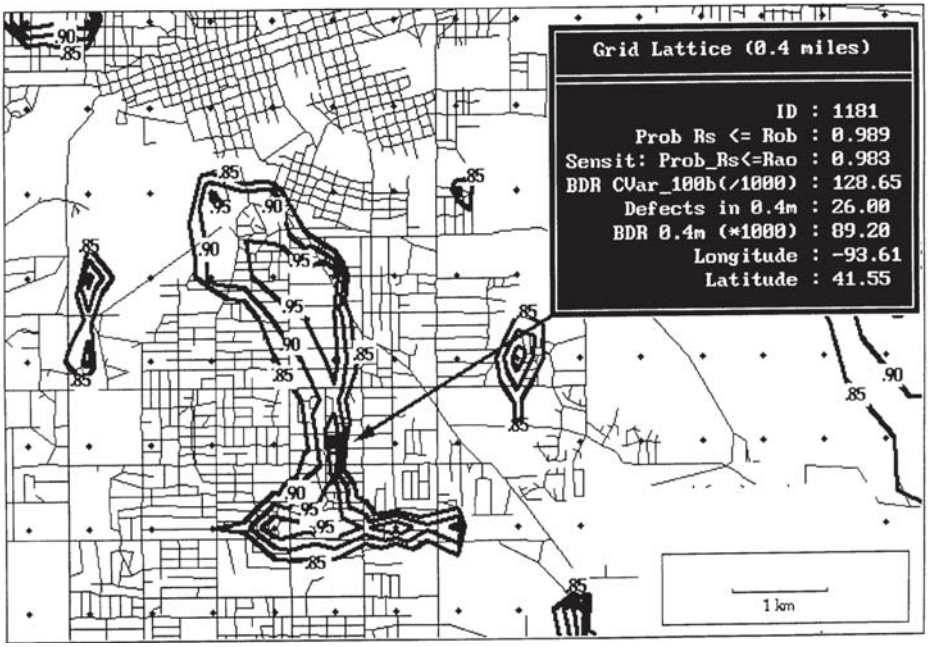


FIGURE 5.11. Areas with statistically significant high rates of birth defects in Des Moines, Iowa, based on the Rushton and Lolonis method. From Rushton and Lolonis (1996). Copyright 1996 by John Wiley & Sons, Ltd. Reprinted by permission.

lated urban areas. DMAP now can implement an adaptive bandwidth in which the size of the circular zone at a particular grid point varies inversely with the size of the local risk population. The bandwidth is increased until it encompasses a certain threshold population. Thus, a smaller bandwidth is applied in highly populated urban areas than in rural areas (Figure 5.12). This spatial adaptive filtering approach addresses the small numbers problem that arises when using a fixed bandwidth in areas of uneven population density. It ensures that each zone contains a sufficient number of cases to accurately estimate health risk. The only disadvantage is that the spatial scale of analysis varies across the map, which complicates interpretation of results.

Object-Based Spatial Clustering Methods

Unlike the methods described in the previous section, which search for clusters by scanning the entire study area, object-based methods attempt to "construct" spatial clusters through a process of aggregation. Nearby cases of disease, or nearby areas that have high counts or rates of disease, may be aggregated together. Objects—points or areas—serve as the building blocks for identifying clusters.

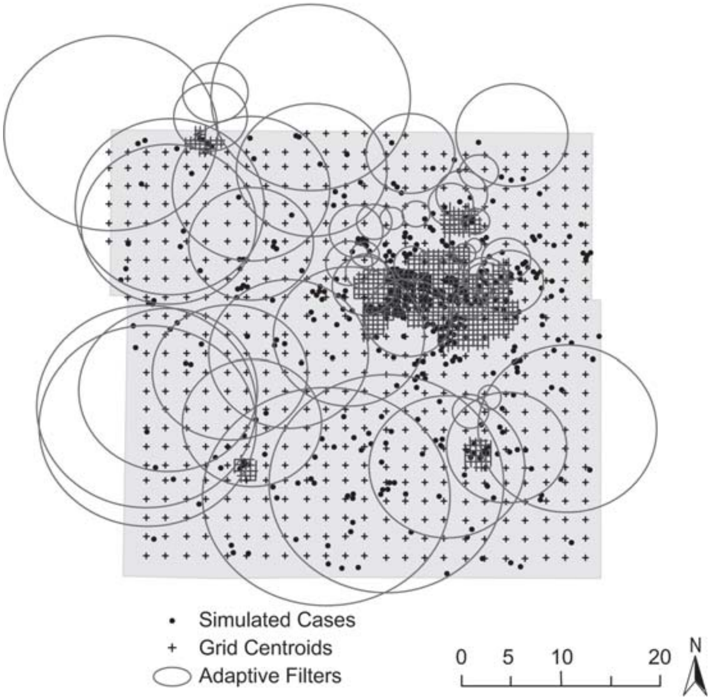


FIGURE 5.12. Spatial adaptive filters differ in size, with each filter including the same number of expected cases. From Cai (2007). Reprinted by permission of Qiang Cai.

BESAG AND NEWELL'S METHOD

Besag and Newell (1991) devised a spatial clustering method that only searches for clusters around cases. Their logical premise is that *clustering can only exist in places where cases exist*, so instead of searching over the entire study region, we need only search around cases. This greatly reduces the amount of spatial search and computation required over field methods, and it provides a finite limit to the number of significant clusters that can be detected.

Assume that we have point data for cases but area data for risk population. The actual locations of health events are known, while data on risk population are available for areas such as counties, census tracts, or blocks. Let k be the minimum number of cases needed to constitute a disease cluster. Besag and Newell's method tests for clustering around a case i by analyzing the number of nearest neighbor areas (M_i) needed to accumulate the k cases closest to i. In other words, if we rank cases according to their distance from i and identify the k nearest cases, M_i identifies the geographic areas that contain those k cases (Figure 5.13). Defining x_j as the number of cases in area j and p_j as its risk population, then the total number of cases in M_i is $X_j = \Sigma x_j$ and the total risk population is $P_i = \Sigma p_j$. To test for clustering around i, we analyze whether the total number of

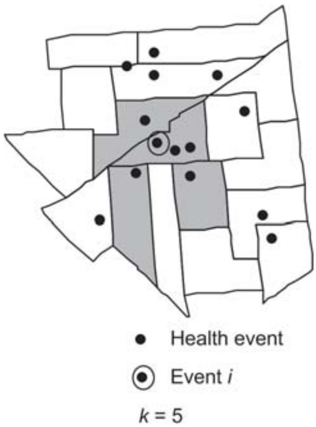


FIGURE 5.13. The Besag and Newell method searches around each health event *i* to find the *k* nearest health events. The areas containing those events are shaded. The risk population within the shaded areas is the denominator for the Poisson test.

cases in M_i is large relative to total risk population. Once again the Poisson test is used to assess significance, with λ set at the incidence rate for the entire study area.

Besag and Newell's method involves a sequence of operations that can be implemented in GIS. At each case location, we first compute distances to all other cases and rank the distances to determine the k nearest cases. Point-in-polygon operations identify the areas in which the k nearest cases are located to identify the set M_i for case i. We then sum cases and risk population across those areas and calculate the statistical test to determine whether the number of cases is significantly high relative to the risk population. If prevalence is high, one can draw a circle encompassing that cluster and display the cluster on a map. Clusters often appear as overlapping concentrations of circles.

A critical issue in Besag and Newell's (1991) method is the choice of k, the cluster size parameter. The value of k must be large enough to identify real clusters of cases, not just isolated groupings. But it must be small enough to permit the identification of distinct clusters within a region. Values of between 2 and 5 are common, representing clusters of 3 to 6 cases including the centroid case. Besag and Newell recommend trying several k values and analyzing the sensitivity of the results to the choice of k.

A modified version of Besag and Newell's method was used to search for spatial clustering of breast cancer cases among long-term residents of West Islip, New York (Timander & McLafferty, 1998). West Islip is a middle-income community of approximately 40,000 people located on Long Island. As in many communities on Long Island, residents of West Islip were worried about their high rates of breast cancer and possible links to environmental factors: hazardous waste sites, contaminants in the water supply, and electromagnetic fields. Taking the matter into their own hands, they conducted a survey of local residents to find out about breast cancer prevalence and risk factors (Grimson, 1999).

The survey data were entered into a GIS, geocoded, and displayed on a community map. To see if cases were clustered within West Islip, a modified version of Besag and Newell's (1991) clustering method was used and implemented in a GIS. Because breast cancer takes many years to develop, the research focused on long-term residents, survey respondents who had lived at their current addresses for more than 25 years. These people would be most affected by hazards in the local environment near their homes. The GIS analysis showed no strong evidence of spatial clustering of breast cancer in West Islip. Four significant overlapping clusters were uncovered in the south-central portion of the town (Figure 5.14), but the clusters disappeared when known risk factors for breast cancer, like family history and age at first pregnancy, were controlled.

Compared to other widely used methods, Besag and Newell's (1991) procedure provides a clearer description of cluster locations. Because the method only checks for clusters around cases, it is more conservative in detecting clusters and is less likely to identify false positives (Fotheringham & Zhan, 1996). By limiting the search process, it is not computationally intensive and thus represents an efficient option for cluster detection.

AMOEBA METHOD: DETECTING IRREGULARLY SHAPED CLUSTERS

The cluster detection methods discussed so far work well in detecting clusters that are compact and regular in shape; however, disease clusters may well be irregularly shaped, reflecting the uneven social and environmental geographies of risk. Several new methods have been developed which are capable of detecting irregularly shaped clusters (Tango & Takahashi, 2005). These methods are object-based in that they detect clusters by agglomerating nearby geographical units that have similarly high rates or numbers of health events. This discussion focuses on the AMOEBA method developed by Aldstadt and Getis (2006), which works with area-based indicators of health or ill health.

Starting from a "seed" area, AMOEBA attempts to grow a cluster outward from the seed location. The algorithm checks areas contiguous to the seed to determine whether the areas' health indicators are similar to that of the seed. Similarity is defined according to the local G_i^* statistic discussed earlier, a widely used measure of local spatial autocorrelation. If adding in the contiguous areas results in a more extreme value of G_i^* (either positive or negative), then the areas are grouped with the seed, thus creating a cluster. All combinations of contiguous areas are evaluated, and the combination that maximizes the absolute value of the local G_i^* statistic is preserved as a cluster. This cluster becomes a new seed for another iteration of the method. As new areas join the cluster, those areas become new starting points for checking contiguous areas. The procedure stops when the value of G_i^* cannot be improved by adding contiguous areas to the cluster.

At that point, AMOEBA shifts to a different seed location and again tries to construct a cluster by agglomerating contiguous areas. The process continues

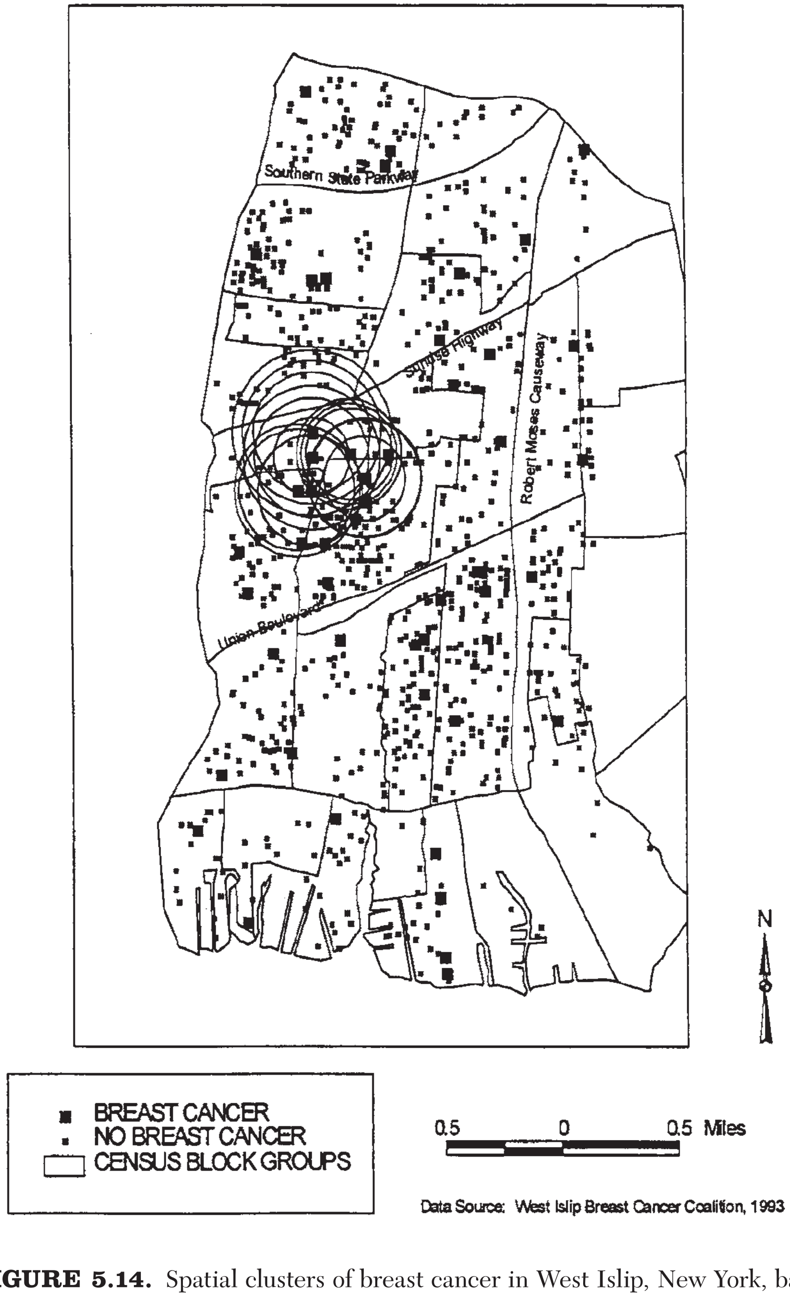


FIGURE 5.14. Spatial clusters of breast cancer in West Islip, New York, based on a modified Besag and Newell method, for $k = 4$, 5, and 6. There were four significant circles in the set of circles for which $k = 4$; there were six significant circles each in the sets of circles for which $k = 5$ and $k = 6$. Reprinted from *Social Science and Medicine, 46*(12), Timander, L., & McLafferty, S., Breast cancer in West Islip, NY: A spatial clustering analysis with covariates, 1623–1635, 1998, with permission from Elsevier.

until the entire study region has been scanned for clusters. The method focuses on identifying geographically distinct clusters, so clusters are checked for overlap. In the end, the nonoverlapping clusters with the highest absolute G_i* values are designated as important clusters, and their statistical significance is tested using Monte Carlo randomization methods.

The main advantage of AMOEBA is its ability to identify clusters of any shape. Comparisons with circle-based methods such as SaTScan reveal that AMOEBA is much more effective at identifying noncompact, irregularly shaped clusters (Figure 5.15). Furthermore, AMOEBA can incorporate any measure of spatial statistical association as the basis for agglomeration, not just the G_i* statistic.

Space–Time Clustering

Time is an important issue in spatial clustering analyses. It is an essential component of cluster definition: a cluster only exists with reference to a specific place and time. Health surveillance often involves scanning real-time health data to

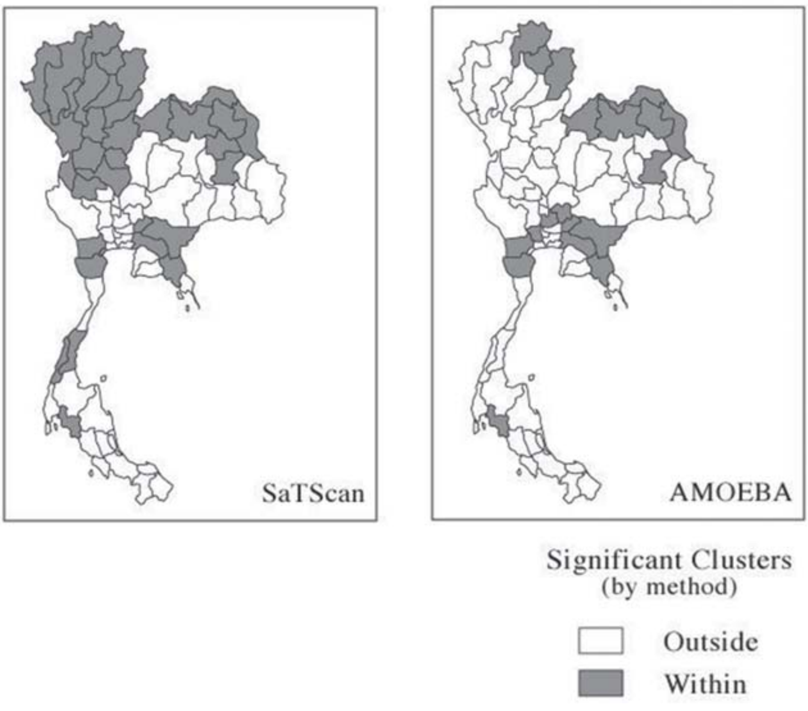


FIGURE 5.15. A comparison of statistically significant clusters found by the SaTScan procedure and the AMOEBA procedure (p ≤ .05). From Aldstadt (2007). Reprinted with Author's permission.

identify ***space–time clusters***, clusters that emerge during particular time intervals at particular places. One can record time in a variety of ways with respect to health data. Typically, time refers to the day/year when a health condition was first diagnosed or the day/year of first onset of symptoms. In a GIS database, time can be recorded in a time field that contains the year and date of occurrence. Other methods for incorporating time in GIS databases are discussed in Chapter 7.

To identify space–time clustering, many of the methods discussed in previous sections can be and have been extended. Generally, these extensions involve the use of a ***time window***—a specified time interval—in addition to the spatial neighborhood window used in spatial clustering tests.

SaTScan, for example, offers statistical tests for spatial, temporal and space–time clustering (Kulldorff, 2010). In the latter, a space–time cylinder scans the study area over time in search of unusually high concentrations of health events. The base of the cylinder represents the circular spatial window for assessing spatial clustering, and the height of the cylinder represents the time interval for analyzing temporal clustering. Noncylindrical search windows can also be used to incorporate the possibility that the spatial extent of spread varies over time. Figure 5.16 shows an application of SaTScan to analyze space–time clustering of a particular fine-scaled genetic type (finetype) of meningococcol disease in Germany. A significant cluster of four cases appears in the second time period.

Similar approaches to analyzing space–time clustering involve local variations of the Knox test which analyzes the number of pairs of cases that are close in space and time. Rogerson (2001) developed a method that tests for local clustering within a varying space–time window. The DYCAST model, which has been used to examine space–time clustering of dead birds and human infections in a West Nile Virus outbreak, adopts a similar approach (Theophilides, Ahearn, Binkowski, Paul, & Gibbs, 2006). These methods share a set of common GIS operations. They involve first laying a regular grid across the study area or identifying a set of discrete points at which clustering will be assessed, then scanning that space to evaluate space–time clustering within a prespecified time–space window. Clearly the spatial and temporal dimensions of the window should reflect knowledge of the underlying disease process. In the West Nile Virus (WNV) study, distance parameters were chosen to reflect the limited spatial mobility of birds that were infected with WNV, and the time dimension reflected the disease incubation period (Theophilides et al., 2006). Another approach is to vary the spatial and temporal dimensions over a range of values and investigate clustering in an exploratory framework.

Space–Time Clustering and Residential Mobility

The methods discussed thus far only associate a single location, usually the residential location at the time of diagnosis, with each health event. This ignores the important concepts of latency and migration, which affect peoples' exposures to social, environmental, and biological risks and the short- and long-term effects

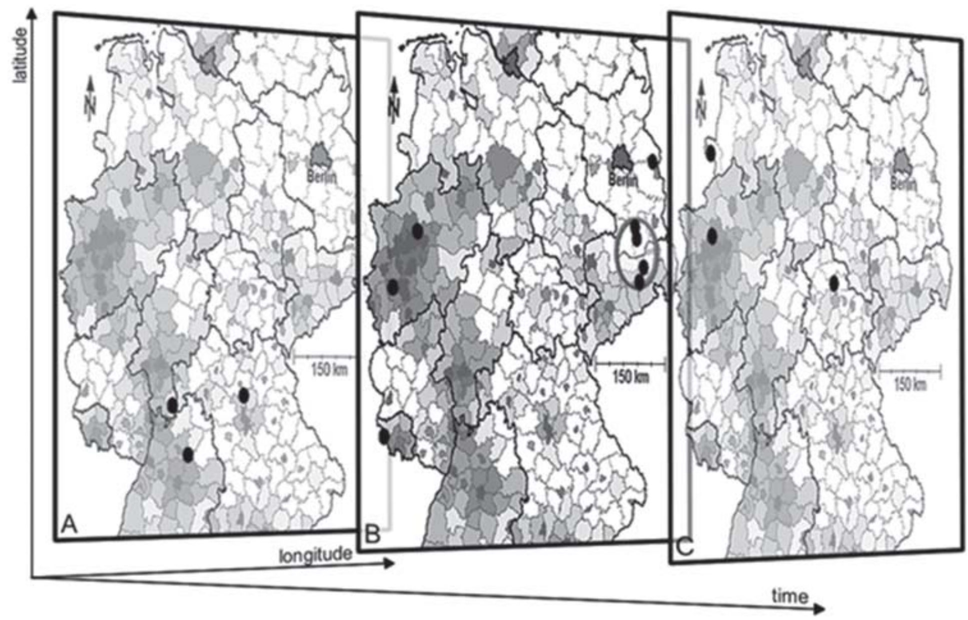


FIGURE 5.16. Using SaTScan for retrospective cluster identification over time. Planes A, B, and C show consecutive temporal windows of 30 days in 2003. Planes A and C do not show spatial clustering, but Plane B shows an accumulation of four cases of meningococcal disease (of the finetype C:P1.5,2:F3–3) in two counties within a circular region encompassing a population of 339,185. The counties of Germany are shaded according to their population densities. From Elias et al. (2006).

on health. ***Latency*** refers to the length of time between exposure to a disease-causing agent and detection or diagnosis. Latency is typically divided into two periods: the ***induction period,*** which is the time between exposure and initiation of disease, and the ***latent period,*** which is the time between disease initiation and disease detection. While some health problems develop quickly, others (e.g., many cancers) take months or years to develop.

Migration bias refers to the effects of migration on the results of a clustering analysis. Most cluster methods only work with single addresses at single points in time. Yet people in the United States move frequently, and they face different environments at each place of residence. A person's current address may have little connection to his or her environmental exposures over the life span. Migration effects are most relevant for diseases with long latency periods that result from environmental, social, behavioral, and genetic factors interacting over long periods of time.

One way to address these issues is to analyze spatial clustering at several distinct points in the life cycle. For example, Sabel et al. (2003) investigated spatial clustering of amylotrophic lateral sclerosis cases in Finland at the time of birth and time of death. The close match between clusters identified based on birthplace and place of death led the authors to hypothesize the role of genetic or environmental factors in disease causation. An alternative approach is to focus selectively on long-term residents and analyze spatial clustering of disease within that "immobile" population, as was done in the West Islip breast cancer study. However, such an approach relies on a potentially ad hoc definition of "long-term" resident, and like other methods, it may be affected by selective migration out of the study area.

More recently, new methods have been developed that analyze the clustering of health events in relation to individual space–time life paths (Jacquez et al., 2005). These methods make it possible to assess whether people who have a disease might have been exposed to a common environmental or infectious agent when they lived or worked nearby earlier in their lives. The building block of these methods is the ***space–time path***—a three-dimensional representation of a person's movement or migration history through space and time (Figure 5.17). Jacquez et al. (2005) proposed a procedure that searches for overlapping segments of space–time paths, segments where people lived nearby in time and space. The procedure is implemented in a case–control study design that enables evaluation of space–time clustering. Overlap is evaluated in relation to a predefined induction period that defines the interval between exposure and disease initiation. The test involves counting the number of nearest neighbors of a particular case whose induction periods overlap the case's induction period. Similar counts are made for controls as well, providing a benchmark for comparison. Large numbers of nearest neighbor cases, as opposed to controls, signify potential disease clustering, and conditional randomization procedures enable significance testing.

The methods proposed by Jacquez et al. (2005) have many features that enhance public health investigations. They can incorporate individual-level information on risk factors, covariates, latency periods, and exposures. They can also be used in an exploratory way to identify the induction and latency time periods that maximize space–time clustering among cases (Jacquez, Meliker, & Kaufmann, 2007). The main challenge in implementing these methods stems from data requirements. Detailed migration histories for cases and controls are needed for model input, but are rarely available. One can generate such information from sample surveys; however, small sample sizes and sampling bias may be problematic given the wide range of time periods and geographic areas that are likely involved in making inferences about clustering. Another option is to make use of the rich data available in countries that have high-quality population registries, such as Denmark and Sweden.

Choosing a Clustering Method

How does a public health analyst choose among these diverse methods? Characteristics of the data to be analyzed affect this choice. Spatial data characteristics

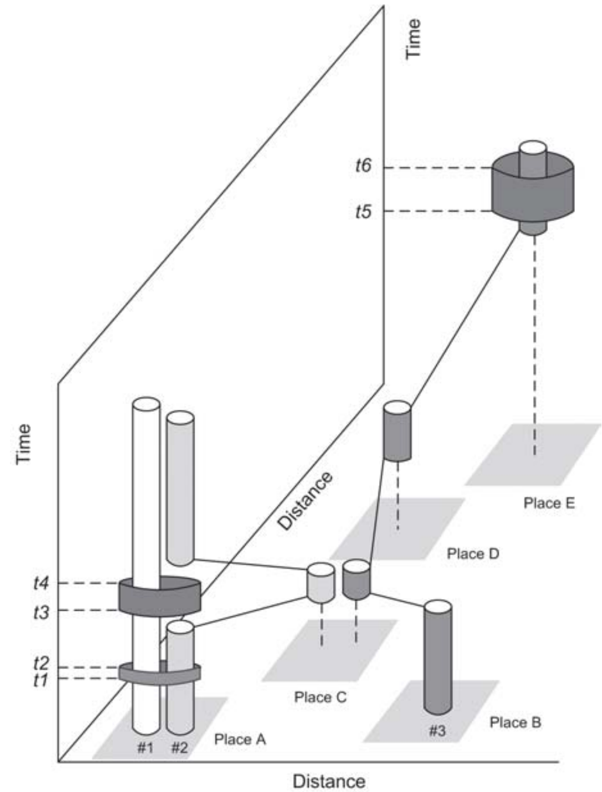


FIGURE 5.17. A topology of residential histories for three people in time and space. Person #1 resided in Place A his entire life. Person #1 was Person #2's neighbor, except when Person #2 attended college in Place C about the same time as Person #3 who grew up in Place B. After college, Person #3 moved to Place D and then to Place E. A brief contamination event in Place A from *t*1 to *t*2 exposed both Person #1 and Person #2. A later contamination event in Place A, which was of greater duration from *t*3 to *t*4, exposed only Person #1 because Person #2 was away at college. The contamination event at Place E from time *t*5 to *t*6 affected only Person #3.

are important. Some spatial clustering methods like kernel estimation are most appropriate for point data, whereas others such as G_i*, LISA, and AMOEBA are suitable for area data. However, the lines between area and point methods are blurring insofar as area data can be converted to points via area centroids, and point data can be converted to area data by Thiessen polygons or by aggregating points into user-defined zones. Thus, with a few exceptions, most data sets can be made to fit most methods. Of course, data conversion introduces bias and errors that may outweigh the benefits of using a particular method. These trade-offs need to be carefully considered in selecting a spatial clustering method.

Software availability is another important factor. A spate of specialized clustering software has been introduced in the past decade, and many programs are downloadable at no cost (Table 5.2). Some of these programs include easy-to-use interfaces and visual displays, whereas others require import of data from and export of results to GIS for mapping and analysis. In addition to these stand-alone software packages, many methods can be implemented using spatial analysis tools within the R Project for Statistical Computing (Bivan, Pebesma, & Gómez-Rubio, 2008). These are a good option for users with high-level statistical knowledge. On the other hand, commercial GIS currently incorporate a limited set of spatial clustering methods in their spatial analysis toolkits, with the methods varying from one commercial GIS software product to another. Users with programming skills can write macros in some GIS software to implement more advanced clustering procedures. In sum, given the wide and increasing array of downloadable tools for spatial cluster analysis, software availability no longer poses a major barrier to mapping and analyzing disease clusters.

How well do the various methods perform? Comparisons suggest that each method has advantages and disadvantages (Fotheringham & Zhan, 1996; Conley, Gahegan, & Macgill, 2005; Tango & Takahashi, 2005). The methods differ in their ability to detect clusters of different shapes and sizes and in areas of varying risk population density. Some methods include options like adaptive bandwidths and adjusting for covariates that may be important for specific kinds of public health investigations. There are also important differences in computational efficiency, technical requirements, and capabilities for data input, output, and geovisualization. Clearly, the match between method and requirements or capabilities of specific cluster investigations is critically important.

The theoretical basis of clustering methods also underlies the choice of a modeling strategy. Most methods make somewhat naive assumptions about the environmental or social processes that generate observed patterns. Clusters of

TABLE 5.2. Software for Spatial and Space–Time Clustering

Program	Availability	Web Link
GeoDa	Shareware	*geodacenter.asu.edu*
SaTScan	Shareware	*www.satscan.org*
DMAP	Shareware	*www.uiowa.edu/~gishlth/DMAP4*
R Analysis of Spatial Data	Shareware	*www.r-project.org/index.html* *cran.r-project.org/web/views/Spatial.html*
ClusterSeer	Commercial	*www.terraseer.com/products_clusterseer.php*
CrimeStat	Shareware	*www.ojp.usdoj.gov/nij/maps/crimestat.htm*

Note. ClusterSeer and CrimeStat are additional software, not discussed in the text, which may be of use in analyzing clusters of health events.

motor vehicle accidents will be arrayed in a linear pattern constrained by the road network. Clusters that are environmental in origin may follow contamination footprints or plumes of air pollution. Communicable disease clusters are shaped by human interactions that are unlikely to fit simple geometric forms. Spatial clustering methods that incorporate these process-based understandings are only beginning to be developed. Use of GIS for investigating specific types of health concerns is discussed in Chapters 6 through 8.

Uses of Spatial Clustering Methods

Spatial clustering methods are exploratory tools that help researchers and policymakers make sense of complex geographic patterns. Knowing whether or not clusters exist and where they are located provides an important foundation for health research and policy formulation. State health departments receive literally thousands of requests for spatial cluster investigations from individuals and community groups every year (Greenberg & Wartenberg, 1991). The methods discussed here can be used to confirm or deny the existence of suspected clusters in an efficient and effective manner.

Responding to community concerns, however, only deals with a fraction of potential clusters and is likely to miss clusters in communities that lack political and economic clout. To overcome this bias, cluster detection methods are increasingly being incorporated in ongoing public health surveillance efforts (Rushton, 2003). With GIS, health departments can monitor health records as they come in and use spatial and space–time clustering methods to search for unusual patterns. These systems are being implemented at the local, state, and national levels for rapid detection of disease outbreaks (Mandl et al., 2004). Of course, once a cluster is identified, only detailed epidemiological investigation can determine whether the cluster is a random event or whether it is linked to some environmental, occupational, or social cause. A small number of clusters will occur by chance, even if health risks are not elevated. Therefore, it is essential that statistically significant clusters be examined in more detail. Spatial cluster analysis is not an end in and of itself; it is a screening tool that assists in guiding public health surveillance efforts.

Incorporating spatial clustering methods in GIS also makes it possible to conduct exploratory analyses to help in identifying the causes and correlates of health problems. Overlaying cluster maps with other spatial databases, including environmental, social, transportation, and facilities data, can provide clues about the causes of disease, while identifying variations in health linked to differences in physical and social environments. These map overlays have always been important hypothesis-generating tools in public health research and policymaking (Croner et al., 1992). Now they can be implemented efficiently and effectively in an automated GIS environment. These overlay procedures are discussed in more detail in the chapters that follow.

CONCLUSION

The past decade has witnessed major advances in our ability to model and detect spatial clusters of health events. In fact, all of the issues presented as challenges to spatial cluster analysis in the first edition of this book are currently being addressed in development of innovative modeling strategies and software tools. There is also a rapidly expanding knowledge base of empirical case studies that use clustering methods for formulating and analyzing public health policies and for understanding the causes of health inequalities. Analysts who want to embark on a spatial cluster investigation have a rich array of resources at their disposal.

This chapter has discussed a representative set of methods for analyzing spatial clustering of health events including field- and object-based approaches to cluster identification and the analysis of clustering in both space and time. Many other methods exist, and we encourage readers to look beyond this chapter to find the method that best fits their data and clustering problem. Good general references include Kingham, Gatrell, and Rowlingson (1995), Fotheringham, Brunsdon and Charlton (2000), Waller and Gotway (2004), and Rogerson and Yamada (2009). What does the future hold? Given the rapid advances in methods and software development and the rapidly emerging geovisualization capabilities of systems like Google Earth, it is likely that analysts will soon be able to link spatial cluster investigations to rich and detailed geospatial information about the places where clusters occur. This will give public health departments an expanded and enhanced set of tools for performing one of their most basic surveillance tasks: the search for clusters of health problems.

CHAPTER 6

Analyzing Environmental Hazards

Environmental health, "the prevention and control of health problems related to the environment" (Thacker, Stroup, Parrish, & Anderson, 1996, p. 633), is an important function of health organizations at all levels. Environmental health problems involve agents that produce adverse health outcomes in humans. These agents can be physical (ultraviolet and ionizing radiation), chemical (lead), or biological (cryptosporidium) in nature. Human populations encounter these agents by breathing, eating and drinking, or coming into physical contact with agents present in the atmosphere, the food and water supply, and the natural and built environments. GIS modeling of health problems involving biological agents is discussed in Chapters 7 and 8. This chapter focuses primarily on physical agents and on ***toxicants***, natural or synthetic chemicals that produce adverse health outcomes.

The process by which an agent in the environment produces an adverse health outcome in a person can be modeled as a hazard–exposure–dose–response process (Figure 6.1). Information systems for monitoring environmental health problems viewed in this way require longitudinal data on the amount, nature, and sources of environmental hazards; the environmental quality in the places where people live and conduct their daily activities; the presence of the agents in human populations; and the adverse health outcomes that can be linked to exposure (Mather et al., 2004; Smolders & Schoeters, 2007). In cases in which these data exist for particular agents, exposure, dose, and response data may not always be available from the same single data source. Instead, data must be drawn from multiple sources and integrated in the surveillance system. Although not always explicit, time and space are the basis for data integration in a way that logically models hazard–exposure–dose–response processes. Integration of geographic data drawn from many sources is one of the main uses of GIS.

This chapter discusses GIS applications in environmental health. GIS have been used to display sources of environmental contaminants of concern for human health and to model the zones of contamination around these sources. Geographic variations in environmental quality measured at monitoring stations

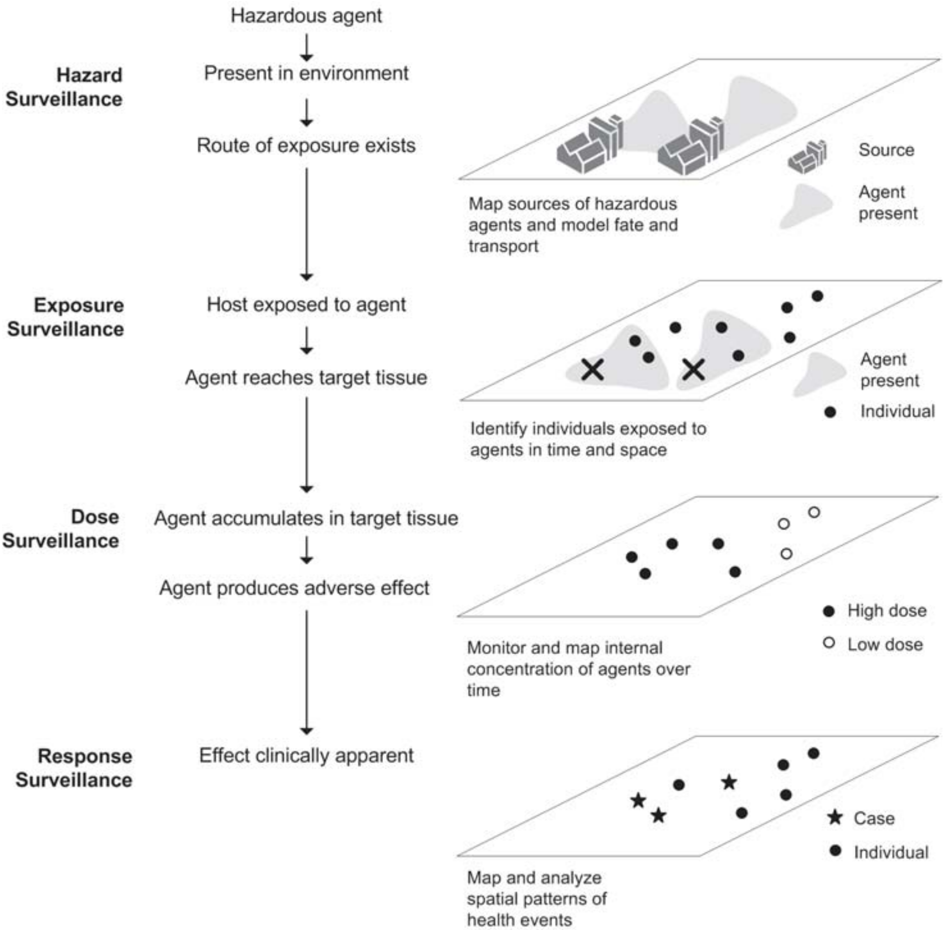


FIGURE 6.1. A geographic model of the hazard–exposure–dose–response model.

have also been modeled using GIS. The geographical distributions of populations at risk and spatial patterns of health outcomes, reviewed in Chapter 4, have also been reported and analyzed using GIS technology.

These applications suggest that complete environmental health surveillance systems are in place for only a few of the thousands of potentially hazardous agents. We have not identified all hazardous agents with demonstrated links to specific diseases, let alone described their presence in the environment in relation to susceptible populations. Somewhat more complete data are available on health outcomes, and these have been used to conduct epidemiological investigations to identify potential hazardous agents. This means that, for particular environmental health problems, surveillance systems will differentially emphasize hazard, exposure, or outcome surveillance.

The first part of this chapter briefly considers how an agent is identified as hazardous based on risk assessment and subsequently becomes a focus for reporting and regulation. The geography of hazard or potential hazard sources and the role of GIS in hazard surveillance are discussed. The following sections in the chapter address some of the most challenging aspects of environmental health surveillance using GIS: designing environmental health surveillance systems, modeling exposure of susceptible populations to environmental contaminants, and linking dose and health response data to exposure.

The final sections of the chapter consider the role of GIS in risk management and issues in mapping environmental hazards. ***Risk management*** involves the selection and implementation of appropriate strategies for the regulation or control of identified hazards based on social and political factors (Ruckleshaus, 1983). Risk assessment and risk management are complementary.

How Environmental Agents Are Identified as Hazards

Before the source locations of hazards and the associated contamination fields can be modeled in a GIS, hazardous agents must be identified. Human experience has been an important source of awareness of the harmful effects of many toxicants. Our recognition of lead, mercury, and other chemicals as hazards to human health is centuries-old. The establishment of quantitative standards for risk, however, is largely a development of the last 100 years. In part, this reflects the growth and geographical dispersion of chemical production on an industrial scale (Cutter, 1993). In addition to naturally occurring chemicals, approximately 4,000 synthetic chemical substances are in commercial use worldwide, accounting for more than 99% of the total volume of synthetics (Moochhala, Shahi, & Cote, 1997). To cite just one example, thousands of pesticide products containing at least one of more than 1,000 registered active ingredients are commercially available (U.S. Environmental Protection Agency, 2007). With the production, storage, transportation, use, and disposal of all these chemicals have come increased risks to workers, other susceptible populations, and the environment.

Quantitative risk assessment is the process of characterizing the health effects expected from exposure to an agent, estimating the probability of occurrence of health effects, estimating the number of occurrences in a population, and recommending acceptable concentrations of the agent in air, water, or food (Hallenbeck, 1993). In the United States, the development of quantitative standards for risk grew out of the Pure Food Act of 1906 (Hattis, 1996). Subsequent federal legislation has been the major impetus for conducting risk assessments. The statutes and their amendments have not generally defined “acceptable risk”; the responsibility for determining acceptable risk was left to the various federal regulatory agencies implementing the legislation.

Globalization of economic activity and economic shifts are affecting environmental regulation. In the 1970s and 1980s, U.S. environmental laws like the Clean Air and Clean Water Acts and the regulations implementing them were taken as models by other countries (Schapiro, 2007). Over the last decade, the European Union has emerged as the world leader in setting standards for the global economy on issues ranging from financial accounting standards to safety to screening for toxic chemicals. The European approach is based on the ***precautionary principle*** of European Union law (Fisher, 2007). According to this principle, the absence of full scientific certainty is not used as a reason for failing to act when there is a body of evidence that the risk of serious or irreversible harm to public health or the environment exists. As a result, many substances banned in Europe and subsequently banned in other countries adopting European standards are allowed to be used in the United States, even though agencies in both countries have reviewed the same scientific studies and have access to toxicity data.

Three basic categories of scientific information are used in quantitative risk assessment: toxicological studies, controlled clinical studies, and epidemiological studies (Moochhala, Shahi, & Cote, 1997). Each of these approaches has advantages and disadvantages. GIS can perhaps provide the greatest support to quantitative risk assessment through epidemiological studies. Toxicological and controlled clinical studies are conducted in laboratory or clinical settings.

Toxicological studies have been a major and controversial source of information for risk assessment. ***Toxicology*** is an experimental science that studies the effects of toxic substances in selected animals or cells. Toxicological studies provide the greatest degree of control over populations exposed, exposure conditions, and measured effects. They are most often used to evaluate agents for which epidemiological studies would be premature because of the lags between exposure and outcome and for which controlled clinical studies would be unethical. Animal testing may also be a regulatory requirement, as in the case of premarket toxicological testing of pesticides (Alavanja, Hoppin, & Kamel, 2004). Aside from the ongoing public debate over the ethical treatment of animals in human health research (Rollin, 2003), the major scientific limitation of toxicology studies is uncertainty in extrapolating the exposure–outcome relationships observed in animals to humans. Screening approaches are clearly needed to keep dangerous products from making it to the market, but evidence from epidemiological studies suggests that many products that meet regulatory approval requirements based on animal testing have significant human health effects (Alavanja, Hoppin, & Kamel, 2004).

Controlled clinical studies, like toxicological investigations, provide the opportunity to control and quantify exposure but focus directly on the effects of agents on the health of human subjects instead of animals (Holgate, Sandström, et al., 2003; Holgate, Devlin, Wilson, & Frew, 2003). The U.S. National Ambient Air Quality Standards for ozone and sulphur dioxide (SO_2) were developed in part based on controlled clinical studies of changes in airway resistance of asthma sufferers who were exposed while exercising (McDonnell et al., 1991). These effects would be difficult to detect in an epidemiological investigation of the general population. The major limitation of controlled clinical studies in producing data for risk assessment is that, for ethical reasons, research must be limited to exposures producing nothing worse than short-term health effects that are reversible. The number of human subjects is generally very small in these studies because of the costs of the research. Some susceptible individuals would never be considered appropriate human subjects for clinical studies because of the potential for harm.

An interesting issue in controlled clinical studies is the extent to which the human subjects are homogeneous with respect to housing and other aspects of environmental quality. A search of the literature suggests that the residential locations of participants in controlled clinical studies are rarely considered explicitly as part of the study sample design. As noted in the Introduction, a random selection of all people will not be a random selection of all places unless the population from which the sample is drawn is uniformly distributed (Goodchild, 1984). Subjects in controlled studies are likely to have a far narrower range of susceptibilities than represented in populations of epidemiological studies, and it is possible that the results of epidemiological studies are driven by populations not included in human clinical studies (Brown et al., 2007). If there are important geographical differences in susceptibilities and exposure, these could be explicitly considered in a spatially stratified sampling scheme for participants in controlled clinical studies.

The method of hazard identification where GIS can make the strongest contribution is environmental epidemiology. ***Environmental epidemiology*** research attempts to associate adverse health outcomes with environmental exposures. Epidemiological investigations are designed to find out whether or not a statistically significant adverse health outcome is observed in an exposed group. The main advantage of these studies is that they measure health effects in people based on actual exposure conditions. Epidemiological studies are particularly useful in situations where exposure concentrations are relatively high during the time period of investigation (e.g., exposure to benzene in workplaces) or when exposed populations are very large (e.g., large urban populations exposed to air pollution), providing the sample sizes necessary to detect small increases in disease incidence with exposure.

There are also limitations to epidemiological studies for risk assessment. Frequently, high-quality hazard information may not be available to assess exposure, as in the case of indoor air quality or water quality at the tap. Also, effects in worker populations may be unsuitable for estimating health effects in the total population because occupational exposures generally involve smaller populations at a limited number of sources and higher doses. Finally, from a public health perspective, it is most desirable to identify hazards *before* exposure has produced adverse human health effects. Epidemiological investigations, relying as they do on the lagged association between adverse outcomes and exposures, are not protective of human health because the adverse impacts have already been manifested.

Among the many research, modeling, and evaluation approaches used to make links between toxicants and human disease, accountability studies and environmental health tracking are emerging as complementary approaches to risk assessment and environmental epidemiology. Accountability studies, which are discussed later in this chapter in the section on GIS and environmental risk management, focus on the level of health benefits that might accrue from actions to reduce toxicants. ***Environmental public health tracking (EPHT)*** "is the ongoing systematic collection, integration, analysis, interpretation and dissemination of data about environmental hazards, exposure to environmental hazards, and health effects potentially related to exposure to environmental hazards" (McKone, Ryan, & Özkaynak, 2009, p. 32). In 2002, Congress provided the Centers for Disease Control and Prevention with funding to develop an environmental public health tracking program in the United States (Centers for Disease Control and Prevention, 2006).

Environmental public health tracking involves locating sources of potential toxicants, analyzing the transport and ultimate fate of toxicants released from these sources, collecting and analyzing samples that offer a measure of the environmental quality at various locations from air, water, and soil, describing the locations and demographic characteristics of exposed and susceptible populations, and modeling the conditions of human exposure. This approach often relies on ***exposure indicators*** such as proximity to roads as surrogate measures of exposure. Temporal and spatial variations in pollutants and in human travel and activity patterns are also emphasized in the environmental public health tracking approach, and GIS are proving to be useful in many of these analyses.

GIS Analysis of Source Locations of Environmental Hazards

One way of classifying hazard sources is based on the geography of the discharge process. ***Point source*** pollution occurs when contaminants are discharged into the environment at a single discharge point, for example, a smokestack or a sewer (Puckett, 1994). Many of the air and water pollution control measures adopted in the early 1970s were directed at point sources because of the volume of pollutants they discharged and the relative ease of identifying them in the landscape. These sources are often modeled as point features in GIS applications.

More recently, pollution control efforts have broadened to include nonpoint sources. ***Nonpoint sources*** contribute pollutants to the air, water, or soil at numerous and widespread locations rather than at a few localized discharge points. Motor vehicle emissions are a nonpoint source of air pollution. Commercial fertilizer and animal manure are important nonpoint sources of nitrogen, which affects water quality. Nonpoint sources are often modeled as line or area features within a GIS.

Some challenges in using GIS to represent the distribution of point and nonpoint pollution sources include quality of positional information, completeness of data, and acquisition of data from different regulatory agencies or other data sources. These factors can make compilation of source data difficult.

Pollutant Release and Transfer Register Databases

Pollutant release and transfer register (PRTR) programs adopted in more than 20 countries provide some of the most comprehensive data on pollution releases to air, water, and land including off-site transfers for waste management (Table 6.1). These systems share several basic characteristics:

- Standardized data for specific facilities at identified locations.
- Standardized data for specific chemicals with data on releases to air, water, and land and data on transfers for each chemical.
- Periodic reporting, preferably mandatory reporting on an annual basis.
- Public access to data, with online access through websites that incorporate GIS functions.

The Toxics Release Inventory in the United States features many of these characteristics.

THE TOXICS RELEASE INVENTORY

One of the most important sources of information on environmental release of toxic substances from point sources in the United States sources is the ***Toxics Release Inventory*** (TRI) developed by the U.S. Environmental Protection Agency (EPA) in 1986 as part of the Superfund reauthorization. Approximately 30 states and cities had already enacted some form of pollution disclosure law by this time (Hearne, 1996). After Congress passed Section 313 of the Emergency Planning and Community Right-to-Know (EPCRA) law, U.S. manufacturers were required to report to the EPA on an annual basis the amounts of toxic

TABLE 6.1. Pollution Release and Transfer Registers

Agency	Register	Online links
U.S. Environmental Protection Agency	Toxics Release Inventory	*www.epa.gov/tri*
Environment Canada	National Pollutant Release Inventory	*www.ec.gc.ca/pdb*
México SEMARNAT	Registro de Emisiones y Transferencia de Contaminantes	*www.semarnat.gob.mx/Pages/Inicio.aspx*
Australian Department of the Environment, Water, Heritage and the Arts	National Pollutant Inventory	*www.npi.gov.au*
European Commission	European Pollutant Release and Transfer Register	*prtr.ec.europa.eu*

chemicals they release into the environment or ship off-site as waste (Doa, 1992). The EPA required manufacturers to submit complete TRI data forms for *each* chemical covered by the TRI if they meet the requirements. For example, a facility required to report on three TRI chemicals would submit three separate forms. An important component of the law was creation of an unrestricted online reporting system. The first year for which TRI data are available is calendar year 1987.

The Pollution Prevention Act of 1990 broadened the TRI to include reports on source reduction, recycling, and treatment. In 2000, more stringent reporting thresholds were established for ***persistent bioaccumulative toxic chemicals*** (PBT). These chemicals, including dioxin, lead, mercury, compounds related to these three, polychlorinated biphenyls (PCB), and some pesticides, are of special concern because they are toxic and not easily mitigated, they remain in the environment for long periods of time, and they accumulate in body tissue.

The law initially covered all manufacturing facilities in all U.S. states and jurisdictions employing the equivalent of 10 full-time employees in industries in ***Standard Industrial Classification*** (SIC) codes 20 through 39 that produced, imported, or processed 25,000 pounds or more of any of the 600 individual chemicals and 28 chemical categories on the TRI list of toxic chemicals or that used in any other manner 10,000 pounds or more of a TRI chemical during the reporting year (U.S. Environmental Protection Agency, 2010a). In 2005, the EPA proposed changing reporting requirements. These proposals were controversial because they allowed some facilities to file a brief certification form (Form A) instead of a detailed reporting form (Form R), allowed some PBT chemicals to be reported using Form A, and proposed changing the frequency of reporting from yearly to every other year (Bazilchuk, 2006). The final rule, which was promulgated in 2006, changed the reporting requirements, but it did not change the reporting frequency (U.S. Environmental Protection Agency, 2006). The 2006 TRI data, the first to be released under the new regulations, showed a 13% increase in Form A reports and a 2% decrease in releases from the levels reported in 2005 (U.S. Environmental Protection Agency, 2008a). Three percent fewer facilities reported in 2006 than in 2005.

In 2006, the EPA adopted the ***North American Industrial Classification System*** (NAICS) codes for identifying industries required to report (U.S. Environmental Protection Agency, 2009a). NAICS has replaced the SIC system for classifying industries. It was jointly developed by the U.S. Economic Classification Policy Committee, Statistics Canada, and Mexico's Instituto Nacional de Estadística, Geografía e Informática, to promote comparability in business statistics among the North American countries participating in the North American Free Trade Agreement (NAFTA). NAICS was introduced in 1997 and is revised every 5 years. Based on these revisions, the EPA updates its list of NAICS codes identifying industries required to report.

An important category of reporting information for the purposes of GIS analysis is facility information. This category includes lon/lat coordinates that identify where the release occurred. In 2005, the EPA modified TRI reporting forms so that facilities were no longer required to report lon/lat and certain facility identifier codes assigned by regulators. These data are now managed in a separate database, the ***Facility Registry System*** (FRS) (U.S. Environmental Protection Agency, 2009b). This system was developed to address problems with accuracy and consistency of facility and location data across state and federal regulatory programs recognized by analysts of the TRI data published in the early years of the program (Burke, 1993; Scott, Cutter, Menzel, Minhe, & Wagner, 1997). It maintains data on more than 1.5 million unique facilities. Data can be accessed online in cartographic and tabular formats. TRI facilities and release data and data on Superfund National Priority List sites can also be mapped through the National Library of Medicine's TOXMAP® site (National Library of Medicine, 2010).

The geographic information on TRI facilities has provided a basis for describing geographic patterns of TRI facilities and releases (Stockwell et al., 1993). Considering the use of TRI data in an environmental hazard surveillance system, however, it is worth pointing out that TRI reporting requirements cover only a selected set of industries that release toxicants in their current operations. TRI data cannot be regarded as providing a complete spatial and temporal picture of environmental contamination.

OTHER POINT SOURCE DATABASES

To create a comprehensive picture of sources of environmental contamination in a particular region, public health analysts have to draw on multiple sources of information. Many state environmental protection agencies have compiled databases of point sources of environmental contamination. In Connecticut, for example, the Department of Environmental Protection maintains a Point Source Inventory. The Point Source Inventory includes all sources in the state capable of emitting more than 5 tons per year of any one of a specified set of pollutants including carbon monoxide (CO), volatile organic compounds, lead, and particulates. Most of the sources are combustion sources, but these are sometimes operated by facilities like hospitals or schools that public health analysts might not readily think of as sources of environmental contaminants. Health facilities have also been identified as point sources discharging chemicals into water by their waste disposal practices for unused pharmaceuticals (U.S. Environmental Protection Agency, 2008b).

Connecticut's Point Source Inventory includes state plane coordinate locations for each stack and a base elevation for the stack. This feature makes it possible to integrate data from the Point Source Inventory with other GIS databases maintained by the Connecticut Department of Environmental Protection. Mapping the locations of point sources reveals how widely distributed they are across the state (Figure 6.2). State environmental regulations and permitting statutes can provide important information on the kinds of sources that are regulated in a state and that might be identified in a database compiled by a state regulatory agency.

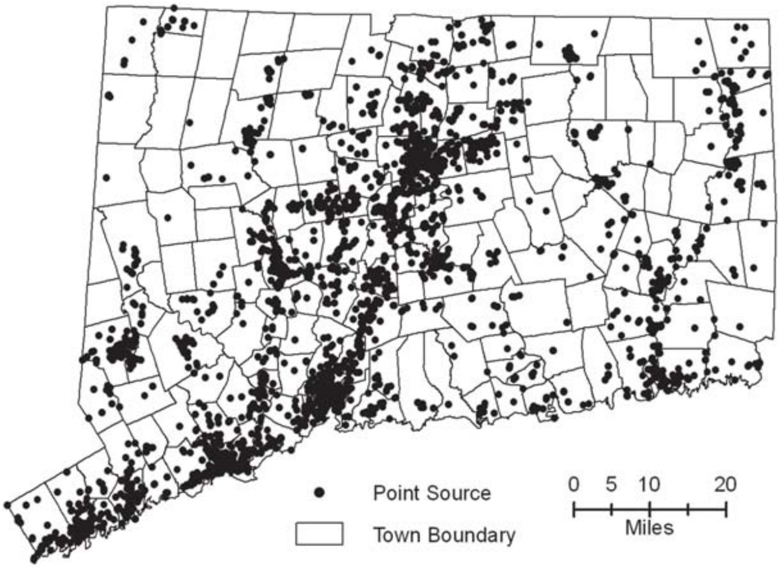


FIGURE 6.2. A map of point sources for air pollution in Connecticut. Data provided by the Connecticut Department of Environmental Protection.

USING GIS TO DEVELOP DATABASES OF POTENTIAL POINT SOURCES

In some cases, regulatory agencies find it difficult to identify all of the entities that may release toxicants and are subject to some kind of permitting or enforcement action. In these cases, GIS analysis has been used as a screening tool to identify the locations of industries that might be discharging wastes. An analysis taking this approach was conducted in seven counties in Pennsylvania to identify industries in the region likely to have shallow injection wells based on their SIC codes from the Dun and Bradstreet Baseline database (Davis & Flores, 1992). A GIS database of the locations of 14 types of industrial facilities was created based on the lon/lat reported in the Dun and Bradstreet database. This database, like the TRI, contains a variable describing how the coordinates of the facilities were determined. A GIS database of principal sewer systems was also created from maps of sewered and unsewered areas provided by local officials in the counties.

The analysts were interested in identifying facilities in unsewered areas because these facilities would likely be using injection wells to dispose of their liquid wastes. The cartographic overlay function of the GIS was used to overlay the two databases. "In this way, the facilities outside sewered areas were easily identified" (Davis & Flores, 1992, p. 117). Using the GIS functions that enable selection of objects by feature attribute, companies that had yearly sales figures of more than $1,000,000 were selected and mapped separately because it was assumed that they generated higher levels of waste. A second group of companies having 20 or more employees was also selected because their septic systems required permitting.

A similar approach was adopted in a study modeling the possible impacts of emissions from facilities too small to be required to report to regulatory agencies (Dolinoy & Miranda, 2004). TRI data for Durham County, North Carolina, for 2000 when SIC codes were still being used were downloaded from the EPA website. Data for facilities that were not required to report to the TRI but had TRI-reporting SIC codes were identified from a commercial city marketing database for the same year. A tax parcel database was used to geocode the locations of the nonreporting facilities using the addresses of facilities listed in the marketing directory. The TRI data yielded 16 facilities, but more than 400 additional facilities were identified from the marketing directory. In addition to increasing the number of facilities, incorporating data from the commercial directory significantly changed the geographical distribution of facilities. Only three of the TRI facilities were located in central Durham, an area of low-income minority population, but most of the small nonreporting facilities were concentrated in this area (Figure 6.3).

Computer models have been developed to estimate quantities of hazardous wastes generated at industrial facilities based on their number of employees and the products they manufacture (Ashact, Ltd. & Dagh Watson, Spa., 1989; Dolinoy & Miranda, 2004). These data can be used with the locations of any facilities captured from directories, remote sensing data, or field surveys. This approach is potentially useful in environmental health analyses conducted in situations where facilities are not required to disclose hazardous material inventories (Lowry, Miller, & Hepner, 1995).

Nonpoint Source Data and Modeling

Data on nonpoint pollution sources affecting air and water are challenging to assemble because the sources are so numerous and widespread. Air pollution comprises gases and particulates, and both types of pollution have been studied using GIS. Vehicles are a major source of emissions affecting both human and environmental health. The transportation sector accounted for one quarter of the increase in greenhouse gas emissions in developed countries between 1990 and 2004; this share is likely to grow over the next 30 years, with most of the growth occurring in developing countries (Walsh, 2008).

To quantify emissions from on-road vehicles in Istanbul, Turkey, devices were used to collect emissions data from the exhaust of different types of vehicles operated in real-world conditions (Gumusay, Unal, & Aydin, 2008). Road grade and coordinates for the position of the vehicle were captured using GPS receivers attached to each vehicle. The location data from the GPS were used to create a point database in a GIS application, and the emission data, measured every second, were integrated with the corresponding GPS locations.

Agricultural runoff including pesticides is a major source of water pollution. California's pesticide regulatory program is recognized as a model program (Cal-

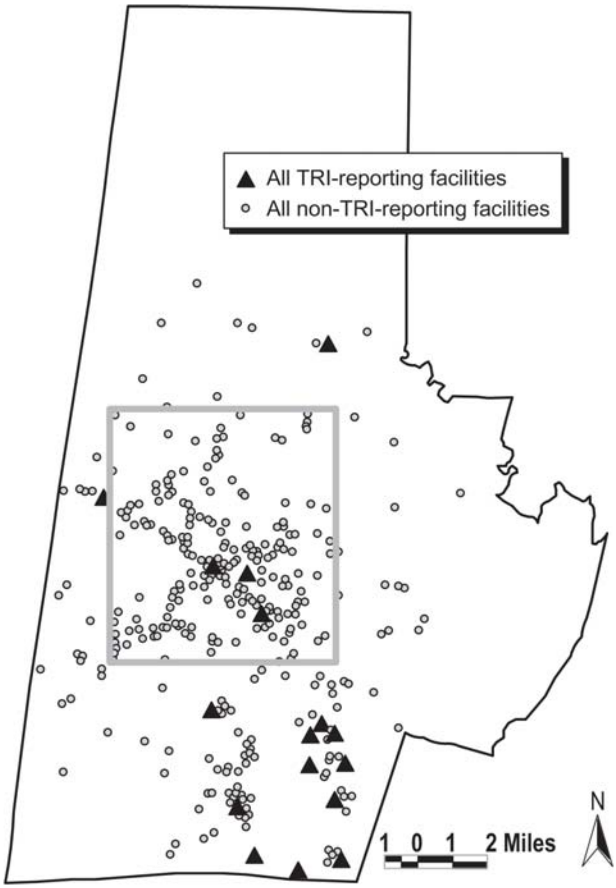


FIGURE 6.3. The locations of all TRI and non-TRI reporting facilities in Durham County, North Carolina, in all SIC codes required to report. The gray box highlights Central Durham County. Including non-TRI reporting facilities significantly changes the number and distribution of facilities. The number of facilities increases, as does the concentration in the center of the county where low-income and minority populations live. From Dolinoy and Miranda (2004).

ifornia Department of Pesticide Regulation, 2000), and data from the Pesticide Use Reporting (PUR) program have been used in many environmental health studies. Full use reporting regulations were introduced in 1990, but limited-use reporting requirements have been enforced since 1950. Home and garden use and most industrial and institutional uses are not included in the reporting, but uses in parks, golf courses, cemeteries, and transportation rights-of-way are included.

Pesticide applicators are required to report month and year of application, date and time of application, acres or units treated, amount of product applied, and a range of geographic identifiers. These include the county and Public Land Survey System section, township, and range, base and longitude, and a site ID. California is one of a number of states using the Public Land Survey System described in Chapter 2. A site identifier, assigned by county agricultural commissioners to a physical plot of land, is also reported.

When these types of data are not available, GIS have also been useful in modeling nonpoint pollution sources. Underground, septic system discharge in soils unsuitable for waste purification has been identified as an important source of groundwater pollution (Bicki & Brown, 1991). In Pennsylvania, where soil wetness, shallow bedrock, slow percolation, and steep slopes limit septic system performance, a GIS analysis was performed to model nitrogen loadings from septic systems on a statewide basis (Nizeyimana et al., 1996). The 1990 Census of Population and Housing, however, reports the number of housing units on septic tanks or cesspools by census tract. These data were used with data on the number of persons and housing units to estimate the amount of nitrogen produced in each tract. The census tract data were integrated with a database of watersheds in the state. This approach was taken as an alternative to pinpointing the locations of more than one million septic systems on a statewide basis in Pennsylvania (Figure 6.4). In smaller areas or watersheds, however, analysts have used

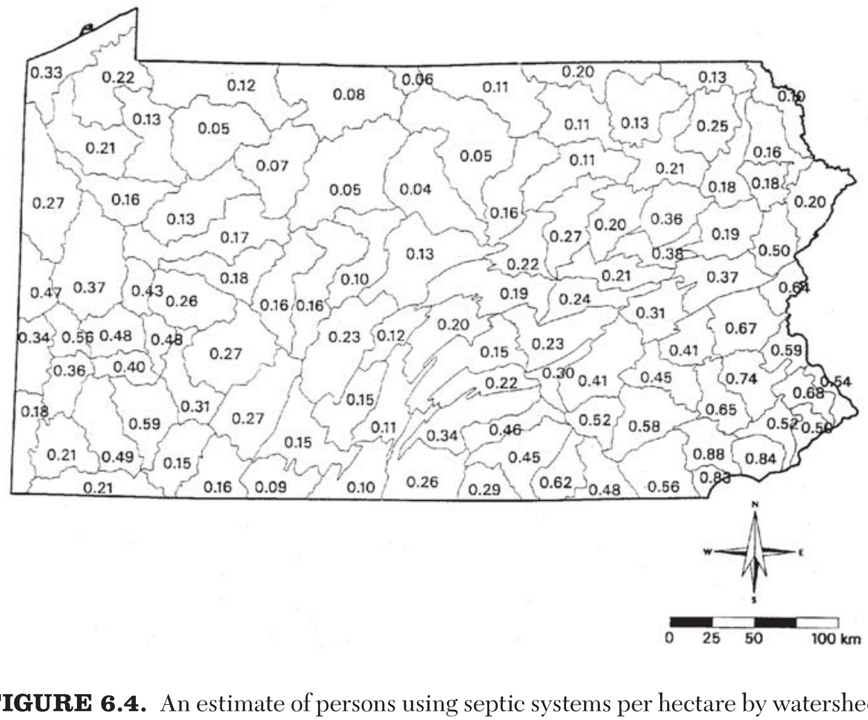


FIGURE 6.4. An estimate of persons using septic systems per hectare by watershed area in Pennsylvania. From Nizeyimana et al. (1996). Used with permission from *Journal of Environmental Quality* (1996).

parcel data or other information to geocode the locations of septic systems using GIS (Stark, Nuckols, & Rada, 1999; Delaware Department of Natural Resources and Environmental Control, 2004).

The Changing Geography of Hazards

Hazard geography is not static. The locations of industrial and agricultural activity change over time, and past patterns of land use may not be evident in the current landscape. Modern industrial and agricultural activities involve the assembly of raw materials from many sources to production sites as well as the distribution of finished products and waste materials to other locations. Accidental release of contaminants may occur at a variety of places as materials are transported (Bowen et al., 2000). GIS have been used to document historical patterns of contamination and to address transportation of hazardous materials.

HISTORICAL PATTERNS OF ENVIRONMENTAL CONTAMINATION

Local land use and zoning controls provide some basis for evaluating potential impacts from new activities making use of toxic substances or generating them during the production process. It may be more difficult to detect historical patterns of contamination, particularly when land use change has occurred. Databases like the TRI and the California PUR have been in existence for a sufficient number of years that it is possible to track changing patterns of environmental hazards over time. It is also possible to reconstruct past community landscapes through other sources. Old telephone directories, Sanborn fire insurance maps (Geography and Map Division, Library of Congress, 1981; Keister, 1993), reports and case studies of contamination events (Colten, 1991), and other archival materials can be used to identify the locations of tanneries, paint manufacturers, metal processors, and other businesses that might have polluted the environment or even the geographical patterns of past contamination events.

A historical geographic information system (HGIS) application was developed for an environmental history project covering the period from 1855 to 2005 in London, Ontario (Gilliland & Novak, 2006). Scanned fire insurance plans provided one data series used to identify industrial sites and their proximity to rivers, parks, and open spaces. These tools can enable analysts to document important changes in urban morphology. Analysis of past activities is also important in less urbanized areas, as research on breast cancer risk and historical exposure to pesticides among women living on Cape Cod, Massachusetts, illustrates (Brody et al., 2004).

Analysts using GIS to study conversion of industrial sites to other uses in New Orleans, Louisiana, found that most sites occupied by polluting activities had been converted to other uses (Frickel & Elliott, 2008). The study covered the period from 1955 to 2006. The study used the parcel as the basic spatial unit of analysis and tracked changes in the uses of the parcels over time. The sites where conversion had occurred were not concentrated in minority neighborhoods. A similar pattern was found in a study of contaminated sites in Baltimore showing concentrations of sites in white working-class neighborhoods with excess mortality rates (Litt & Burke, 2002).

Public health professionals have grappled with the issue of converting ***brownfields***, defined as chemically, physically, or biologically contaminated abandoned or underutilized commercial or industrial properties, for residential use (Greenberg, 2002; Greenberg, 2003) even when neighborhood residents are not opposed to the development of housing (Greenberg & Lewis, 2000). An assessment of European brownfields evaluated differences in definitions of brownfields across countries and regions (Oliver, Ferber, Grimski, Millar, & Nathanail, 2005). In countries with relatively low population densities, dealing with contamination was the primary issue of concern. In western European countries with high population densities, there was more pressure to convert brownfields to residential use. The identification and redevelopment of brownfields is also a significant issue in developing countries (Moore, Gould, & Keary, 2003).

HAZARDOUS MATERIALS TRANSPORTATION AND DISASTER RELEASES

The Hazardous Materials Transportation Uniform Safety Act of 1990 provides a legislative basis for regulating shipments of hazardous materials in the United States. To track transportation-related releases of hazardous materials, federal regulations require reporting to the Pipeline and Hazardous Materials Safety Administration (PHMSA) (U.S. Department of Transportation, Pipeline and Hazardous Materials Safety Administration, 2011). Notice involving infectious substances—etiologic agents—may be given to the director of the Centers for Disease Control and Prevention. The database of Incident Reports can be searched and the search results downloaded in comma-delimited textfile format. Geographic identifiers include city, county, state, and postal code of the incident, and route information with address or intersection data that could be used to geocode the locations of incidents associated with highway transportation. Incidents related to other transportation modes including pipeline, air, rail, and water are also included.

Releases of toxicants can be caused by natural disasters such as forest fires, floods, and earthquakes (Young, Balluz, & Malilay, 2004). Some of these releases are direct, resulting from the event itself. Dioxin produced by forest fires is an example of a ***direct release*** that cannot be prevented. An ***indirect release*** resulting from a disaster may be intentional or unintentional. Pesticide spraying to control insects following flooding is an example of an indirect release intended to prevent a health threat, like vector-borne disease, that is considered more serious than the threat from the pesticide. ***Na-tech event*** (natural-technologic event) releases are also indirect releases, but they are considered unintentional. Na-tech events are nevertheless seen as preventable sources of environmental contamination because action can be taken to improve the technologies used to store and ship toxicants to prevent unintended releases.

GIS and related technologies played a role in assessing the potential impact of flooding from heavy rainfall on discharges from waste pits used by confined animal feeding operations in North Carolina (Wing, Freedman, & Band, 2002). If the waste pits are breached or flooded, both chemical wastes and pathogens can be released. The lon/lat coordinates of more than 2,200 operations in the eastern part of the state were checked and corrected to represent the feeding operations as point features. The state Division of Water Quality provided information on which waste pits were breached or flooded from 15 to 20 inches of rainfall caused by Hurricane Floyd in September 1999. The state's Division of Emergency Management provided data on inundated areas derived from satellite imagery approximately one week after the hurricane. More than 10% of feeding operations modeled as points were located in inundated areas. Although not all of these experienced breaches or flooding as a result of the rainfall, many are at risk for off-site discharge of waste from flooding. This finding has implications for the management of hazards because feeding operations are permitted as non-discharge facilities that retain waste on site. The interaction of natural events, production processes, and regulation can lead to complex environmental impacts that, in turn, affect human health.

Integrating Databases Describing Sources of Contaminants

GIS applications have used pollutant release and transfer data in conjunction with other source data to describe the sources of environmental contaminants more completely. As part of a project to monitor environmental conditions within the Greenpoint/Williamsburg section of Brooklyn, New York, data on the locations of approximately 20 TRI sites were integrated with data on the locations of other potential sources of environmental contaminants (Osleeb & Kahn, 1999). These included a sewage treatment plant and incinerator, a low-level radioactive waste repository, more than 200 hazardous-materials processors, a major expressway, and a large number of chemical and petroleum bulk storage tanks.

As these applications show, GIS have been used to describe the sources of environmental contaminants. Concerns remain about the accuracy of self-reported data from pollution release and transfer registers, and new technological processes like nanotechnology may not yet be regulated (de Marchi & Hamilton, 2006; Wardak, Gorman, Swami, & Rejeski, 2007). Even when pollution source databases are relatively complete, the location of pollution sources only partially describes the geography of environmental hazards. Contamination zones around these sources also need to be evaluated.

Modeling Fate and Transport and Environmental Quality in a GIS

Fate and Transport Modeling

After the locations of sources of toxicants have been identified and analyzed, it is necessary to understand how these agents affect environmental quality *in situ* and elsewhere. ***Fate and transport models*** are used to investigate what happens to agents that are released into the environment (Dunnivant & Anders, 2006). These models require geographic and physical descriptions of the source and information on the rate of release into the atmosphere, the hydrosphere (surface water and groundwater), and the lithosphere (land). Given the location of the source, the meteorology of the receiving air (including rainfall patterns), the hydrology and hydrogeology of the receiving water features, and the physical and geologic characteristics of the receiving land can be evaluated.

GIS have often been used for pre- and postprocessing of data used in fate and transport models (Pistocchi, 2008). Applications taking this approach to describing the areas ultimately impacted by discharges into the air, water, or soil generally consist of two major components: a chemical dispersion model and a GIS database (Chakraborty & Armstrong, 1995). Given information about the types of chemical released and local environmental conditions (Figure 6.5), the chemical dispersion model provides the dispersal "footprint." The plume footprint can then be incorporated into a GIS database by locating the source of the release as the origin of the footprint and computing the planar coordinates of the footprint polygon. By overlaying the plume footprint with census data, the characteristics of the population within the footprint area of risk or exposure can be modeled.

A variety of dispersion models are available. Some of these are relatively simple, while others are complex three-dimensional models capturing vertical as

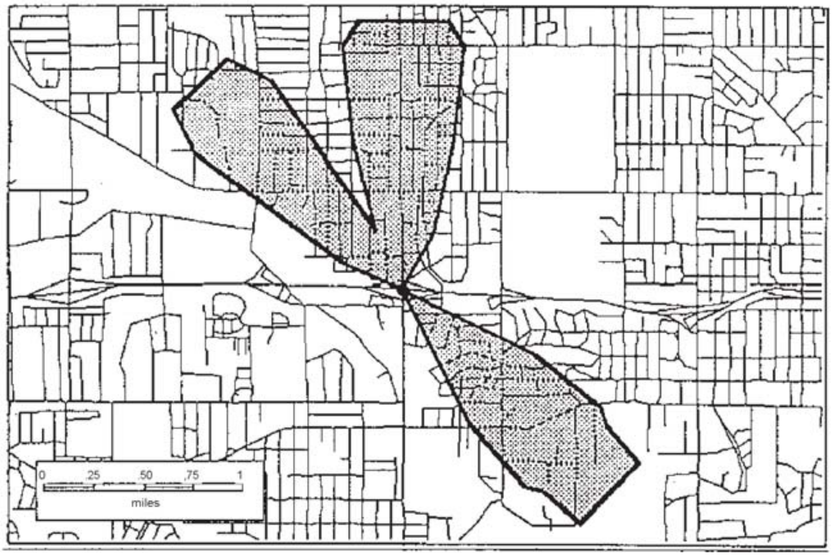


FIGURE 6.5. A composite of 12 monthly dispersion footprints generated for the same accident location in Des Moines, Iowa, reflects seasonal variations in prevailing wind direction. From Chakraborty and Armstrong (1995). Copyright (1995), with permission from Elsevier.

well as horizontal spread (Brutsaert, 1982). Composite plume models are developed from a set of dispersion footprints generated for the same source but incorporating different air temperatures, relative humidities, cloud covers, and wind speeds and directions, reflecting variability in climatic conditions. If the locations of releases are known, composite plume models can be developed around each origin. By identifying the locations of 45 intersections in Des Moines, Iowa, with the highest numbers of truck accidents and then developing composite plume models for each intersection based on long-term average monthly climate data, Chakraborty and Armstrong (1995) were able to identify residential populations most at risk for exposure to gases released following a collision.

Fate and transport modeling is also an important technique for investigating degradation of water quality. A number of models have been developed to study watershed hydrologic processes and nonpoint source pollution (Borah & Bera, 2003). Some of these are continuous, and others are designed to measure the impacts of single rainfall events like major storms. Most of these models were developed in the 1970s and 1980s before the widespread availability of GIS, but subsequent research has integrated many of them with GIS.

Solute transport models incorporating GIS for spatial data compilation, analysis, and visualization have been developed at a variety of spatial scales from the individual farm to the multistate region (Wagenet & Hutson, 1996). In a study of a relatively small geographic area (7 kilometers × 10 kilometers in upstate New York), a GIS was used to overlay slope, land use, and soils databases to produce a composite identifying those soils found in agricultural areas with slopes less than 10%. The hydraulic properties of the soils were used to identify main hydraulic groups. Pesticide data for four chemicals along with pesticide application rates and typical corn planting dates and growth patterns were obtained for the region. In this application, the GIS was particularly useful for preprocessing input data for the solute transport model from a variety of sources and for postprocessing the results of the various simulations and generating maps.

The ***Soil and Water Assessment Tool (SWAT),*** was used to investigate two subwatersheds in the Sandusky Watershed region in Ohio (Grunwald & Qi, 2006). This watershed is part of the Great Lakes Basin, and surface runoff of suspended sediment and agricultural chemicals is a major source of nonpoint source pollution affecting water quality in Lake Erie. The spatially distributed modeling approach made it possible to produce maps showing the geographical patterns of simulated and independently validated suspended sediments and nutrients across the watershed. The simulated data were not equally valid throughout the study area.

Assessments of environmental conditions resulting from fate and transport modeling can be confirmed or contradicted by field measurements. Comparative studies of fate and transport for different chemicals across different study sites are providing evidence that local conditions affect transport. A study of phosphorus, nitrogen, and pesticide transport to streams in five agricultural basins in the United States showed that climatic, hydrological, and agricultural management practices were geographically variable and affected pesticide loads as a percent of actual use (Domagalski et al., 2008). GIS data and tools are increasingly used to develop spatial multimedia fate models (Pennington, Margni, Ammann, & Jolliett, 2005). In addition to describing the sources of environmental contaminants and their fields of impact, measuring environmental quality is an important component of environmental health analysis.

Environmental Quality

The cumulative effects of physical processes and human activities in the environment manifest themselves in environmental quality. While the sources of environmental contaminants and their dispersion patterns have generally been modeled as object data (points, lines, and areas), environmental quality is a field variable, as defined in Chapter 2. Dimensions of environmental quality—atmospheric conditions, water quality, and soils—are continuous and can be observed everywhere. These phenomena are usually measured based on a partitioning of the surface that creates a spatial framework for locating monitoring stations. Sometimes, the observed measures of environmental quality are mapped only for the monitoring locations. As part of the CDC's National Environmental Public Health Tracking Program, U.S. Geological Survey and National Water-Quality Assessment (NAWQA) Program data on domestic well water quality was mapped for more than 12,000 domestic wells in private water supply areas across 16 states (Bartholomay, Carter, Qi, Squillace, & Rowe, 2007). Point data for wells were mapped with other GIS data on aquifers, land cover, and population density to aid interpretation of the water quality data in terms of human health concerns. More commonly, however, monitoring data are interpolated to create statistical surfaces of environmental quality. These methods contribute information to studies of the relationships between environmental quality and human health outcomes by modeling environmental quality. Other types of models are used to link environmental conditions to exposure and to human health outcomes, as discussed later in this chapter and in Chapter 11.

In numerical analysis, ***interpolation*** is the process of creating new data observations within a discrete set of known observations. ***Spatial interpolation*** involves analyzing measurements at known locations to estimate values of the measured phenomenon at other locations where no measurements have been taken (Chang, 2009). These points where no values have been measured are generally modeled by creating a grid that covers the region for which the phenomenon is being interpolated and are referred to as ***grid points***. Once the ***control points***, the points with known values, are in hand, the analyst must choose an interpolation method. ***Global interpolation*** methods like trend surface analysis use values from all of the known points to estimate the value at each unknown point. In most health applications, ***local interpolation*** methods are used. These methods use values observed for only a sample of neighboring known points to estimate the value at each unknown point. The sample of points may be determined in a number of ways: by finding the closest known points to the unknown point, by finding known points within a specified distance of the unknown point,

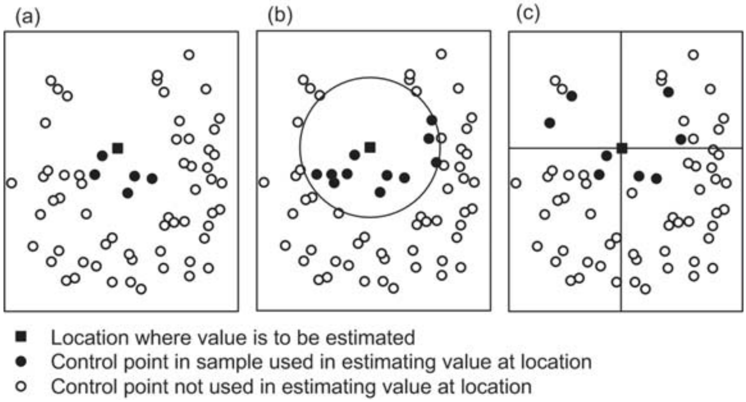


FIGURE 6.6. Search methods for identifying control points for local interpolation. The analyst can select the k closest points, in this case the five closest points (a). Alternatively, a search radius of k units identifies 11 points (b). In both of these approaches, the sample control points lie mostly to the south of the location for which a value is being estimated. The analyst can use a quadrant requirement (c) to select the k closest points within each of four sectors. An octant requirement is similar but would require sample points in each of eight sectors.

or by finding known points within each of four quadrants around the unknown point (Figure 6.6).

Interpolation methods can also be categorized as exact or inexact. ***Exact interpolation*** methods result in the estimated values at known locations that are exactly the same as the known values. ***Inexact interpolation*** methods may result in estimated values for known locations that are not equal to the known values at the locations. Finally, interpolation methods have been classified as deterministic and stochastic. ***Deterministic interpolation*** methods provide no measures of error associated with the estimated values. ***Stochastic interpolation*** methods, however, provide information on errors in estimated values with estimated variances.

Inverse distance weighted (IDW) interpolation is an exact local interpolation method. The IDW function for estimating values is

$$z_j = \left(\sum_{i=1}^{s} z_i \frac{1}{d_{ij}^k} \right) \Bigg/ \sum_{i=1}^{s} \frac{1}{d_{ij}^k}$$

In the equation, z_j is the estimated value at point j, z_i is the known value at point i, s is the number of points with known values used in the estimation, and d_{ij} is the distance between point j and point i. The weights are inversely proportional to the power k of the distance. In this way, the influence of the surrounding known values on the unknown value at a point is described. As k increases, less weight is given to known values at more distant locations in estimating the value at a point. The number of points with known values used in the estimation also affects the degree of local influence.

This method of spatial interpolation was used to model patterns of cesium deposition on Europe after the Chernobyl disaster (De Cort et al., 1998). Analysts sampled the six closest neighbors within a maximum radius. Because the underlying structure of the deposition pattern was not known, analysts estimated the value of k empirically. Data were randomly selected from the complete data set and inverse distance weighted using different k values. Estimated values for points with known values were then compared to the known values in order to assess which exponent best reproduced known values. An exponent value of 2 provided the best fit over a range of samples.

Kriging is a statistical spatial interpolation method. Unlike IDW, kriging considers not only the distances to control points, but also the spatial autocorrelation of measurements among the control points. ***Spatial autocorrelation*** refers to the similarity or association of values over space as measured in statistics like the G_i* statistic discussed in Chapter 5.

To measure spatial autocorrelation, kriging uses the measure of ***semivariance*** (variance divided by 2). If y_i is the measurement at control point i, d_{ij} is the distance between control points i and j, and h denotes a distance among control points, then the semivariance at distance h is

$$\sum_{d_{ij} \le h} (y_i - y_j)^2 / 2n(h)$$

where $n(h)$ is the number of pairs of control points that are distance h apart. The semivariance thus measures the variance of values for control points separated by distance h. A small semivariance indicates that measurements at distance h are similar to each other, whereas a large semivariance indicates a large disparity in measured values. As h varies, different values for the semivariance will be obtained.

The semivariogram is a graph that shows the values of semivariance at different values of h (Figure 6.7). In general, the semivariance increases with increasing h, reflecting spatial dependence in the phenomena being investigated. Nearby control points will tend to have similar values, and thus a small semivariance, whereas distant locations tend to have less similar values and thus a larger variance. The ***nugget*** is the semivariance where distance equals 0. Semivariance at this distance can represent measurement error or variation at a scale not captured by the analysis. The ***range*** is the distance at which the increase in semivariance with increasing distance levels off. The ***sill*** is the semivariance value reached at the range.

Semivariograms may be used alone, to explore spatial autocorrelation in data, or with other methods. The semivariogram range was used to determine the maximum search radius for the six neighboring control points used to estimate cesium by the inverse distance-weighted method in the atlas of cesium deposition on Europe from Chernobyl (De Cort et al., 1998).

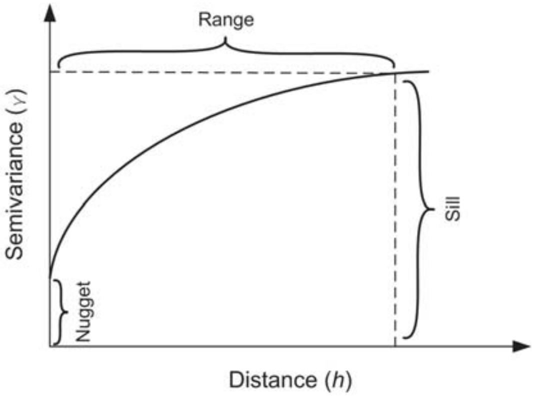


FIGURE 6.7. The semivariogram graphs the relationship between semivariance and distance. The curved line represents the mathematical function fit to the observed semivariance values plotted by distance.

In kriging, the semivariogram is fitted with a mathematical function and used to generate a set of spatial weights or $\lambda_i(s)$ for each control point i and grid point s. Computing these weights involves matrix algebra (Isaaks & Srivastava, 1989). Once the weights have been computed, we estimate the value $y(s)$ at grid point s as a weighted linear combination of the values at various control points (Figure 6.8). The estimated value at grid point s is

$$y(s) = \sum \lambda_i(s) \times y_i$$

Computing these estimated values for all grid points s produces a fine mesh of values that appear as a continuous surface when mapped, as in Figure 6.8.

Kriging is a well-developed spatial interpolation method that is widely used in the earth sciences and geography (Chang, 2009). It has been used for exploring environmental health problems like lead poisoning (Griffith, Doyle, Wheeler, & Johnson, 1998), and to model temporal peaks in the spread of infectious disease, as described in Chapter 7. Although complex, kriging is generally considered the best method for creating a continuous surface map of estimated values from measurements taken at discrete control points. The main advantages of kriging are that:

- Unlike estimated values from other interpolation methods, the estimated values can fall outside the range of the known data values.
- Kriging gives a standard error or kriging variance for the estimated grid points values (Figure 6.9), making it possible to compute confidence intervals around the predictions.
- Kriging incorporates and indeed models the spatial dependence in the data.

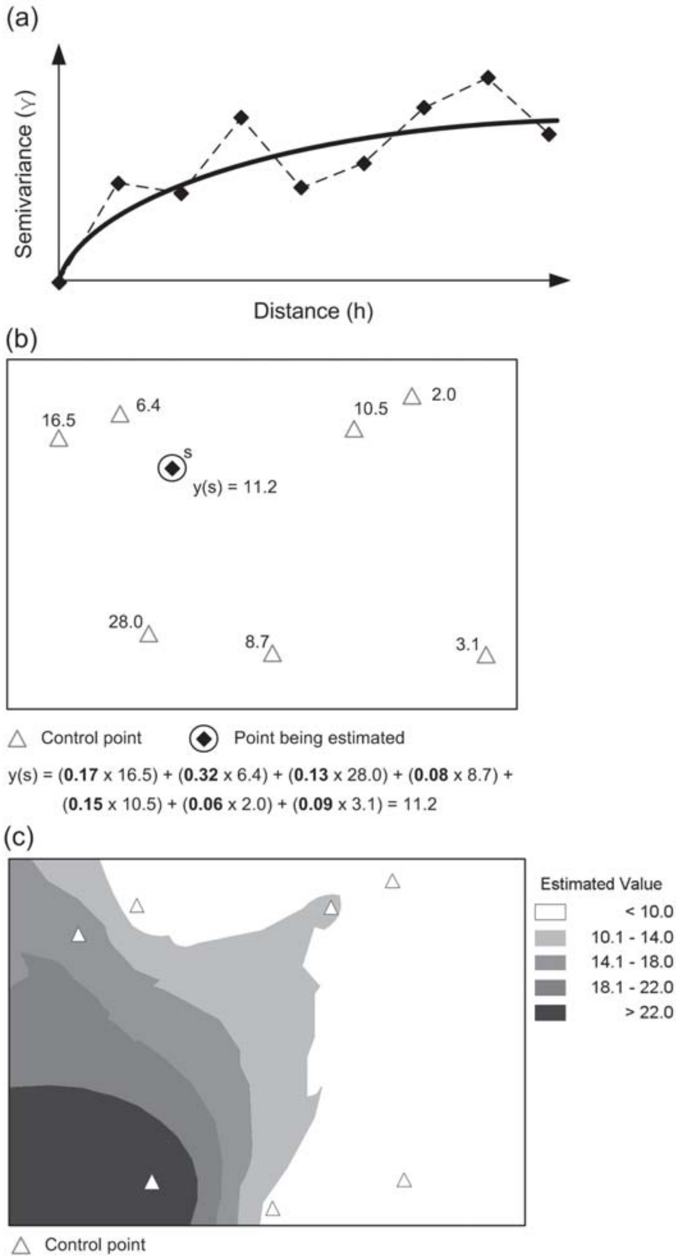


FIGURE 6.8. A schematic example of using kriging to estimate the value at particular places based on known values at control points. The mathematical function is fit to the semivariogram (a). The known values are multiplied by their respective kriging weights (λ), which come from the semivariogram (b). Kriging weights are highlighted in bold. The results of the kriging analysis can be displayed as a continuous map (c).

Despite these advantages, kriging, like all statistical methods, requires caution in its application. The accuracy of kriging estimates depends heavily on the accuracy of the semivariogram that generates them. Before embarking on kriging, it is important to check that the semivariogram is properly specified and fits well the data being modeled.

SAMPLING NETWORKS FOR MEASURING ENVIRONMENTAL QUALITY

An important issue in describing environmental quality, and ultimately its association with patterns of human health and disease, is the design of the sampling network. In a study of outdoor air pollution and asthma in Brooklyn and Queens, New York, there were too few air monitoring stations for measuring inhalable particulate matter and ozone, and the stations were too clustered in the western portion of the study area to interpolate air quality adequately (Weisner, 1994). The absence of adequate monitoring networks also affects environmental quality measurements for catastrophic events (Service, 2003). Sampling network design is also important in studies where watersheds or airsheds span more than one country (Miller et al., 2010).

Designing a monitoring network involves defining the number, locations, sample pattern, and sample frequency of sampling sites (Olea, 1984). Statistical analyses have generally focused on detecting statistically significant variations in environmental quality, that is, evaluating the statistical accuracy of the estimations. Some evidence suggests that more frequent sampling does not necessarily greatly increase the power of the statistical tests or the precision in the estimates (Hsueh & Rajagopal, 1988; Loaiciga, 1989). Redundancy attributable to spatial autocorrelation in environmental quality measurements can also inflate variance estimates (Griffith, 2008). How many sampling sites are necessary and where they should be located to yield an accurate representation of some dimension of environmental quality in a region is both a statistical question and a spatial question, and environmental analysts recommend using spatial–statistical methods in the design of sampling networks (Nelson & Ward, 1981; Beach, 1987; de Gruijter, Brus, Bierkens, & Knotters, 2006).

In terms of where samples should be taken, a "network of good geographical coverage is essential, implying the use of tessellation stratified random sampling" (Griffith, 2008, p. 496). One such design is based on a hexagonal tessellation (see Figure 2.1) or partitioning of the continuous environmental quality surface over the study area and sampling at a random location within each hexagon (Overton & Stehman, 1993). Essentially, then, samples should be taken everywhere. As the foregoing discussion of kriging illustrates, environmental quality values estimated for points where measurements have not been taken have higher error associated with them when there is little local evidence from nearby control points for making the estimation. Unfortunately, resource constraints and other geographical factors limiting access to potential monitoring locations mean that this approach to deciding where samples should be taken is rarely used.

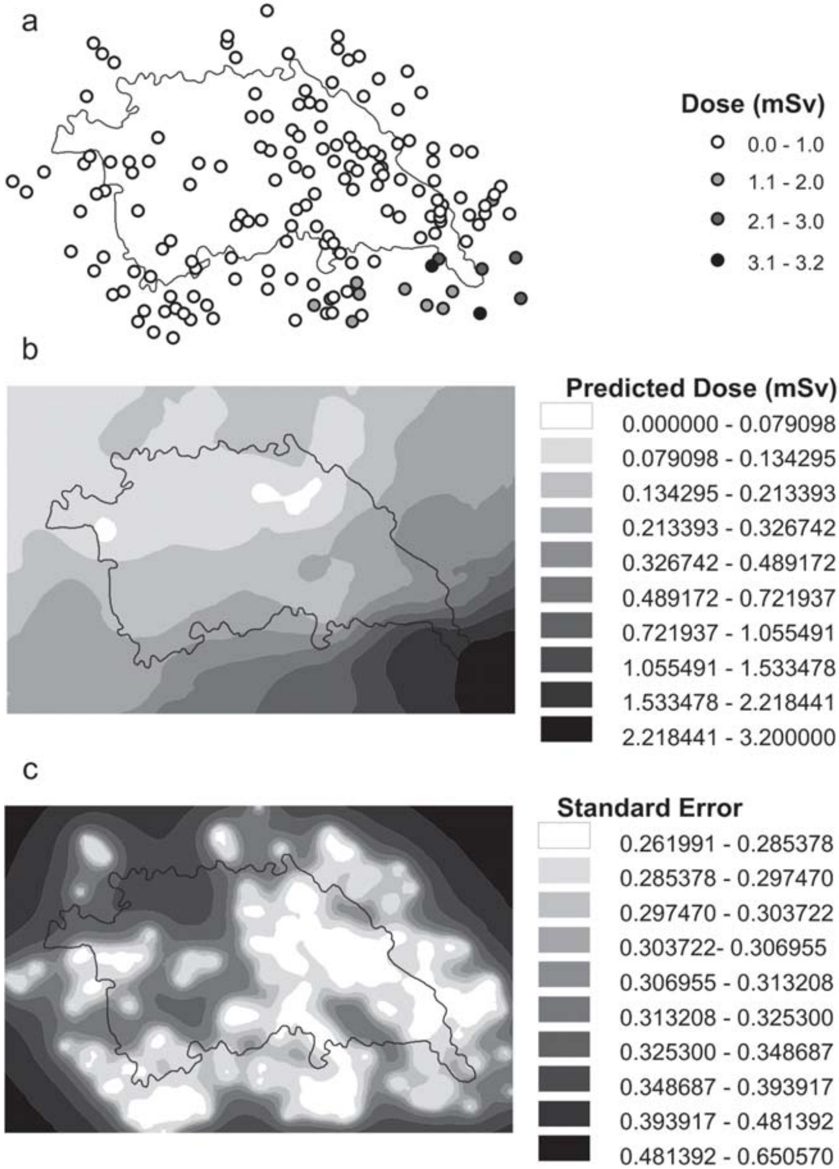


FIGURE 6.9. An analysis of annual effective equivalent dose of radiation, measured in millisieverts (mSv), from external sources in Mozyrskiy Rayon, Gomel Oblast, Belarus. Figure 6.9a shows control points located within and outside of the area with the highest dose levels in the southeast in the direction of the Chernobyl nuclear power plant. Figure 6.9b shows the results of a kriging analysis using ordinary kriging. Figure 6.9c shows the associated standard error map for the predicted values. Error is greatest in the areas with fewer control points spaced farther apart. Data from Serebriakova (2005).

The geographical extent of the tessellation unit is determined by the number of sample points. If the spatial distribution of the phenomenon of interest were known, it would be relatively easy to determine the appropriate number and configuration of control points. For purposes of interpolation, a regular tessellation of sampling points can detect only variations in a surface like environmental quality with "wave lengths of twice the sampling distance or more" (Tobler, 1969, p. 243). But the only way to find out what the spatial patterns of environmental quality are is to sample environmental conditions at different places.

It is possible, nevertheless, to explore the issue of the necessary number of sample points by modeling hypothetical patterns of environmental quality. One such study sought "to demonstrate how sample density affects the spatial and statistical representation of a unitless groundwater quality variable in a hypothetical basin" (Luzzader-Beach, 1995, p. 384) using a variance reduction approach (Rouhani, 1985). The hypothetical basin was designed to approximate in size and dimensions a basin that might be found in northern California and exhibited three groundwater quality patterns of different geographic extent (one large and two small areas). A rectangular grid was overlaid on the hypothetical water quality map, and a model well from which water samples could be taken was located at each intersection in the grid.

Sections of the grid were grouped into hypothetical "townships" to facilitate a spatially stratified random sampling scheme for 22 sample densities ranging from 100% of the hypothetical wells down to 5% of the hypothetical wells. In addition, a sample was drawn at 2.8% coverage to simulate a policy recommended by the California Department of Water Resources Task Force to sample one well per township. For each level of sampling density, the pattern of water quality resulting from the sampled wells was mapped by kriging.

The results of the analysis supported the point that relatively few sample sites are required to accurately represent an environmental feature like ambient groundwater quality in a hypothetical basin. A standard of one well per township, however, was inadequate to capture variation in groundwater quality. The smaller, more localized groundwater patterns were harder to detect as sampling density decreased (Figure 6.10), and a density of five wells per township appeared to be the threshold density for ensuring adequate sampling.

Because the sampling networks for measuring environmental quality are rarely adequate, it is important to investigate how many statistically independent samples have been collected, which is different from determining the number of sampling sites needed. A spatial filter binomial regression model can be used to determine the number of independent samples in the presence of spatial autocorrelation (Griffith, 2008). This model was used to analyze a database of more than 3,500 soil samples taken in Syracuse, New York, and tested for heavy metals. The sample locations were geocoded by GPS. The primary motivation for the research was to determine how many locations were necessary to estimate parameters of a spatial autoregressive model of soil contaminant level as a function of a set of attribute variables rather than accurately modeling environmental quality per se.

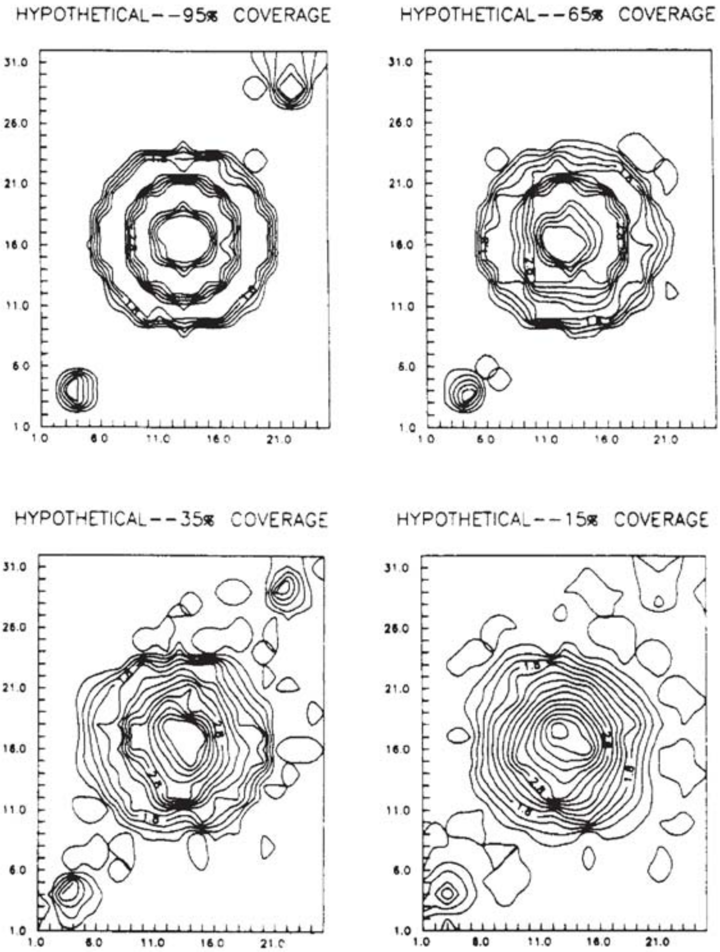


FIGURE 6.10. Contour maps of groundwater quality based on diminishing sample sizes (expressed as the percent of available sample wells in a hypothetical basin) show different patterns of groundwater quality. From Luzzader-Beach (1995). Copyright 1995 by Springer. Reprinted by permission.

INTEGRATED INDEX OF ENVIRONMENTAL QUALITY

In the same way that GIS have been used to develop composite databases of the sources of environmental contaminants (Osleeb & Kahn, 1999), systems have also been developed to integrate data measuring individual components of environmental quality. One integrated index of this type was developed in the Netherlands specifically to support land use zoning that was sensitive to environmental issues (Sol, Lammers, Aiking, De Boer, & Feenstra, 1995). High values of the index at a location would be used to support planning restrictions on housing units, for example, while low values might indicate areas where no restrictions on residential development would be needed. The construction of an integrated index of environmental quality involves five basic steps: (1) identification of polluting agents, (2) assessment of the magnitude of health effects, (3) summation to combine the effects of different agents producing comparable health effects by the same mechanism, (4) valuation of combined health effects to express them as dimensionless units on an arbitrary numerical scale, and (5) aggregation to combine the values associated with the identified agents. Individual and composite index scores can be mapped using a GIS, as in the Netherlands study.

An important advantage of this approach is its integrative aspects. If a regulatory system monitors individual pollutants and compares observed levels to a corresponding regulatory standard, it may be that no individual pollutant exceeds its standard. Use of the integrated index is important in this context because it can be used to identify areas where environmental quality as a whole is unacceptable, even though individual pollutants do not exceed their standards. A major difficulty of this approach, aside from the uncertainty of risk in the assessment process, is the difficulty of finding suitable methods for modeling of human intake and for valuation of health effects.

GIS and Exposure Modeling

Toxicants in the environment can affect a person's health only if the person is exposed. ***Exposure*** may consist of a single occurrence, may be repeated, or may be long term and continuous. The ***dose*** is the quantity of the agent a person is exposed to. The ***effective exposure time*** is the minimum time interval required from exposure to produce a health effect. With exposure to some toxicants, the effect may be almost immediate; with other substances, effects may not be induced for years. The latent period is the time interval from first exposure to observed health effect. The ***latent period*** is a function of many factors, including the dose and dose rate; characteristics of the person like age, sex, and length of time exposed; and the frequency and nature of health observations. ***Threshold toxicants*** are substances that are known or believed to cause adverse health effects above a specified dose or dose rate. ***Nonthreshold toxicants*** are known or believed to cause adverse health effects at any dose. In order for adverse health effects to be detected, studies must follow subjects for longer than the minimum latent period. Studies that do not follow individuals over the lifetime most likely underestimate true risk. Risk to human health is not just a function of the toxicity of the agent; it also depends on the likelihood of people coming into contact with the agent.

The Agency for Toxic Substances and Disease Registry has developed materials for taking an exposure history for use by medical practitioners because many environmental diseases have nonspecific symptoms (Yu, 2008). The form suggests some of the complexities in developing valid exposure histories for a person. In addition to recording basic demographic information such as age and sex, the exposure history form explores the substances someone has come into contact with, but does not attempt to quantify the level of exposure. The form also asks for an occupational history and a history of the residential environment.

The basic demographic characteristics are an indicator that individuals vary in susceptibility to toxic substances. Questions that probe occupational environment, residential environment, and recreational activities remind us that the home location is not the only place where individuals are exposed to toxicants because the home is only one node in a person's activity space, described in the Introduction. Questions detailing employment history highlight the importance of assessing exposures throughout the lifetime.

A study of cancer incidence in the population of a mixed residential and industrial suburb in Denmark was able to account for individual residential mobility during the study period (Poulstrup & Hansen, 2004). The Danish Central Population Register provides data on current addresses of residents as well as historical addresses and dates that individuals moved to or from the addresses. These address data were geocoded to be accurate within several meters. A high degree of residential mobility was observed in the study community. At the end of the 13-year observation period, only half the population lived at the same address and less than one third of the population lived within the study area. Although few countries maintain population registers that track residential mobility at the individual level, researchers can sometimes collect residential data for subjects included in epidemiological studies of adults (Han et al., 2005) or of children (Kohli, Noorlind Brage, & Löfman, 2000). As noted in Chapter 5, GIS modeling of residential histories is being used to detect clusters of health events as well as to model exposure (Jacquez et al., 2006).

Ideally, the environmental health analyst has quantitative data on characteristics of the exposed population, including numbers by age and sex, and on the route and duration of exposure along with the concentration of the contaminant. In most epidemiological investigations, however, exposure data are "problematic" (Hallenbeck, 1993). Often, no exposure data are available for individuals for the locations and time periods of interest or data are of limited validity and reliability due to measurement problems.

Susceptible people live and engage in activities in different geographical contexts. As the previous sections have illustrated, there is also geographical variation in the distribution of pollution sources and environmental quality, which means that conditions in these zones differentially expose people to contaminants. The intersection of these geographies—the ***geography of susceptibility*** and the ***geography of exposure***—has been termed the ***geography of risk*** (Jerrett & Finkelstein, 2005).

Factors that confound or modify the relationship between the toxicant and the health outcome may operate both at the individual and the contextual levels. When susceptible individuals are concentrated in certain areas, the observed health effects in the areas reflect the composition of the population, the ***compositional effect***. When socioeconomic or environmental characteristics of the areas themselves influence health, there is a ***contextual effect***. These concepts are elaborated further in Chapter 11.

In terms of the geography of exposure, actual exposure data at the individual level are rarely available, although there is increasing interest in methods of personal exposure assessment and there have been advances in sensors for monitoring chemicals in the personal ambient environment (Weis et al., 2005). Most exposure assessments involve modeling. A review of models for exposure to air pollution within urban areas identified six modeling approaches for assigning exposure (Jerrett et al., 2005): proximity-based measures, geostatistical interpolation models; land use regression models (Ryan & LeMasters, 2007); line dispersion models; integrated models of emissions and meteorological conditions; and hybrid models combining one of the other methods with exposure monitoring at the individual or household level.

GIS applications in environmental health, emphasizing colocation in time and space of susceptible populations and facilities that release harmful substances, have been open to criticism for apparently substituting proximity to hazardous facilities for quantitative data on actual exposure at the individual level. Systems like Landview® 6, a joint project of the U.S. Bureau of the Census, the U.S. Geological Survey, and the Environmental Protection Agency, enabling users to draw a circle around a user-selected point and generate a demographic and environmental profile of the area within the circle, may be useful for environmental risk management, but they fail to exploit the full capabilities of GIS for more accurate modeling of spatial processes in the hazard–exposure–dose–response sequence. If, for example, we knew the location of a contaminated public drinking water well, we would not want to assess the exposed population based on a circular buffer around the well location. Water from the well is not equally likely to travel in every direction around the well, as the circle implies, because the water is delivered through a distribution network (Aye & Archambault, 1997). GIS can model that distribution network to produce a more accurate representation of the areas served whose water might be contaminated (Figure 6.11).

Cumulative distribution functions have also been used to explore proximity to pollution sources (Zandbergen & Chakraborty, 2006; Chakraborty & Zandbergen, 2007). The cumulative distribution function method is a graphical technique that plots percent of population on the y-axis of the graph and distance to nearest hazard on the x-axis of the graph. A cumulative distribution frequency function shows, for example, the percent of children in a community who live within a specified distance to the nearest TRI facility over the range of all such distances.

An important research area in exposure modeling is the analysis of error. These errors include not only errors in classifying the level of exposure and the health response, but, as highlighted in the discussion of sampling above, spatial autocorrelation in measurement and in measurement error.

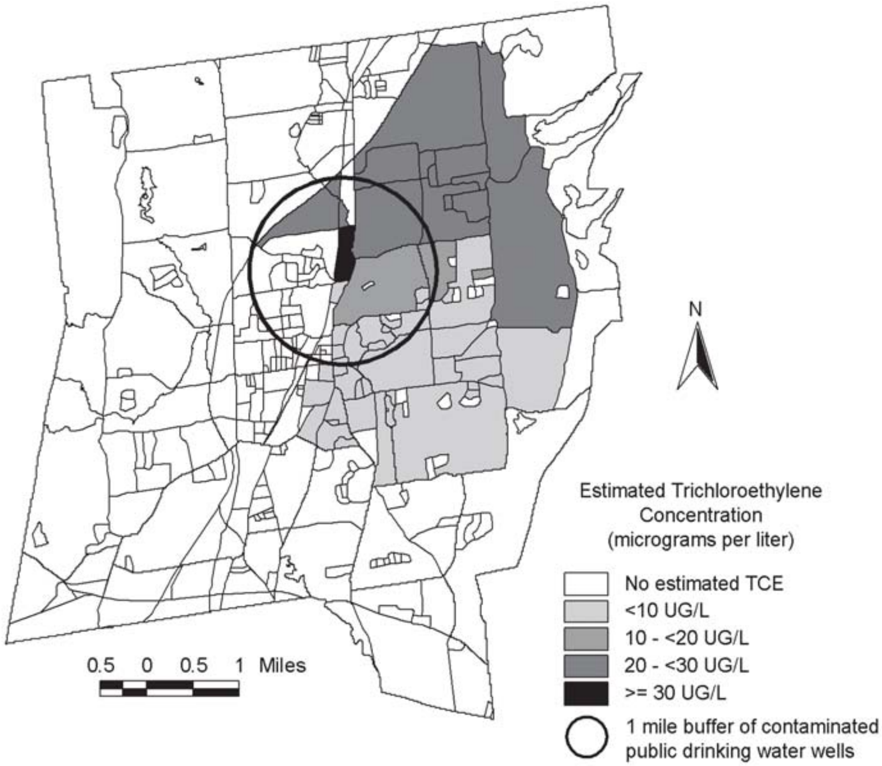


FIGURE 6.11. Census block areas that received contaminated drinking water from wells adjacent to a National Priority List hazardous waste site. A hydrological analysis of the water supply system estimated that areas northeast of the contaminated wells received the highest exposures. The geographical pattern of water contamination impact was influenced by water usage, competing sources of water, and hydrologic pressures in the water distribution system, and not simply distance from the contaminated wells. From Aye and Archambault (1997).

Estimating Populations in Hazard Zones

GIS have been used in a wide range of environmental health studies to estimate populations in zones of different environmental quality. In some cases, population data have been aggregated to areas, usually by residence. In other cases, health analysts have had access to disaggregate population data for geocoded points representing residences and other locations such as schools or for parcels.

One area of environmental health analysis has been the investigation of possible links between electromagnetic fields (EMFs) and a variety of human health problems. EMFs arise from the operation of electrical generators, distributors, and appliances, and are omnipresent in postindustrial societies. Epidemiological research has focused attention on the EMFs associated with electrical power systems, including generating stations, high-voltage transmission lines, and distribution lines.

To map and analyze EMFs associated with high-voltage transmission lines in Hartford County, Connecticut (Cromley & Joy, 1995), a GIS database of transmission lines located in the study area (Figure 6.12) was created by digitizing maps of line location and type obtained from the utility and registering the digitized lines to a less complete but more spatially accurate database of transmission lines compiled by the Connecticut Department of Environmental Protection. The FIELDS program, a software system available from Southern California Edison Company, was used to calculate the EMF field around each transmission line segment in the study area. The FIELDS program requires input on phase coordinates (horizontal and vertical), number of subconductors per bundle, conductor diameter, bundle diameter, line kilovolts, phase current in amps for the relevant time period, phase angle in degrees, ground wire coordinates (horizontal and vertical), ground wire diameter, ground wire current amps, and ground wire phase angle. These data were also obtained from the utility. Variables used to calculate the exposure fields were based on what a line was rated to carry during the winter and summer months. Actual readings are taken along line segments every 15 seconds by the utility, but these data would be too voluminous to process.

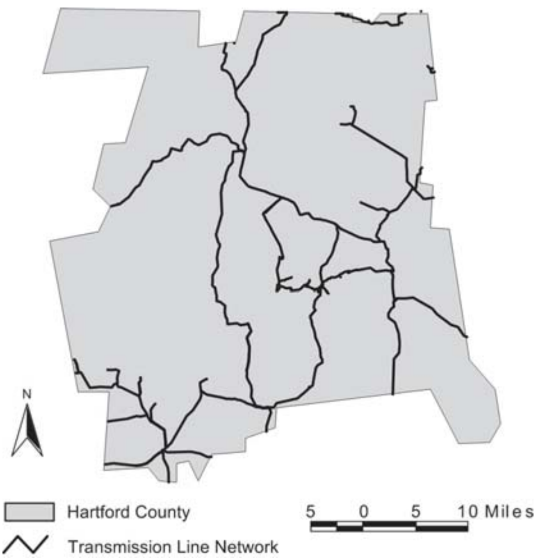


FIGURE 6.12. Electrical transmission lines in Hartford County. From Joy (1994). Copyright 1994 by K. P. Joy. Adapted by permission of the author.

Based on the results of the analysis for each line segment, the center of each line segment right-of-way as represented in the GIS was buffered to the distance calculated for an exposure field in excess of 2 milligauss. This buffered area minus the line segment right-of-way—where no development can exist—represented the area of 2 milligauss or greater exposure. This data layer represented the area where homes or schools might be exposed to EMFs at the specified level of interest.

To find out how many children in the study area might live or attend school within this zone, two analyses were performed. The first, to estimate the number of children living within the zone of interest, used aggregate data on population. The second, to estimate the number of children attending school within the zone of interest, used point data for school locations.

Working with Area Population Data

The estimate of children exposed by residence involved areal interpolation. ***Areal interpolation*** refers to a set of techniques to estimate the distribution of a phenomenon (in this case the number of children under 18 years of age) across a set of spatial units called ***target units*** (in this case the 2-milligauss exposure zones) based on the observed distribution of the same phenomenon across a set of spatial units called ***source units*** (in this case 1990 census blocks). A common approach to areal interpolation is the ***area weighting method***, relying on the concept of map overlay (Lam, 1983). In this approach, the variable "number of children under 18 in the census block" is weighted by the proportion of the census block's area that lies in the target EMF zone. The resulting number of children is then assigned to the EMF zone as part of that unit's population (Figure 6.13).

Areal interpolation can be enhanced by incorporating ancillary data (Flowerdew & Green, 1989). For example, if we know based on the distribution of streets or houses that no one actually lives in a certain part of the census block, we can remove that uninhabited zone from the source unit to derive a better estimate of the population residing in the area of overlap. This method is sometimes referred to as ***filtered area weighting***.

To estimate the number of children under 18 exposed based on residence, a database of 1990 census blocks was digitized using coincident features from Connecticut Department of Environmental Protection road, stream, and town boundary databases to build a census block database of greater positional accuracy than could be compiled from the TIGER/Line® files. With this database, it was possible to identify by polygon overlay all of the 1990 census blocks in each study area town within 1 mile of a transmission line. Clearly, some areas in a town are not close to transmission lines, but it may also be that no one actually lives in those areas. To develop a better picture of where people actually reside in the census blocks, a 300' buffer was created around the street network—excluding major highways—to represent the area in which residential development is most

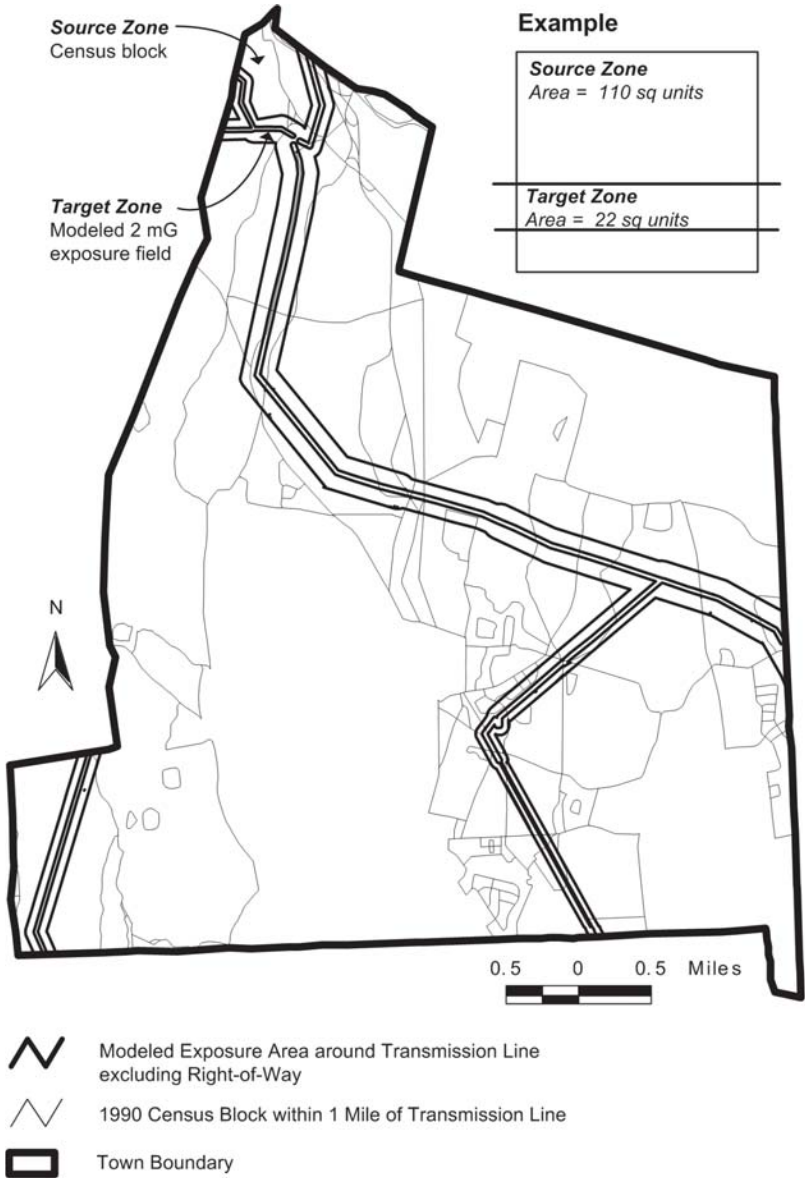


FIGURE 6.13. Areal interpolation by the area-weighting method to determine population within a risk area. In the example, 20 % of the population in the source zone would be estimated to be within the target zone because the target zone covers 22/110 or 20% of the source zone. The map shows the complex arrangement of 1990 census blocks as source zones with target zones determined by modeling the 2 mG exposure field around transmission lines excluding the power line right-of-way where no development can occur. From Joy (1994). Copyright 1994 by K. P. Joy. Adapted by permission of the author.

likely to occur on the basis of residential setback and lot configurations in the study area (Figure 6.14). The 1990 census block population for children under 18 was assigned uniformly to this buffered street area, and this area was overlaid with the database containing the 2-milligauss exposure fields. The estimated number of children exposed was calculated by multiplying the total number of children in the block by the percentage of the buffered street region in that block that coincided with an exposure field (Figure 6.15). Approximately 2% of children under age 18 were estimated to be exposed to transmission line EMFs based on this analysis, which also revealed the neighborhoods where that exposure was likely to occur.

Excluding population in a region from areas where there are no streets is an example of dasymetric mapping. ***Dasymetric mapping*** techniques exhaustively partition space into zones where the zone boundaries delineate areas where the variable value of interest is close to constant (Holt, Lo, & Hodler, 2004). Population density in a study region may not be constant across the region if population is more likely to be concentrated in some parts of the region than others. Land cover is commonly used as an ancillary data layer to identify areas where population can be excluded such as areas of surface water (Mennis, 2003). Dasymetric mapping combines data from the source zone with the ancillary data to produce a map of population distribution by zone within the study region (Figure 6.16).

As an alternative to buffering vector street networks and localizing census tract or block populations within the buffer areas, GIS and related technologies have been used to create high-resolution raster population distribution databases. LandScan USA, which grew out of earlier modeling efforts, uses a wide range of input data sources and dasymetric modeling techniques to create daytime and nighttime population distribution data with a spatial resolution of 3 arc seconds (approximately 90 meters) (Bhaduri, Bright, Coleman, & Urban, 2007). The LandScan™ data set is a worldwide population distribution data set based on a 30-minute by 30-minute lon/lat grid (Oak Ridge National Laboratory, 2008). Only the most recent LandScan database is distributed because the data are not considered adequate to support studies of migration or change detection.

Parcel data have also been used to identify more accurately than other methods the areas where people actually reside and to eliminate those areas where no one lives. Research on the relationships between residential proximity to a variety of sources of air pollution in the Bronx, New York, and asthma used cadastral data, in part because of the extremely high level of land use heterogeneity among neighboring parcels in this urban setting (Maantay, Maroko, & Porter-Morgan, 2008). Results from the analysis using cadastral data were consistent with findings from an analysis using areal interpolation with ancillary data. Both showed that proximity to air pollution sources is correlated with hospital admissions for asthma in the Bronx. The odds ratios were elevated, however, in the analysis using cadastral data, and the difference was especially pronounced for the effect of proximity to limited access highways.

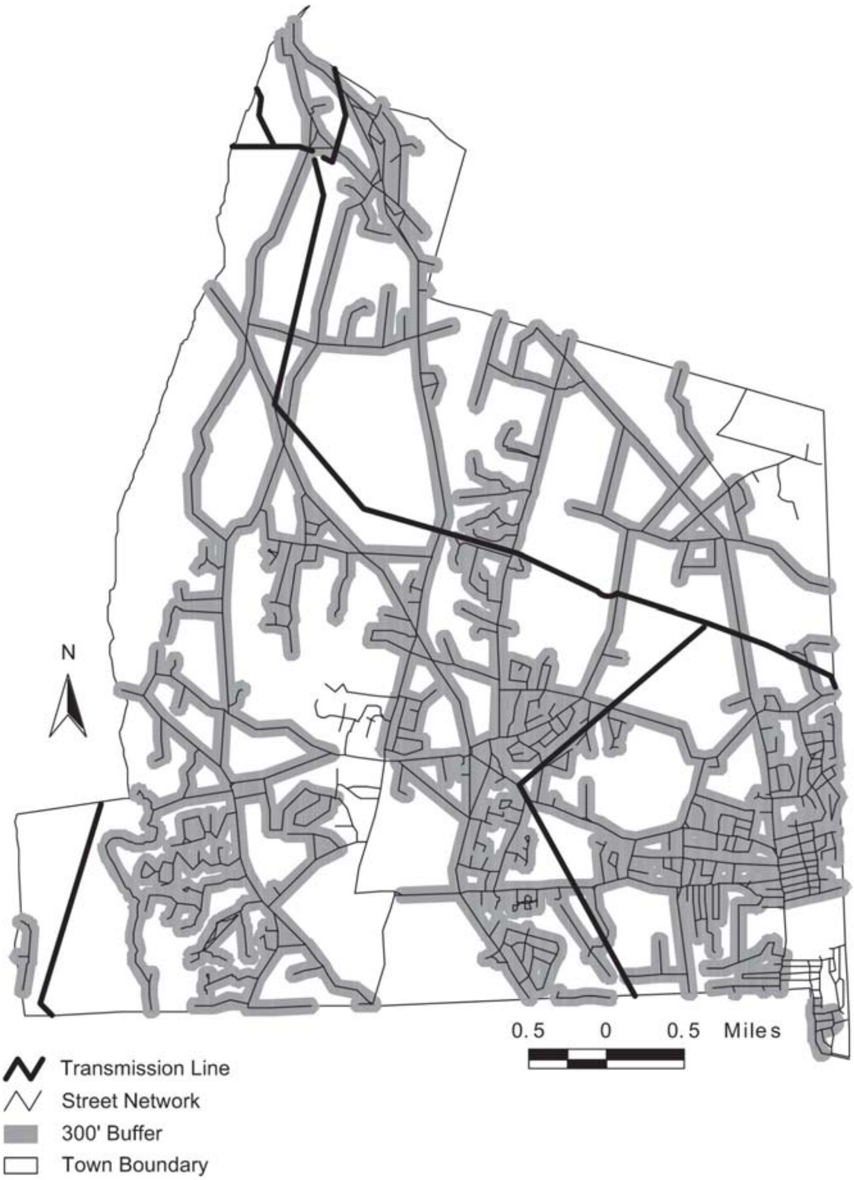


FIGURE 6.14. The buffered street network is a form of ancillary data that can be used to localize the population within a source zone like a census block to improve estimates based on areal interpolation. From Joy (1994). Copyright 1994 by K. P. Joy. Adapted by permission of the author.

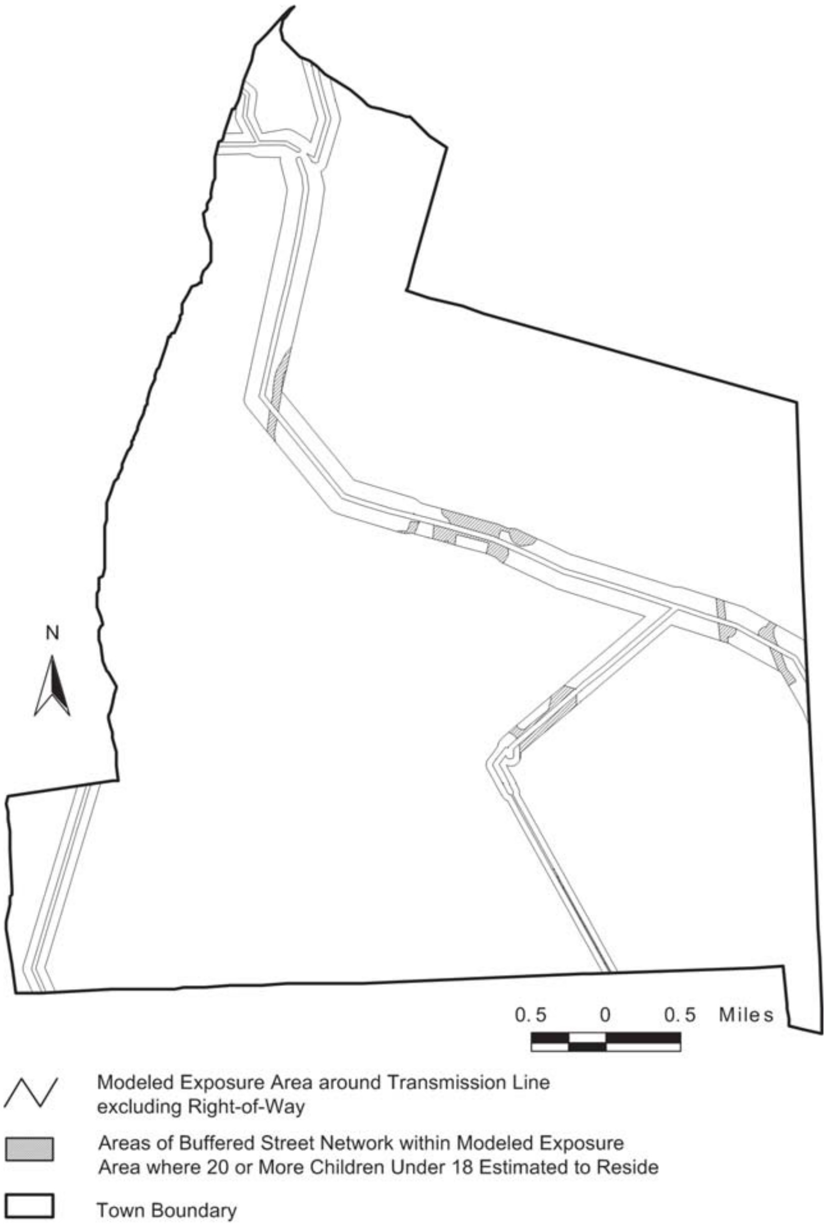
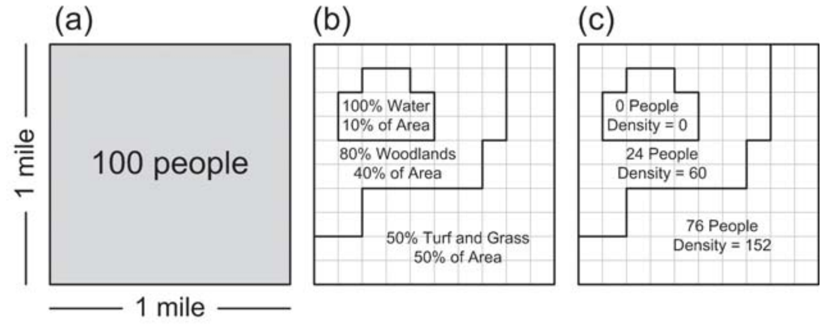


FIGURE 6.15. Areas within a modeled exposure zone where children might live based on position of the street network in relation to the exposure zone and the child population of the census block where the street segment is located. From Joy (1994). Copyright 1994 by K. P. Joy. Adapted by permission of the author.



Percent of area where people can reside in each land cover zone:
0 x 10% = 0% of the water area
20% x 40% = 8% of the woodlands area
50% x 50% = 25% of the turf and grass area

Total percent of land where people can live summed over all areas:
0% + 8% + 25% = 33%

Proportion of total area where people can reside by land cover type:
0% / 33% = 0% (water)
8% / 33% = 24% (woodlands)
25% / 33% = 76% (turf and grass)

Number of people in each land cover zone:
0% x 100 people = 0 in water zone
24% x 100 people = 24 people in woodlands
76% x 100 people = 76 people in turf and grass

Population density in each zone:
0 / 0.10 = 0 people/square mile in water zone
24 / 0.40 = 60 people/square mile in woodlands zone
76 / 0.50 = 152 people/square mile in woodlands zone

Population density in entire study area is 100
(0 x 0.10) + (60 x 0.40) + (152 x 0.50) = 100

FIGURE 6.16. An example of dasymetric mapping to model population distribution. A census unit with 100 people per square mile (a) has a land use pattern shown in the raster land cover data layer (b). Based on these two layers, the number of people and the population density of each zone can be estimated (c).

WORKING WITH POINT, GEOCODED ADDRESS, AND PARCEL DATA

The study of EMFs and child exposure considered exposures at home and at school. School lat/lon data were obtained from a gazetteer and projected based on the Connecticut State Plane Coordinate System. Information on 1990 school enrollment was collected from the Connecticut Department of Education and individual schools. To estimate the number of children attending school within an EMF exposure zone, a simple point-in-polygon overlay was performed to identify schools (points) located within EMF zones (polygons).

Locations of residences or other features such as schools can also be obtained from parcel data, address-match geocoding, or address points, as discussed in Chapter 3. A body of research has compared the methods for determining point locations (Zandbergen, 2007; Zandbergen & Green, 2007; Zandbergen, 2008; Whitsel, 2008), specifically in the context of environmental health research. Positional accuracy and completeness of point databases are two key measures of data quality that can affect detection of relationships between exposures and health outcomes.

Simulation studies can be used to investigate the impact of inaccuracies in point data in specific study settings. One such study, conducted in a county in Iowa, focused on residential addresses in rural areas because there is evidence that geocoded points in these areas are less accurate than geocoded points in urban communities (Mazumdar et al., 2008). The zone of environmental contamination was given. Health outcome data were generated for hypothetical individuals living at residential locations that were geocoded and then validated using a properly registered orthophoto image. The relationship between exposure and health outcomes was assessed. Then, the same relationship was estimated using two other commonly used approaches to geocoding health outcome data, address-match geocoding using TIGER/Line data from the U.S. Census, and geocoding using an E911 database. The E911 database was an address point database where the address location corresponded to the location where an emergency responder would turn off a public road to the private drive leading to the residence associated with the telephone number from which an emergency call originated. Because parcels in rural areas are so large and parcel centroids would have a high positional error, parcel geocoding was not evaluated. The results of the simulation show that analyses of the data produced by the less accurate geocoding methods were able to reproduce the disease odds ratio obtained using the more accurate point data, but the strength of the relationship declined with decreasing geocoding accuracy.

Even though the specific results of this research cannot be generalized to other study settings, the methods for evaluating the sensitivity of results based on different methods of determining point locations can be widely applied. Regardless of the methods used, when individual health records are incomplete and a matching point location cannot be found, the implications of subject loss also need to be evaluated (Gregorio, Cromley, Mrozinski, & Walsh, 1999).

Assessing Population Characteristics

Many health problems are age and sex specific, and methods that yield an estimate only of the total population within a zone of contamination are insufficient. Techniques have also been developed to estimate populations by age and sex. One approach uses raster data on population distribution. A total population surface was developed using dasymetric mapping techniques for a 90-meter grid. Additional 90-meter grids of age-sex proportions were estimated separately. A grid for age–sex proportion for the male cohort aged 55 to 64 was analyzed with the grid modeling the total population surface to yield a population surface of males aged 55 to 64 (Cai, Rushton, Bhaduri, Bright, & Coleman, 2006). This cohort was of interest for studying prostate cancer.

Screening Tools for Disproportionate Environmental Risk

Data on hazards, modeled fate and transport, pollutant health risk, and population have been combined in databases that can be used as screening tools in environmental studies. These tools are designed to identify which toxicants, emission sources, and locations are of greatest concern for human health. The U.S. Environmental Protection Agency has developed and updates two screening tools: Risk-Screening Environmental Indicators and National Air Toxics Assessments.

Risk-Screening Environmental Indicators (RSEI) incorporate amounts of chemical released, fate and transport of the chemical, route and extent of human exposure to the chemical, toxicity of the chemical, and the number of people affected (U.S. Environmental Protection Agency, 2010b). RSEI is based on TRI-reported data for release estimates for more than 600 chemicals and more than 50,000 reporting facilities. The most recent version is based on TRI releases for 1996 through 2007. Toxicity data for more than 400 chemicals are included in the model, along with population data from the U.S. Census. Toxicity weights are based on chronic health effects such as cancer rather than acute health effects or environmental impacts. The model produces RSEI scores, which are the products of the estimated dose of a chemical multiplied by the toxicity weight multiplied by the exposed population. Scores can be reported in the form of sorted lists, tables, graphs, and maps.

While RSEI includes releases to air, water, and land, the ***National Air Toxics Assessments (NATA)*** screening tool focuses on air toxics (U.S. Environmental Protection Agency, 2010c). A national inventory of emissions from outdoor stationary and mobile sources is compiled. From this inventory, ambient concentrations of air toxics are estimated along with exposed populations. Toxicity is assessed based on chronic exposure. Ambient and exposure concentrations and estimates of risk are generated at the census tract level for air toxics in each state. The most recent assessment based on data for 2002 was released in 2009.

These screening tools provide an alternative to assessments based on proximity and amount of hazard. Not all chemicals are equally hazardous to human health. Relatively small releases or low ambient concentrations may be a concern if the chemicals are highly toxic and the estimated exposed population is large. Data from these screening tools have been used in a number of studies addressing issues of environmental justice (Ash & Fetter, 2004; Apelberg, Buckley, & White, 2005; Morello-Frosch & Jesdale, 2006; Abel, 2007; Linder, Marko, & Sexton, 2008; Chakraborty, 2009).

Environmental justice is the fair treatment and meaningful involvement of all people with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies. Fair treatment means that no group should bear a disproportionate share of environmental hazards resulting from the operations or policies of industrial, governmental, or commercial entities. Meaningful involvement means that people have an opportunity to influence decisions affecting their environment and health. The environmental justice movement began in the United States in 1990 when the Congressional Black Caucus met with EPA officials to discuss evidence that minority and low-income communities bore higher environmental risk burdens than the general population (U.S. Environmental Protection Agency, 1992). Environmental justice concerns have since broadened in scope to include sustainability (Agyeman & Evans, 2004), and environmental legal instruments such as the Aarhus Convention have been developed to address these concerns in other regions (Justice and Environment, 2011).

RSEI data were used in a study of the Philadelphia Metropolitan Statistical Area, a nine-county region, to characterize who lived near the facilities producing the worst pollutants (Sicotte & Swanson, 2007). After testing and correcting for spatial autocorrelation, researchers developed models to test four hypotheses of differential impacts on racial/ethnic minorities, disadvantaged populations, working-class populations, and industrial workers. The results supported the hypotheses that places with a high percentage of minority populations, who were low income, less educated, and unemployed, and a high percentage of industrial workers were near the most hazardous facilities. Interestingly, there was evidence of spatial variability in these relationships. The percent black increased with the log hazard score in all nine counties together and within areas in three counties separately. Within one county, however, there was a negative relationship between percent black and hazard score. Spatially varying processes are apparent in many studies of health disparities, as discussed in Chapter 11.

GIS analyses to model exposure often incorporate data on health outcomes. The development of biomonitoring data sources may make it possible to investigate, after exposure areas have been identified, which agents and how much of the agents are present in people even before the emergence of disease.

GIS and Dose

Biomonitoring involves analyzing the organisms that live in an area to make inferences about ecological conditions in the area. Biomonitoring of aquatic invertebrates to assess the water quality of streams, lakes, and wetlands has a long history. In a study conducted in Cologne, Germany, analysis of airborne pollutants that had accumulated on pine needles demonstrated that passive sampling of natural vegetation could provide environmental monitoring data of high temporal and spatial resolution (Urbat, Lehndorff, & Schwark, 2004). Although biomonitoring of plants and animals has been used to assess patterns of ecosystem health and environmental quality, human biomonitoring has been adopted as a method for documenting how much of various contaminants may actually be absorbed by the human body.

Human biomonitoring programs have been established to sample blood, urine, breast milk, and body tissue to assess human exposure to natural and synthetic chemicals (Angerer, Ewers, & Wilhelm, 2007). When an individual is exposed to a contaminant, the ***potential dose*** is the amount of the substance contained in what is being swallowed, breathed, or touched. The ***internal*** or ***absorbed dose*** is the amount of the substance that passes through the absorption barriers of the body through physical and biological processes. ***Body burden*** is the total amount of a substance observed in a body at a given time. To some health researchers, a contaminant that is present in large quantities in the environment but is less likely to be absorbed may be considered less a concern than other contaminants with greater uptake in the body.

In the United States, the Environmental Health Laboratory of the National Center for Environmental Health of the Centers for Disease Control and Prevention conducts the National Biomonitoring Program. The program published the first *National Report on Human Exposure to Environmental Chemicals* in 2001. The program relies on blood and urine samples collected from the population sampled in the National Health and Nutrition Examination Survey (NHANES). The fourth report, published in 2009, is based on NHANES data from 1999–2004 and provides information on more than 200 chemicals (Centers for Disease Control and Prevention, 2009c). Data are presented for the total population and by age, sex, and race/ethnicity. No data are reported for states or communities. The current NHANES sample design does not permit examination of exposure levels by locality, state, or region, by proximity to sources of exposure, or by season.

After funding planning grants to 25 state and regional organizations, the Centers for Disease Control and Prevention funded New Hampshire, New York, and the Rocky Mountain Biomonitoring Consortium, including Arizona, Colorado, Montana, New Mexico, Utah, and Wyoming in 2003, to develop biomonitoring programs. California established the first state biomonitoring program in 2006. The California Environmental Contaminant Biomonitoring Program is designed to evaluate the presence of toxicants in a representative sample of the state's population, identify trends in the levels of toxic chemicals over time, assess the effectiveness of regulatory and public health programs to decrease exposure to specific chemicals, and provide opportunities for public participation in environmental health issues (Office of Environmental Health Hazard Assessment, 2008).

A number of other countries have established human biomonitoring programs (Porta et al., 2008). The German Environmental Surveys have been carried out since the mid-1980s (Schulz et al., 2007), and the European Union is developing a biomonitoring framework (Smolders & Schoeters, 2007; Smolders, Casteleyn, Joas, & Schoeters, 2008). In Canada, a biomonitoring component has been added to the Canadian Health Measures Survey carried out from 2007 to 2009 (Health Canada Santé Canada, 2007).

Biomonitoring data can be used to reconstruct exposure and to assess biological doses and their effects. Linking biomarker data to exposure data can be challenging, especially when there are many potential sources and pathways (Georgopoulos et al., 2009). Some databases, like NHANES, include complementary exposure data. Biomonitoring data have been used with fate and exposure data to evaluate the relative contributions of different organophosphorus pesticide exposure pathways in a cohort of pregnant Latina women in Salinas, California (McKone et al., 2007). Relative to NHANES data, a statistically significant added intake of pesticides was noted and attributed to nondietary exposures from local agriculture.

Analysts have also used GIS modeling to develop measures for comparison to biomarker-derived measures of pesticide exposure. California PUR data were linked with geocoded residential history data for cases and controls in Kern County, California (Ritz & Costello, 2006), using GIS methods for linking the data (Rull & Ritz, 2003). Lipid-adjusted dichlorodiphenyldichloroethylene (DDE) serum levels were taken for the study subjects. DDE is a metabolite of DDT, which has a physiological half-life of sufficient length that it can serve as a biomarker for long-term pesticide exposure. The estimate of environmental exposure was used along with information on the personal history of loading or handling pesticides and use of pesticides in the home to predict DDE level. The ***sensitivity*** of the GIS model was poor; only 38% of subjects with high serum DDE levels were identified. The ***specificity*** of the GIS model, however, was good; 87% of the unexposed subjects were correctly identified. GIS models may be useful in studies where biomonitoring data are not available.

Despite increasing interest in biomonitoring data, health researchers have also urged recognition of the limitations of these data (Needham, Barr, & Calafat, 2004; Bhatia, Brenner, Salgado, Shamasunder, & Prakash, 2004). People who receive the same exposure may receive different biological doses due to ***pharmacokinetics***, differences in absorption, distribution, metabolism, and elimination of a chemical. Individuals who receive the same biological dose may experience different effects due to ***pharmacodynamics***. Individual factors including age, sex, and genetics, environmental factors including geography and housing, and behavioral factors may all be important in explaining how environmental exposures affect health outcomes.

GIS and Outcome Surveillance

Physical and chemical agents like those studied in the earlier examples in this chapter may produce several different kinds of adverse health and environmental effects (Hallenbeck, 1993; Stockwell et al., 1993; Smolders & Schoeters, 2007). Cancer is an effect due to toxic mechanisms operating in nonreproductive cells. Developmental effects, like death of the fetus; structural abnormalities, like cleft palate, that are observable at birth; or postnatal functional disabilities of the central nervous, respiratory, and intestinal systems can also result from toxic mechanisms operating in nonreproductive cells. When DNA-interactive mechanisms operate in reproductive cells, hereditary effects can occur. Finally, organ and tissue effects, like damage to the liver, kidneys, or lungs or to nerve tissue, are due to nongenotoxic mechanisms operating in nonreproductive cells. The following sections offer examples of how GIS techniques have been used to investigate these various health outcomes.

Cancer

When the sources of toxicants are known, studies of the relationship between environmental conditions and health outcomes like cancer have used focused cluster methods. Unlike the clustering methods discussed in Chapter 5, which are used to investigate general patterns of health outcomes or identify specific clusters anywhere in a region of interest, a ***focused test*** addresses clustering in relation to identified locations where increased disease risk might be expected (Waller & Gotway, 2004; Lawson, 2006). In this type of cluster analysis, the source locations or foci are identified first, and the analysis then looks for evidence of a high occurrence of disease around the foci.

The null hypothesis is that the locations of cases of disease reflect an underlying heterogeneous Poisson distribution with a constant level of individual risk from region to region; that is, there is no cluster of cases related to the location of the pollution source. The alternative hypothesis is that there is an increase in individual risk based on the exposure values assigned to the individual's region. The exposure values can be modeled as binary, with all unexposed regions having a value of 0 and all regions within a specified distance having a value of 1, or the exposure value can vary as a continuous function of distance from the source (Wartenberg & Greenberg, 1990).

Statistical power has been identified as an important issue in cluster detection. Because the dispersion of contaminants around a pollution source can vary in size and shape, the resulting cases of disease may also exhibit a variety of spatial patterns. Depending on the test applied and the size and shape of the cluster, the statistical power of focused tests may vary (Puett et al., 2005).

A focused test was used to study a municipal solid waste incinerator in Besançon, France, emitting high levels of dioxin, and found significant clusters of both soft-tissue sarcomas and non-Hodgkins lymphomas (Viel, Arveux, Baverel, & Cahn, 2000). In this study, the focused tests were used first to identify elevated risk around the specific source. Then, a space–time interaction test was used to determine whether clustering was in evidence throughout the time period of interest. Finally, nonfocused clustering methods were used to identify the location and significance of any clusters not located near the facility. Combining these approaches allowed the investigators to identify an elevated risk around the facility, to assess whether an observed cluster corresponded to the latency period, and to rule out the existence of clusters located elsewhere, providing support for a relation between plant location and cancer due to dioxide emission.

The source locations of interest in health outcome studies may be points, lines, or areas. A GIS developed to model household magnetic fields from power lines found a significant association with childhood leukemia in Los Angeles County after analyzing both exposure measurements based on wire codes and 24-hour measurements of the magnetic fields taken in the bedrooms of cases and controls (Bowman, 2000). To develop the wire code model for magnetic fields associated with electric transmission and distribution systems, magnetic field measurements were fitted by nonlinear regression to a function of wire configuration attributes (Bowman, Thomas, Jiang, Jiang, & Peters, 1999). These measurements were 24-hour bedroom measurements taken at 288 homes. Case–control data on childhood leukemia in Los Angeles County were reanalyzed to investigate associations between observed magnetic fields and the predicted magnetic fields and childhood leukemia (Thomas, Bowman, Jiang, Jiang, & Peters, 1999). Although the measured fields were not associated with childhood leukemia, the risks were significant for predicted magnetic fields above 1.25 milligauss and a significant dose–response effect was noted.

Using GIS to display the spatial structure of the transmission system and its wire code characteristics appeared to assess the leukemia risk associated with a child's long-term residential magnetic field exposure better than 24-hour measurements. The GIS approach enabled assessment of exposure with more subjects and more previous residences because it did not entail obtaining measurements in the field at hundreds of homes. This increased the power of the analysis. Also, 24-hour EMF measurements are strongly affected by short-term fluctuations in usage that do not yield a reliable picture of long-term exposure. The GIS model also creates a tool for retrospective studies of exposure because it only requires data on the wire code configuration and the locations of residences. Also, the model can be used to investigate possible links between other cancers, like breast cancer, and EMFs.

Developmental Effects

Reproductive outcomes are believed to be sensitive to many environmental influences (Stallones, Nuckols, & Berry, 1992). A hospital-based case–control study of stillbirths in a community in central Texas investigated the effects of chronic inhalation of low levels of arsenic (Ihrig, Shalat, & Baynes, 1998). A plant in the community producing arsenic-based agricultural products had been in operation for more than 60 years. Arsenic exposure levels were estimated from airborne emission estimates by using an atmospheric dispersion model linked to a GIS. Exposure was then assessed based on the residential address of the mother at the time of delivery. Exposure was included as a categorical variable in a conditional logistic regression model. An exposure-and-race/ethnicity interaction variable was also included to reflect the fact that members of particular groups within the population may be concentrated in certain residential neighborhoods. The prevalence of stillbirth was significantly higher among Hispanics living in high exposure areas.

A study in Harris County, Texas, used a space–time clustering method implemented in SaTScan™ to identify a significant cluster of orofacial cleft palate births in the northwestern part of the county (Cech, Burau, & Walston, 2007). In this area and some other parts of the county, residents were served by drinking water wells, and elevated levels of radon-222 and radium-226 were detected in the wells and in tap water. Areas in the county supplied from surface sources showed no concentrations of radionuclides above water quality standards.

Hereditary Effects

With research advances in genetics, studies have investigated the potential for inducing mutations in the ***human germline***, the sequence of cells containing genetic material that can be passed to children (Somers & Cooper, 2009). Studies have also documented that exposure to air pollution damages DNA in human sperm (Rubes et al., 2005) and that a particular genotype influences susceptibility to the effects of air pollution on sperm (Rubes, Selevan, Sram, Evenson, & Perreault, 2007). To date, research on the effects of pollution on germline cell damage in animals and humans has not generally incorporated GIS or spatial statistical modeling techniques, but GIS has a potential role to play in large-scale epidemiological studies of pollution, genetic conditions, and inherited disorders.

Elevated Blood Lead Levels

Data on health problems like cancer or birth defects can be obtained from tumor registries or vital statistics registration systems, as discussed in Chapter 3. Data on other health effects, like elevated blood lead levels, are often collected as part of screening programs. ***Screening*** is "the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures … [that] can be applied rapidly and usually cheaply" (Eylenbosch & Noah, 1988, p. 279). Unlike vital statistics databases or tumor registries, which are relatively complete, screening databases can be highly biased in representing the geographical distribution of a health problem, particularly if screening is not mandatory. Figure 6.17 reveals how misleading maps of the distribution of health problems can be if spatial biases in screening are not explicitly described.

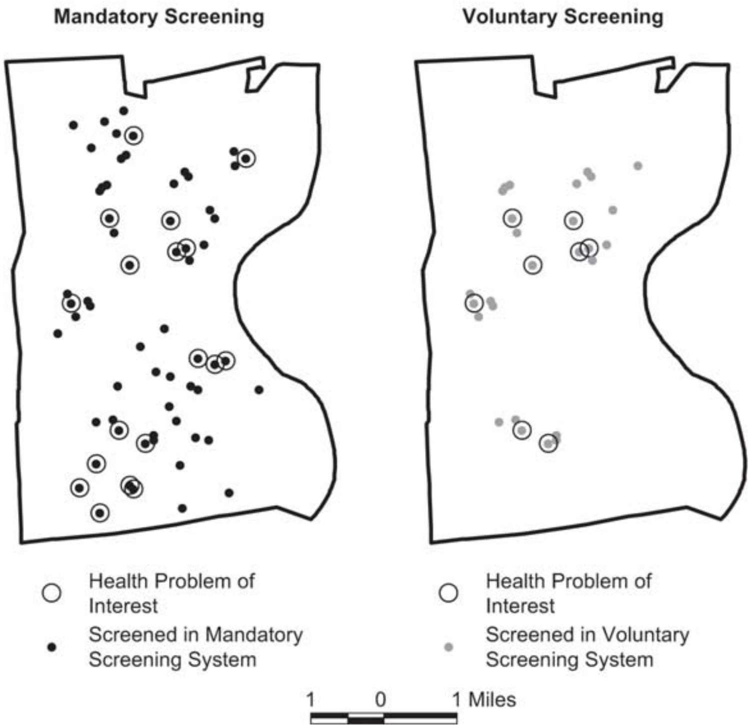


FIGURE 6.17. A mandatory screening system identifies the distribution of health problems of interest within the screened population. If the screening system is voluntary, fewer people may be screened and the distribution of identified health problems of interest may be biased. In the example above, an apparent concentration of cases in the north end of town is probably a result of more people having been screened there. Cases in the south end of town were not detected because the voluntary system was less effective in screening people in that section of town.

For some health problems, like lead poisoning, links to a toxicant have been established, and environmental risk factors have been identified based on our understanding of the distribution of the toxicant and how people come into contact with it in their daily lives. For these problems, GIS analysis has been used to map known risk factors and health outcomes to support the design of public health intervention strategies. In New Jersey, analysts mapped data related to several known risk factors for lead poisoning (Guthe, Tucker, & Murphy, 1992).

Data from the TRI on industrial sites emitting lead and from a state database of hazardous waste site locations were obtained. Traffic volume estimates from the New Jersey Department of Transportation were used to identify segments in the road network with high traffic volumes. EPA lead emission factors were then applied to estimate lead emissions from vehicles. In addition, data identifying census tracts exceeding threshold levels for number of structures built before 1940 and number of children under 5 years of age were compiled to depict exposure to lead paint in residences. The GIS was used to integrate these data and display the risk factors in relation to the distribution of children with high blood lead levels.

Overall, the results of the analysis showed a spatial correlation among sources of lead, susceptible populations, and health outcomes in the study area. In addition to identifying neighborhoods where numerous sources of lead and high blood lead levels were observed, the analysis highlighted other regions within the study area where frequencies of elevated blood lead levels were higher or lower than expected based on the sources of lead present. These results suggested limitations in the screening data and in the lead source databases.

The spatial patterns observed aided the New Jersey Department of Health in developing soil sampling and lead exposure research and in community outreach efforts to prevent lead exposure. There is evidence that uncovering elevated blood lead levels can lead to effective interventions. In the United States, the prevalence of elevated blood lead levels in children aged 1 to 5 based on NHANES data fell from 8.6% in the 1988–1991 survey period to 1.4% in the 1999–2004 survey period (Jones et al., 2009).

GIS and Environmental Risk Management

GIS can make contributions to risk assessment, primarily by supporting better modeling of geographical distributions of hazards, susceptible populations, exposures, and health outcomes. GIS also has a role to play in ***environmental risk management***, a social and political process that involves the selection and implementation of strategies for the regulation or control of identified hazards. Once priorities have been set for regulation or control of particular hazards, GIS can be used to identify the locations of entities producing toxicants for targeting intervention activities.

Clearly, there are opportunities for integrating risk assessment models into GIS. More than 1 billion tons of hazardous materials are transported in the United States annually across all modes (Lepofsky, Abkowitz, & Cheng, 1993). These shipments occur in all regions of the country. A highway transportation risk assessment software system uses a GIS to allow analysts to select highway or rail routes interactively and obtain a route-specific risk assessment for shipment of radioactive materials (Moore, Sandquist, & Slaughter, 1994).

In California, the Highway Patrol developed a GIS to support risk assessment and risk management of hazardous materials shipments (Lepofsky et al., 1993). Particular attention was paid to poisonous gases. The geographic database included a digitized network of all interstates, most federal and state highways, and selected county and local roads extracted from the National Highway Planning Network and augmented to include all highways in the state and some additional county and local roads.

Related technologies can also be used to monitor the consequences of regulatory actions. The efforts of the Chinese government to address air pollution problems in advance of the 2008 summer Olympics created an opportunity for using satellites to track the impact of air pollution control measures. Researchers from the U.S. National Aeronautics and Space Administration (NASA) were able to detect that levels of nitrogen dioxide fell nearly 50% and levels of carbon monoxide fell 20% in the vicinity of Beijing during the two months when factories were temporarily closed and travel by car was prohibited (Voiland, 2008). After the restrictions were lifted, the levels of pollutants increased again.

During the late 1980s, comparative risk assessment emerged as an approach to prioritizing environmental management efforts (Finkel & Golding, 1994). The "risk-based" paradigm is concerned with evaluating whether the relative efforts—regulation, allocation of resources for monitoring, and mitigation—devoted to reducing risks are in reasonable proportion to the seriousness of the risks being compared.

The EPA's interest in comparative risk assessment provoked a debate raising methodological, procedural, and implementation concerns about adopting comparative risk assessment as the basis for environmental protection resource allocation, and a fundamental objection to "ranking" instead of "addressing" environmental problems (O'Brien, 1994). Recently, accountability studies, seeking to link environmental management actions to outcomes, have been suggested as an approach in environmental assessment (McKone, Ryan, & Özkaynak, 2009). Respiratory health, for example, may be affected by motor vehicle emissions and emissions from specific types of facilities. An accountability study would attempt to account for the level of health benefit that would result from actions reducing specific types of pollution from specific sources, presumably providing support for prioritizing some types of regulatory action over others. This approach seems to be a newer version of the comparative risk assessment paradigm.

Although GIS can aid the process of quantitative risk assessment through epidemiological investigation and can be used to support targeting geographic areas for protection (Habicht, 1994), the role of GIS in risk management is not inextricably tied to the risk-based paradigm. GIS can be used to depict the geographical structure of an environmental problem of interest, for example, delineating watershed and wellhead protection areas (Chernin, 1995; Rifai, Hendricks, Kilborn, & Bedient, 1993; Hammen & Gerla, 1994), regardless of the problem's rank in a list of environmental problems prioritized by risk.

Issues in Environmental Health Mapping and Analysis

GIS have made it easier for environmental analysts to produce hazard maps demonstrating the sensitivity of hazard patterns to changes in modeling criteria (Tim, 1995). However, "GIS-generated maps may be viewed by users as having greater reliability than is warranted" (Wagenet & Hutson, 1996). The U.S. Environmental Protection Agency emphasizes the limitations of screening tools such as Risk-Screening Environmental Indicators and National Air Toxics Assessments. Depending on the modeling criteria, vulnerability and hazard maps produced from the same data may appear very different. It is important to define and understand exactly the databases that a particular map displays.

Because maps are often interpreted as precise portrayals of reality rather than an analyst's view of data, efforts to map hazards, environmental quality, and community vulnerability should be undertaken with particular care. Many residents of environmentally degraded areas do not need analyses and maps to describe the hazards in their communities or the adverse psychological and financial impacts that may exacerbate both the direct health effects of exposure to contaminants and the incidence of stress-related diseases. Community groups with vested political and economic interests will not always welcome these maps, even when the data are complete and accurate, the analyses are conceptually sound, and the results are robust.

Environmental health research seeking to understand the varied settings in which people live is inextricably connected to social concerns about environmental quality and to disparities that result in differential burdens on communities. The collection, interpretation, and communication of data from human biomonitoring studies are raising many ethical issues in the United States and in other countries (Sepai, Collier, Van Tongelen, & Casteleyn, 2008; Morello-Frosch et al., 2009). Even in environmental health research that does not rely on biomonitoring data, the range of ethical issues is wide (Sharp, 2003): choosing toxicants and geographic localities to study, involving members of communities of interest, interpreting and disseminating research findings that are sometimes ambiguous or inconclusive, assessing biological mechanisms by which toxicants affect health, exploring the role of genetics in susceptibility to toxicants, and evaluating interventions to improve the health of the environment and the people who live in it. In this context, it is important to recognize the role GIS can play not just in mapping environmental health problems, but in managing and making accessible to the public the large, spatially referenced databases of environmental information that citizens have the right to know.

Conclusion

The decline in incidence of infectious disease in many regions in the 1950s and 1960s shifted the focus of public health and medical research efforts to diseases like heart disease, cancer, asthma, and other chronic health problems that were increasing in incidence. Many researchers have investigated environmental risk factors as explanations for these observed health outcomes. This research is now being supported by the use of GIS to model complex hazard–exposure–dose–response processes in time and space.

GIS is only one technology in a set of emerging technologies leading to personalized exposure assessment. Over the last decade, in addition to the adoption of macroscale laser and infrared radiation sensors to assess population exposures to air pollutants, microscale sensors such as personal dosimeters have been used to monitor levels of toxicants in the residence, workplace, and personal environment (Weis et al., 2005). These developments and the expansion of human biomonitoring are contributing to a shift away from compliance and operational monitoring of pollution to exposure assessment in environmental health.

An interesting development in our understanding of chronic disease is a growing body of research suggesting that some health problems like arteriosclerosis may be partly the result of infectious agents (Ewald & Cochran, 1999). This introduces a new level of information that may need to be considered to understand how environmental conditions, human behavior, and infectious agents work together to cause disease and how we can design public health interventions to reduce the occurrence of disease. The role of GIS in analyzing infectious disease is considered in Chapters 7 and 8.

CHAPTER 7

Analyzing the Risk and Spread of Infectious Diseases

The resurgence of infectious diseases, some new and unfamiliar, others with a long history in human populations, has been identified as a global threat to human health at the end of a century of scientific and medical advances that had seemingly conquered infection as a cause of death (Morens, Folkers, & Fauci, 2004). In addition to the immediate public health need for improved infectious disease surveillance and response, the renewed concern for infectious disease has been accompanied by a reevaluation of contemporary risk factor epidemiology (Pearce, 1996; Susser & Susser, 1996a; Susser & Susser, 1996b). Although the risk factor approach recognizes that many public health problems are multicausal, critics argue that focusing even on multiple individual risk factors has disconnected epidemiology "from an examination of the broader historical and social forces that help to shape disease patterns in populations" (Nasca, 1997) and from the "way in which people interpret their health-related behavior" (Lawson & Floyd, 1996). In response, attention has turned to theories of social epidemiology that emphasize contextual, political-economic, social, and environmental influences on health (Krieger, 2001).

Factors cited as contributing to the reemergence of infectious disease include land use change affecting vector and host habitats and human interaction with vector and host populations, urbanization, transportation technology affecting migration and population mobility, and changes in the ways water and food are delivered (Mayer, 2000; Morens, Folkers, & Fauci, 2004). Many of these factors have an important geographical dimension.

The biological basis of infectious diseases is crucial for understanding where and why the diseases occur. Infectious diseases are caused by living microorganisms, the disease ***agents*** or ***pathogens***. They spread among human or animal ***hosts***, either by direct transmission or via a ***vector*** that transmits the disease from host to host. Thus, the occurrence and spread of infectious disease depends on hosts' exposure and susceptibility to pathogens and interactions among hosts, agents, and vectors.

Chapters 7 and 8 explore how geographic information systems have been used in the study of emerging and reemerging infectious diseases, particularly those transmitted directly from person to person (Chapter 7) and those transmitted by vector (Chapter 8). For many infectious diseases, unlike cancers, birth defects, and other health problems associated with exposure to toxicants, etiology is known and the diseases spread via contact among hosts or their exposures to vectors. These differences in disease process give rise to different modeling approaches.

This chapter examines the use of GIS in analyzing infectious diseases that spread directly from person to person. These "nonvectored" diseases include some of the most significant public health concerns in the United States—HIV/AIDS, tuberculosis, measles, influenza, and gonorrhea, among others. Their uneven geographical distributions reflect the social and environmental conditions that affect risk and susceptibility and the social interactions and behaviors that facilitate transmission. A distinctive feature is that the diseases spread over time from person to person and place to place, often in epidemic and pandemic forms.

Nonvectored diseases spread through a variety of different means. Many, like HIV/AIDS, syphilis, and impetigo, spread directly through skin or sexual contact. Others can exist in the environment, persisting in air, water, soil, or food. These diseases spread via our most basic, everyday behaviors—eating, drinking, breathing, and working. Airborne transmission occurs when pathogens spread from host to host through the process of respiration, as occurs with influenza, tuberculosis, or the common cold. Other diseases, such as cryptosporidiosis and cholera, are transmitted through contaminated water or food. Some pathogens can persist in soil, giving rise to the transmission of diseases like tetanus.

The mode of transmission is a critical factor in any GIS assessment of nonvectored diseases. It influences the kinds of geographical questions asked about the disease, the types of analyses performed, and the types of data layers that need to be included in a GIS. In mapping and analyzing waterborne diseases, for example, the water distribution network, including reservoirs, mains, and private wells, is crucial. By comparison, airborne diseases emerge out of geographical patterns of human contact and interaction, as linked to housing conditions, crowding, and personal contact in facilities such as schools, daycare centers, jails, and workplaces. Sexually transmitted diseases (STDs), on the other hand, reflect intimate relations that are embedded in broader concepts of identity, sexuality, and gender. The relationships among people that channel disease spread are hidden from view, not clearly visible in the day-to-day movements of people that comprise traditional spatial patterns of human interaction.

Transmission of infectious diseases also depends on the host's immunity. ***Immunity*** is the host's ability to resist the pathogenic effects of infection. Some immunity is acquired in response to repeated infection, as occurs with malaria, influenza, or the common cold. ***Immunization*** is a way of artificially inducing acquired immunity by injecting small doses of toxins in the host. ***Innate immunity*** refers to the body's ability to harness its own biological resources to ward off infection. Although immunity rests partly in biology, for example, the host's genetic makeup, it also has an important social basis. Nutrition and malnutrition, exposure to environmental contaminants, and past exposure to infections affect the immune response. These reflect people's access to nutritious food, their home and work environments, and the stresses and risks they face in their daily lives. Understanding how these social and geographical processes influence the immune response, the susceptibility to and risk of infection, and the severity of illness is crucial in explaining the geography of infectious diseases.

The spread of infectious diseases among hosts can be described in general terms using ***epidemic models*** (Thomas, 1992). These models chart the extent of disease spread as a function of the sizes of susceptible and infected populations, the degree of mixing between them, and the transmission rate and incubation period for the disease. The simplest models, SIR models, assume a population consisting of three groups: susceptibles (S), infecteds (I), and recovereds (R), and even mixing of the S and I populations. According to these models, an epidemic begins slowly, but builds up as the number of infecteds increases (Figure 7.1). The peak transmission occurs when the S and I populations are approximately equal in size. After the peak is reached, the epidemic diminishes as the number of susceptibles declines. More complex models have been developed, and these have proven fairly accurate in charting the course of epidemics over time. However, many models are "single region" and thus ignore the geographical dimensions of disease spread (Thomas, 1992).

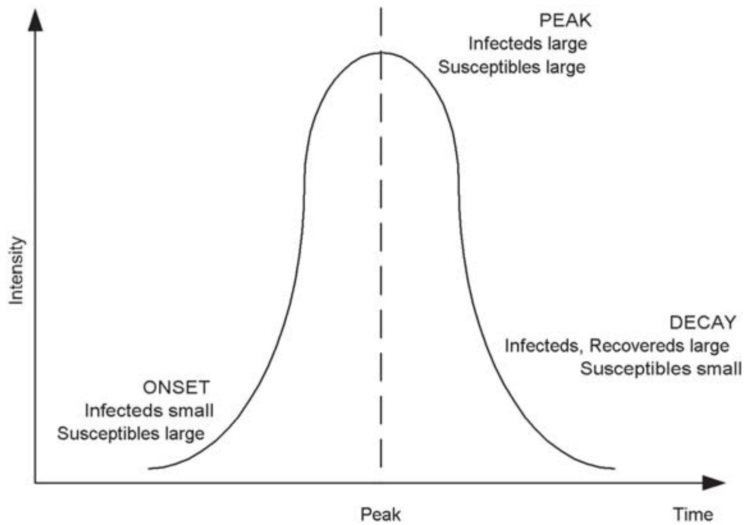


FIGURE 7.1. An epidemic curve, showing changes in susceptible and infected populations.

Spatial Diffusion

The geographical patterns of interaction between infected and susceptible hosts are crucial for understanding how and where infectious diseases spread. ***Spatial diffusion*** describes the movement of phenomena—people, goods, ideas, innovations, and diseases—through space and time. Spatial diffusion of disease occurs when a disease is transmitted to new locations. Sometimes diseases follow a pattern of ***contagious diffusion***, spreading gradually outward from a point of origin to nearby locations (Cliff & Haggett, 1988). Contagious diffusion reflects the localized nature of human spatial interaction: people are more likely to interact with their neighbors than with those located farther away. Constraints on mobility related to age, low income, disability, or poor access to transportation may lead to highly localized patterns of spatial interaction. Diseases also spread contagiously between cities and surrounding suburbs, following commuting flows and social interactions. The prevalence of AIDS in suburban communities is strongly correlated with the volume of workers who commute to central cities (Figure 7.2). These connections among diverse communities are critically important for health policy: the health problems of central cities and suburbs are inextricably linked by flows of people and interactions among them. Inner city clusters of communicable disease can act as epidemic pumps that spread disease to suburban areas (Wallace & Wallace, 1998).

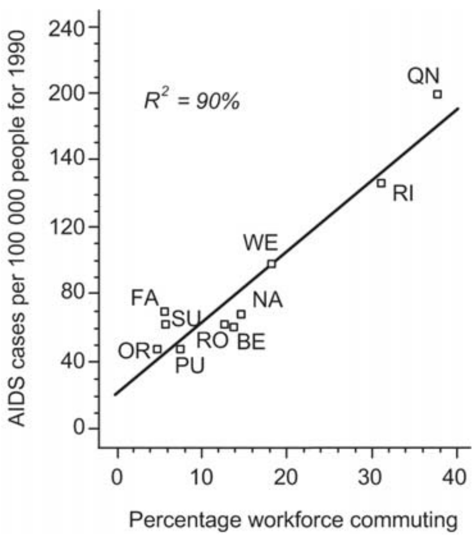


FIGURE 7.2. For 10 affluent counties in the New York metropolitan region, cumulative AIDS cases through 1990 per 1,000 people are highly correlated with the percentage of workforce commuting into Manhattan. Spatial diffusion of communicable diseases is often channeled through commuting flows. From Wallace and Wallace (1998). Copyright 1998 by Verso Books. Reprinted by permission.

Examining disease diffusion patterns, we find that diseases often “jump” from place to place, rather than following the gradual outward expansion of contagious diffusion. In ***hierarchical diffusion***, diseases spread via the urban hierarchy, starting in large cities and spreading over time to medium-sized cities, then to smaller cities and towns. The large populations, strong transportation connections, and movements of people among large cities channel hierarchical diffusion. Transmission occurs over long distances—for example, from New York City to Chicago and Los Angeles—propelled by the strong interactions among large urban populations. ***Network diffusion*** refers to the spread of disease through transportation or social networks. As with the other types of diffusion, network diffusion reflects the geographical and social structuring of human interactions. The roles of social and spatial interactions in network diffusion vary among particular diseases and among geographical and social settings. Analyzing spatial clustering of cholera in Bangladesh, Giebultowicz, Ali, Yunus, and Emch (2011) found that geographic clustering was much more prominent than clustering within social networks. Their findings suggest that in this context, cholera diffused mainly through local environments rather than through interpersonal, social network-based interactions.

Centuries ago, most diseases spread primarily through contagious diffusion; today, however, patterns of disease spread often show a mix of hierarchical, network, and contagious diffusion (Figure 7.3). Mixed patterns are clearly evident in the spread of diseases such as measles, influenza, and HIV/AIDS, both in the United States and worldwide (Gould, 1995; Cliff & Haggett, 2004). GIS-based tools for visualizing and modeling spatial diffusion are discussed later in this chapter.

Mapping Case Distributions

Where are infectious diseases most prevalent? Which populations and geographic areas are most in need of treatment and prevention programs? These questions motivate much geographical analysis of diseases that spread directly from person to person. These analyses require data on case locations that come from some type of surveillance system as discussed in Chapter 3. Before mapping case distributions, we need to consider how cases in the database were identified. Changes in case definition over time can affect the stage during an illness when disease is diagnosed and the number of individuals eligible for a diagnosis. The definition for AIDS originally developed in 1982 was revised in 1985 and 1987 (Chang, Katz, & Hernandez, 1992). The 1992 definition, which became effective January 1993, increased the number of AIDS cases reported significantly, as was expected (Centers for Disease Control and Prevention, 2007a). The new definition brought in many women with AIDS who had been excluded under earlier definitions that failed to include “women’s diseases,” like pelvic inflammatory disease, as opportunistic infections. These changes underline the need to make case definition explicit in GIS databases and metadata.

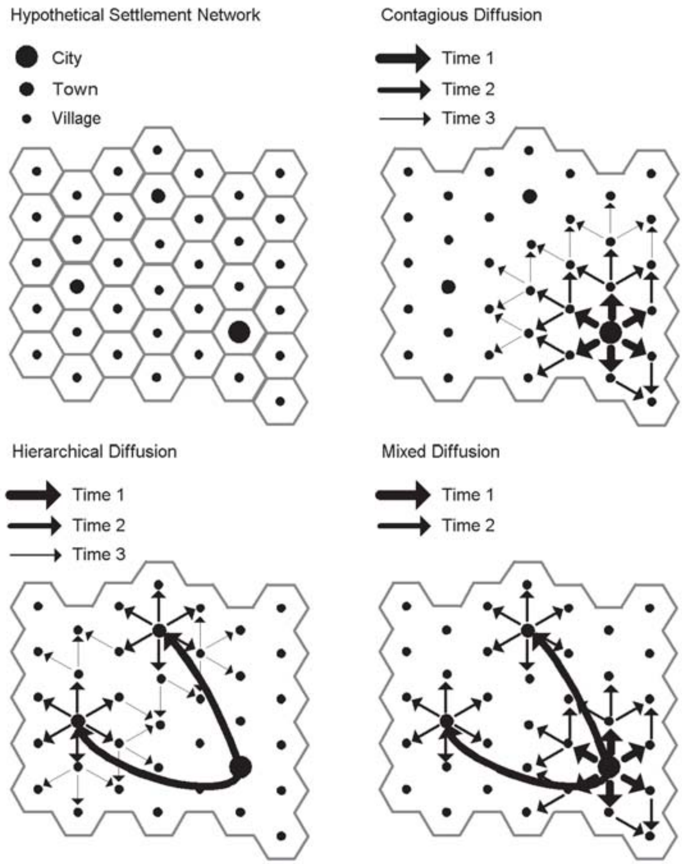


FIGURE 7.3. Spatial diffusion patterns.

A variety of methods, including those discussed in Chapters 4 and 5, can be used to create maps of case distributions and rates for infectious diseases. One example comes from a study of tuberculosis (TB) in Cologne, Germany. TB is a serious contagious disease that spreads via airborne transmission through close personal contact. Effective drug treatments are available for most types of TB. Concern about the rising incidence of TB in some parts of Germany led to an analysis of the spatial clustering of cases using GIS (Kistemann, Munzinger, & Dangendorf, 2002). Tuberculosis cases were first identified through a retrospective surveillance system and then geocoded using an address-matching procedure. The incidence of TB by subdistricts (small administrative areas) was computed by overlaying the geocoded cases on a subdistrict map and summing cases by area. Age-standardized incidence rates were computed by subdistrict. The map of incidence rates revealed clustering of high TB districts in the city center, especially in areas of economic disadvantage and immigrant concentra-

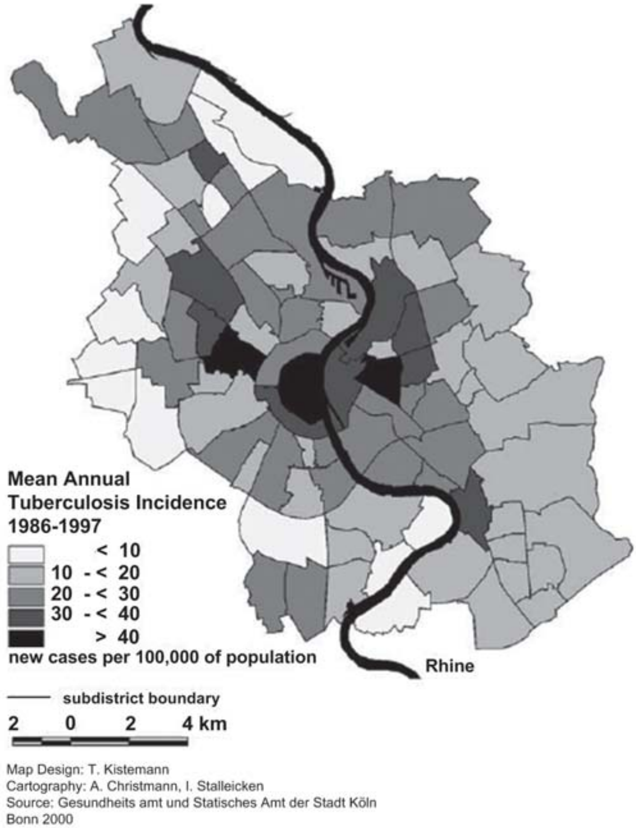


FIGURE 7.4. Average annual incidence rate for tuberculosis by subdistrict in Cologne, Germany, 1986–1997. Reprinted from *Social Science and Medicine, 55*(1), Kistemann, T., Munzinger, A., & Dangendorf, F. Spatial patterns of tuberculosis incidence in Cologne (Germany). 7-19, (2002), with permission from Elsevier.

tion, suggesting the need for more effective treatment and outreach programs in those areas (Figure 7.4).

GIS and Disease Surveillance

Infectious disease surveillance is a critical task for public health departments, and surveillance increasingly involves tracking and visualizing disease reports and incidence data through space and time. GIS are being used in novel ways to support these surveillance efforts. Many geospatial surveillance systems rely on web-based tools and components for disease reporting, data integration, and geovisualization. One of the most extensive, ProMed, is an e-mail and Internet-based reporting system for collecting and disseminating information about disease outbreaks across the globe (International Society for Infectious Diseases, 2011). Reports are posted on the organization's website and disseminated by e-mail to a large network of subscribers. The website includes a mapping tool for creating "pin-maps" of outbreak locations.

There are also national and regional surveillance systems that enable more complex kinds of geovisualization. EpiScanGIS is an online system for meningococcal disease surveillance in Germany (Reinhardt et al., 2008). The system incorporates demographic and residential data about the patient and genetic information about the disease agent. Linked to this detailed database are a series of tools for mapping and spatial analysis, including the SaTScan™ module for detecting space–time clusters that was discussed in Chapter 5. EpiScanGIS was used to create the map sequences of meningitis clusters in Germany shown in Figure 5.16.

With surveillance data for communicable diseases in place, public health professionals can use GIS to develop more effective immunization programs. During a measles epidemic in Auckland, New Zealand, public health analysts displayed measles surveillance data on GIS maps in order to direct epidemic control efforts (Jones, Bloomfield, Rainger, & Taylor, 1998). Intensive vaccination campaigns were targeted to neighborhoods where the incidence of measles was highest. In the final phase of the epidemic, mobile vaccinators were sent to streets where new cases of measles were occurring.

Although most infectious disease surveillance data include the residential address as a geographical identifier, for some diseases transmission can occur outside the home. In many countries, the spread of HIV is linked to long-distance truck drivers and migrant workers who spend much time away from home (Ferguson & Morris, 2007). Their risks of acquiring and transmitting HIV extend over a far-flung network in which the home is just a single node. A study of crack cocaine in North Carolina found high rates clustered along the main interstate highway (Cook, Royce, Thomas, & Hanusa, 1999). For transmission processes like these, a map of case locations geocoded to residence, while depicting where infected people live, only partially represents the geography of disease risk.

In preparing maps of case locations, analysts must also be sensitive to reporting and sampling bias. Surveillance data come from physicians, laboratories, and health care facilities that, though mandated to report, may choose not to do so. Significant underreporting exists for diseases that carry social stigma, like STDs. The implications for mapping depend on the geographical distribution of reporting bias. If underreporting is uniform over space, then rates will be low across the board. However, research on STDs shows strong class differences in underreporting, since high-income people "are more likely to visit a private physician and private physicians are more likely than public clinics not to report an infection" (Thomas & Tucker, 1996, p. S139). This bias leads to lower rates of reporting in high-income areas, and it exaggerates the apparent degree of clustering in low-income areas. The result is an uneven geographical distribution of reporting bias and an uneven pattern of error in case maps.

Mapping Variability in Disease Agents

An important feature of infectious diseases is that the agent is a living organism that evolves over time to enhance its reproductive success. A single agent can have a multiplicity of genetic forms, and these evolve in response to changes in the environment, host, and medical treatments. The best-known example of this is the Influenza A virus whose rapid mutation leads to the emergence of new epidemic strains on an almost yearly basis (Morens, Folkers, & Fauci, 2004). The agents for tuberculosis, malaria, staph, and strep have evolved drug-resistant forms that are resistant to conventional treatments. Mapping the genetic diversity of infectious agents is an increasingly important application area for GIS, reflecting the growing ties between the genetics and GIS research communities (Sloan et al., 2009).

Where are new and resistant strains emerging, and why are they emerging in those areas? Many public health agencies are using GIS to track the changing geographic distribution of infectious agents and their evolution over time. A study of tuberculosis in Capetown, South Africa, combined DNA fingerprinting with GIS to examine the geographical distributions of TB strains (Richardson et al., 2002). Clusters of tuberculosis isolates, each representing a particular strain, were identified based on similarity of genetic sequences; then these strains were mapped and analyzed in GIS. Although most strains were geographically dispersed, some small focal areas of clustering were observed. Mapping of TB strains and human contact networks suggested the presence of an endemic strain in the region and continuing spread of that strain within and beyond the study community through direct contact.

Spatial and temporal variability in disease strains can be analyzed to help in understanding the processes by which infectious diseases evolve and how outbreaks emerge. Carrel, Emch, Jobe, Moody, and Wan (2010) used GIS to construct matrices describing the spatial, temporal, and genetic similarity of strains of H5N1 highly pathogenic avian influenza viruses in Vietnam. Correlation tests showed a strong association between geographic distances among strains and their genetic distances, a finding that highlights the role of geographic location in viral evolution. Viruses located nearby in space were more similar genetically than those located farther apart. The authors also observed a critical distance threshold of 1,100 kilometers, roughly the distance between Hanoi and Ho Chi Minh City, beyond which genetic similarity greatly decreased, suggesting distinct zones of viral evolution in the northern and southern regions of the country.

The social and environmental reasons why new strains emerge can also be explored through GIS. An analysis by van Eldere, Mera, Miller, Poupard, and Amrine-Madsen (2007) examined whether drug-resistant *Streptococcus pneumoniae* was more likely to emerge in response to high levels of antimicrobial consumption in Belgium. Cases of resistant and nonresistant *S. pneumoniae* were geocoded to postal codes, and rates of multidrug resistance by postal code were related to patient and area characteristics using hierarchical modeling methods.

Results showed a positive association between antimicrobial consumption and the proportion of drug-resistant S. *pneumoniae* by postal code. Though ecological in nature, the findings suggest the links between high usage of antimicrobials and the emergence of drug resistance.

Identifying Core Areas

A special application of infectious disease mapping is the analysis of core groups and core areas. Studies of sexually transmitted diseases have drawn attention to the clustering of cases in "core" population groups. A **core group** is a "geographically and socially defined sexual network" (Thomas & Tucker, 1996, p. S134) within which the rate of infection is disproportionately high. As concentrated, high-risk populations, core groups have great importance in efforts to control and prevent STD infection (Bernstein et al., 2004). ***Core areas*** are the geographical analogues of core groups. They are areas in which STD incidence and transmission are unusually high. For diseases like gonorrhea that have an effective cure, the high transmission density in core areas maintains infection and creates a reservoir for epidemic spread (Zenilman, Ellish, Fresia, & Glass, 1999). As with core groups, core areas are critically important for planning and targeting STD treatment and prevention efforts.

GIS can be used in identifying core areas and analyzing the patterns of disease transmission within core groups. This approach was employed in a GIS to analyze core areas for gonorrhea in Baltimore, a city with one of the highest incidence rates of STDs in the United States (Becker, Glass, Braithwaite, & Zenilman, 1998). The GIS worked with disease surveillance data for gonorrhea. Residential addresses of persons diagnosed with gonorrhea were geocoded and assigned to their corresponding census tract. Using 1990 census data to estimate denominator populations, incidence rates were computed by tract. The rates were arranged in rank order, and core areas were defined as the 13 tracts in the upper quartile of the distribution. Core areas contained 15.5% of gonorrhea cases and just 6.5% of population. Incidence rates in the core tracts were more than double the citywide rate.

Core areas can also be identified using the spatial clustering methods discussed in Chapter 5. To delineate core areas for gonorrhea in Baltimore, researchers applied the SaTScan method (Jennings, Curriero, Celentano, & Ellen, 2005). Gonorrhea cases were geocoded to census block groups, and SaTScan was used to identify places where the number of cases was high after adjusting for at-risk population. Four statistically significant clusters emerged (Figure 7.5), and these core areas accounted for 55% of gonorrhea cases over a 6-year time period. A similar approach was adopted by Scribner, Johnson, et al. (2008), who used LISA methods in identifying core areas for HIV in New Orleans.

Successful mapping of core areas requires an appropriate definition of core incidence. Some researchers have used counts of cases in identifying core areas to pinpoint areas of highest incidence (Rothenberg, 1983). When counts are used, however, the core areas that emerge may be areas with large populations, but

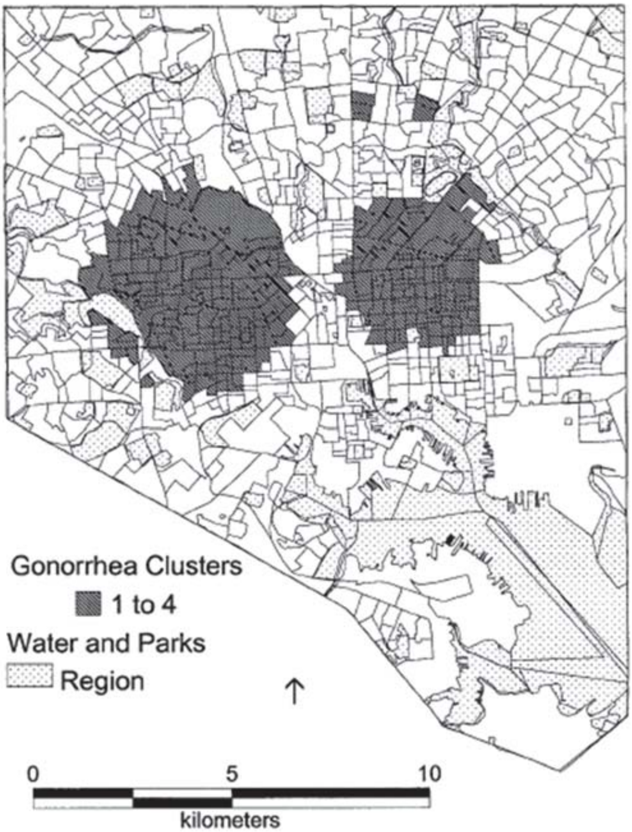


FIGURE 7.5. Core areas for gonorrhea in Baltimore, 1994–1999. Core areas were identified as clusters of census block groups in which there was a significantly elevated risk of gonorrhea based on the SaTScan spatial clustering method. From Jennings, Curriero, Celentano, and Ellen (2005). Copyright 2005 by Oxford University Press. Reprinted by permission.

low disease rates. Alternatively, one can use disease rates in defining core areas to focus attention on places where the risk of infection is high. The disadvantage of emphasizing rates is that high-rate areas may contain relatively few cases. In the Baltimore study by Becker et al. (1998), which used rates instead of counts, focusing public health resources in core areas would reach just 15.5% of reported gonorrhea cases. In some situations neither rates nor counts may be highly effective in identifying areas of sustained transmission. Bernstein et al. (2004) argue that ***repeat infections*** (people who have been infected and reinfected) are most critical for core area transmission. Places containing a high concentration of repeat infections represent a continuing source of infection and thus are important for targeting public health interventions.

Core areas are critical to disease transmission. Their high rates of transmission are thought to sustain diseases during nonepidemic periods and act as reservoirs for infection to areas outside (Rothenberg, 1983). Given detailed data, the geography of STD transmission patterns inside and outside of core areas can be explored to describe the densities, locations, and distances between sexual partners. Such issues were examined in a study of sexual partnerships in Baltimore (Zenilman et al., 1999). The research involved geocoding residential addresses of sexual partners, computing the Euclidean distance between partners, and comparing those distances with distances to randomly chosen residential locations. Residents of core areas lived much closer to their sexual partners than did persons living outside the core, and both distances were significantly less than random. Sexual partnerships were highly localized, especially in core areas, suggesting the potential benefits of geographically targeted prevention programs.

Identifying and mapping core areas raises several important broader issues. Displaying the areas on maps may unfairly stigmatize the residents of core areas and lead to place-based redlining and discrimination. The areas can become labeled as places where disease risk is high, places to be avoided by businesses, service providers, and individuals. Part of the problem stems from defining places as either in- or outside the core when in fact disease incidence is not so polarized. Typically the vast majority of cases are located outside core areas, and the vast majority of core area residents do not have the disease. In the Baltimore example (Becker et al., 1998), almost 85% of people with gonorrhea lived outside core neighborhoods, and well over 90% of core area residents did not have gonorrhea. These patterns are typical of STDs and should be emphasized when presenting maps of core areas.

Core area mapping projects also need to recognize the social and economic conditions that underpin high prevalence. The higher risk of disease in core areas is often rooted in patterns of social deprivation, including high unemployment, low incomes, deteriorated housing, and poor access to health care. The withdrawal of fire services, housing, and jobs from New York City's inner-city neighborhoods in the 1970s triggered the emergence of core areas that then became nodes for the outward spread of HIV/AIDS, tuberculosis, and other communicable diseases (Wallace & Wallace, 1998). As the work of Wallace and Wallace vividly portrays, core areas are not zones of "deviant" behavior, but rather products of a multitude of social and political inequalities and misguided social policies. From this perspective, developing effective intervention programs requires an understanding of the social and political ecology of core areas and how people's access to work, housing, and services in core communities structures their health.

Mapping the Ecology of Risk

The ties between local place environments and the behaviors and interactions that sustain infectious disease transmission are increasingly being investigated with the help of GIS. GIS are useful in linking data on environmental conditions and quality as discussed in Chapter 6, along with characteristics of the built environment, social and economic characteristics of area populations, and information on social institutions and interactions. The specific kinds of social and environmental indicators to be included depend on biological and social characteristics of the disease and how it is transmitted.

Cryptosporidiosis is a gastrointestinal disease transmitted by direct contact and contact with contaminated water. GIS was used to better understand how and why the risk of cryptosporidiosis in England and Wales varies geographically and the role of environmental factors (Lake et al., 2007). Cases of cryptosporidiosis were geocoded to residential postcodes (small geographic zones coded for postal delivery) along with a similar number of randomly chosen controls. Using GIS, researchers extracted data on a variety of environmental indicators for each case and control postcode including drinking water supply, socioeconomic indicators, rural-urban location, and amount of cryptosporidium applied to nearby agricultural land through animal manures. Two species of human cryptosporidium, *C. homninis and C. parvum*, were analyzed, and statistical analysis showed important differences in environmental associations. This study is one of many illustrating the role of GIS in linking and managing diverse types of environmental and epidemiological information for infectious disease investigations.

Risk behaviors such as unprotected sex and drug and alcohol use also vary from place to place. In understanding these risk behaviors, researchers are increasingly focusing on the ***social production of risk***—how social and environmental characteristics such as local environmental quality; access to social resources, jobs and services; density of social networks and support systems; and broader social, economic, and political inequalities influence risk behaviors for communicable diseases (Galea, Nandi, & Vlahov, 2004; Rhodes, Singer, Bourgois, Friedman, & Strathdee, 2005). Understanding people's responses to and interactions with these local place environments is also critically important (McLafferty, 2009).

GIS can assist in documenting where risk behaviors occur and teasing out their roles in disease transmission. Spread of HIV/AIDS among women is an important but relatively neglected issue in the United States. To better understand how and why women acquire HIV infection, researchers in San Francisco utilized sentinel surveillance data for women booked in the county jail systems. Women were tested for STDs including HIV and were queried about a variety of risk behaviors (Kim et al., 2008). Women's residential locations were geocoded and mapped using GIS. A strong association between injection drug use and HIV was observed. A map overlay of the density of women requiring injection drug use services with the locations of two women-only syringe exchange programs (Figure 7.6) revealed that the programs may not be well located to serve their target population.

Substance use is an important risk behavior for several infectious diseases, and it is a significant health concern in its own right. Using GIS, Mennis and

FIGURE 7.6. Density of women needing services for injection drug use in San Francisco and locations of women-only syringe exchange programs. From Kim et al. (2008). Copyright 2008 by Springer. Reprinted by permission.

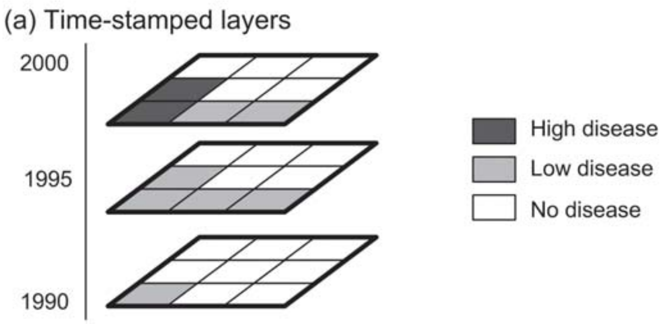
Mason (2011) demonstrated that substance use among a sample of adolescents living in west Philadelphia was associated with characteristics of the spaces in which adolescents engaged in everyday activities (Mennis & Mason, 2011). The adolescents were asked to identify locations in their daily life that they perceived as "safe" and "risky," and locations were geocoded. Locations of other potentially important neighborhood features such as bars, liquor stores, and check-cashing stores were also recorded. Distances were measured from the respondent's home and other important sites to these facilities. Statistical analysis showed that characteristics of adolescents' perceived "risky" places enhanced the risk of substance use among some age–gender groups.

Social mapping—mapping by participants of their own neighborhoods and everyday spaces—is an important tool for examining the relationships between local environments and risk behaviors. It is particularly valuable for understanding the geography of risk for disadvantaged and "hidden" populations such as injection drug users (Singer et al., 2000). Although most social mapping has involved paper maps, use of GIS to represent information from participants about their own everyday spaces has increased. A study of HIV/AIDS risk among female sex workers in Vancouver used social mapping to explore how women's access to health services was influenced by their perceptions of neighborhood risk and violence (Shannon et al., 2007). Women were asked to identify on a GIS map: places where they live and work; locations where they access health/support services and obtain syringes; and places they avoid because of violence and fear of local policing. GIS overlays of the information revealed a strong geographic correspondence between health service availability and violence/policing. Places where health services were located were also places women avoided due to violence and policing. Women's access to health care and syringe exchange services was reduced by their perceptions of fear and violence in the surrounding neighborhood. Incorporating data on people's attitudes and perceptions in GIS is an important contribution and one that is relevant in many GIS and public health investigations.

Analyzing Temporal and Geographic Trends in Disease Outbreaks

An important issue in monitoring communicable diseases and planning interventions is to understand the patterns of spread through space and time. A public health department might want to know when the peak outbreaks of disease typically occur and how those peaks move from place to place. In addition, the health department may want to know how the incidence of disease in different places has changed over time and to forecast future patterns of spread. These kinds of questions involve analyzing spatial and temporal information simultaneously. Time adds a "third dimension" that does not always fit comfortably in the two-dimensional world of GIS. Although challenges remain, GIS capabilities for managing and modeling spatiotemporal data have improved greatly in the past decade.

GIS typically utilize ***time-stamp*** approaches in handling temporal data (Yuan, 2009). In the ***snapshot approach***, a time-stamp is applied to each data layer (Figure 7.7a). The GIS contains layers of information for different points in time, for example, AIDS incidence rates by county for 1990, 1995, and 2000. Different geographic objects can be depicted in each snapshot layer. Snapshot data can be analyzed using map sequences and animation as discussed in subsequent sections. The ***time-stamped tuples*** approach assigns time-stamps to the rows of a geographic matrix. For example, in Table 1.1, spatial data on water mains include a time-stamp indicating the date of construction. The spatial objects remain constant, but their attributes vary over time. In the ***time–space composite model,*** GIS data comprise combinations of spatial and temporal attributes, with rows representing spatial objects and columns representing points in time (Figure 7.7b). A value in the GIS data matrix indicates the characteristic of a location at a particular point in time. The ***spatiotemporal object model*** assigns attributes of a spatial object to time-stamped individual values (Figure 7.7c). The elements of this model are ***space–time atoms*** that represent unique combinations of spatial and temporal attributes. More recent approaches involve analysis of events, actions, and processes dynamically in GIS via methods like agent-based modeling. The methods for visualizing and forecasting the spread of communicable diseases discussed in this section grow out of these alternative approaches to handling spatiotemporal information.



Area ID	1990	1995	2000
1	Low disease	Low disease	High disease
2	No disease	Low disease	Low disease
3	No disease	Low disease	Low disease
4	No disease	Low disease	High disease
5	No disease	No disease	No disease
6	No disease	No disease	No disease
7	No disease	No disease	No disease
8	No disease	No disease	No disease
9	No disease	No disease	No disease

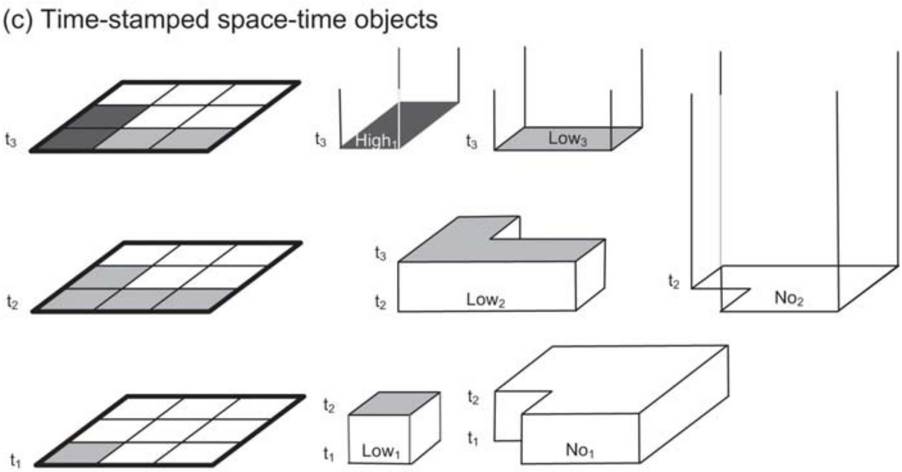


FIGURE 7.7. Methods for incorporating time into GIS databases. In Figure 7.7a (snapshot model), individual data layers showing areas of no, low, and high disease are "time-stamped." In this figure, the geographic unit boundaries (for example, census tracts) do not change over time, but they could. In Figure 7.7b, the GIS database is organized so that rows represent geographic areas and columns represent time periods or time-stamped attributes. An attribute value is recorded for each space–time composite. If the geographic boundaries of the units did change over time, the rows would represent the geographic units resulting from the intersection of all geographic units over time. In Figure 7.7c, the data on three space–time objects showing areas of no, low, and high disease are decomposed into six space–time atoms (two space-time atoms of no disease, three space–time atoms of low disease, and one space–time atom of high disease). Adapted from Yuan (2009). Copyright 2009 by Elsevier. Reprinted by permission.

Mapping Peak Incidence

Many communicable diseases like influenza and the common cold show distinct seasonal patterns of occurrence. For these types of diseases, the patterns of spread reflect when and where the disease was introduced, weather conditions, and the interactions and movements of infected and susceptible populations. One way to visualize temporal and spatial spread is to prepare a map showing for each location the date (month or week) of ***peak disease incidence***—that is, the week or month when the largest number of cases occurred. Thus, locations are time-stamped with the date of peak incidence.

The timing of peak activity for rotavirus infection in the United States was mapped in this way (Torok et al., 1997). Rotavirus is a major cause of gastroenteritis among children and infants. In the United States, the timing of peak rotavirus infection shows strong seasonal and geographical trends, with epidemics beginning in the Southwest in the fall and spreading north and east. To display these trends, data from the National Respiratory and Enteric Virus Surveillance system gathered from 69 laboratories in 42 states were analyzed. The weeks of the year were represented by numbers (1 to 52), and for each laboratory, the week of peak infection was recorded. Kriging, as discussed in Chapter 6, was used to display the geographical distribution of peak times (Figure 7.8). The earliest peaks occurred in the Southwest in the late fall, and the peaks spread north and west through the winter months. In the eastern and northern United States, the epidemic reached its maximum in late spring (Bosley, 1997).

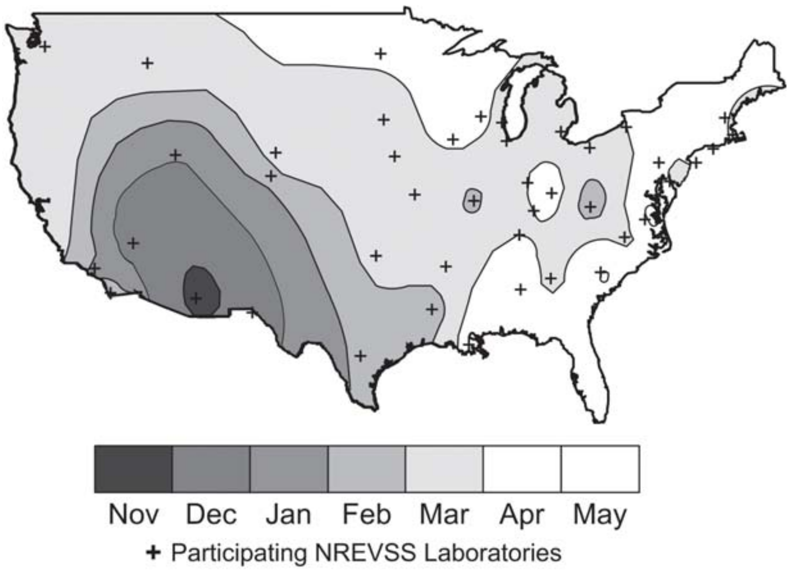


FIGURE 7.8. A kriged map of temporal peaks in rotavirus infection in the United States. From Bosley (1997).

Map Sequences

Map sequences have been widely used in studying the spread of infectious disease. A ***map sequence*** is a series of maps, displayed side-by-side, which show the disease distribution at different points in time. Each map represents a cross section or slice through time of the geographical pattern of disease, corresponding to the snapshot model of space–time data. In their research on the diffusion of measles, Cliff and Haggett (1988) created map sequences of measles in Iceland during various epidemic periods (Figure 7.9). Patterns of contagious and hierar-

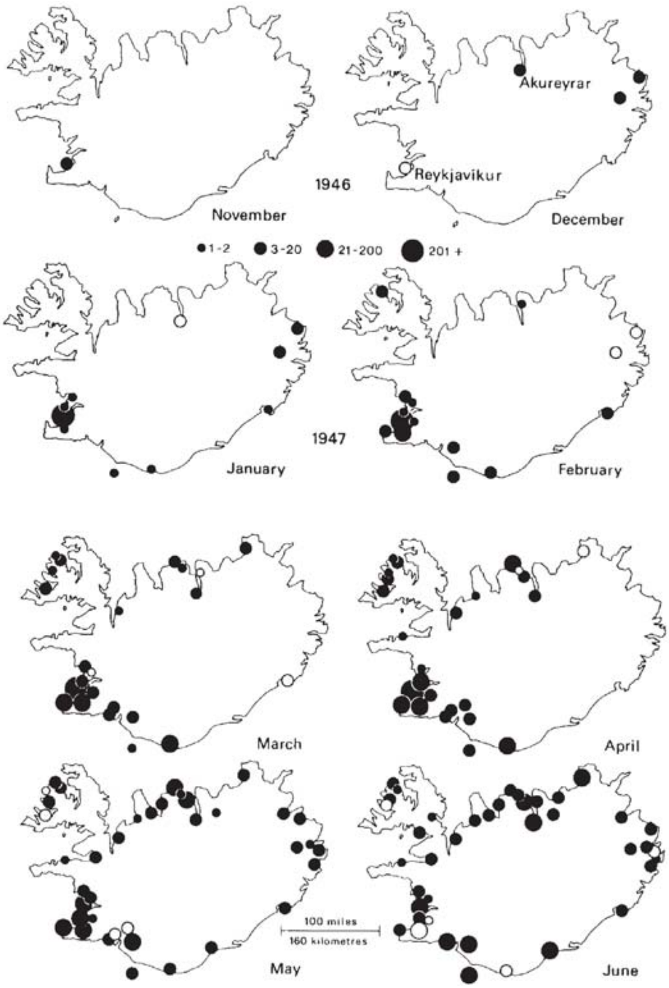


FIGURE 7.9. A sequence of measles cases in Iceland by month from November 1946 through June 1947. From Cliff and Haggett (1988). Copyright 1988 by A. Cliff and P. Haggett. Reprinted by permission.

chical diffusion were clearly evident. Contagious diffusion dominated the map sequences for early epidemics when transportation connections were limited. In contrast, map sequences for the 1940s showed increasing evidence of hierarchical spread, reflecting higher rates of mobility and transportation access. More recent examples of map sequences can be found in the public health literature, including, for example, map sequences of shigellosis in Chicago (Jones, Liberatore, Fernandez, & Gerber, 2006) and of heroin use in Scotland (Field & Beale, 2004).

Map sequences can be created for almost any type of spatial health information that varies over time. The maps can display case locations, incidence rates, or rates of change over time. Ancillary features like transportation routes or commuting flows can be added to show the connections between transportation improvements and the spread of disease.

One of the main advantages of map sequences is the ease of comparing geographical patterns at various points in time. All maps are on the same page, allowing the viewer to shift from one map to the other searching for similarities and differences and to focus easily on time periods that hold special interest. At the same time, it may be difficult to discern time trends as the viewer shifts his or her gaze from one image to another. Despite this shortcoming, map sequences offer a useful method for analyzing disease diffusion.

Animated Maps

Animated maps are maps characterized by continuous change while the map is viewed (Slocum et al., 2009). Time-snapshot maps are displayed dynamically, in sequence, forming a constantly changing image or animation. The field of animation has advanced rapidly in recent years, stimulated by developments in computer hardware and graphics software. These advances are fueling changes in mapmaking as cartographers gain access to one of the first effective tools for representing continuous change through space and time.

Use of animated maps to show the spread of infectious diseases has increased greatly, and many animated map sequences are available on the web. A recent example is a "map movie" depicting the spread of H1N1 Swine Influenza by county in the United States beginning in late April 2009 (Macurek, 2010). The early maps are very sparse, depicting only small clusters of cases, primarily in large cities. Over time, the map "fills in" as the disease spreads to new locations through processes of contagious and hierarchical diffusion. At the peak of the epidemic there are large geographical concentrations of cases, not only in major cities but also along the United States–Mexico border, reflecting the human social and spatial interactions that propel influenza transmission.

Designing map animations involves several considerations beyond the traditional visual variables of static mapping, including duration, rate of change, and order (DiBiasi, MacEachern, Krygier, & Reeves, 1992). ***Duration*** refers to the length of time each map is in view (Slocum et al., 2009). In map animations with a short duration, each image disappears quickly, producing a smooth but constantly changing animation. Lengthening the duration gives the viewer more time to study each map, but the animation appears choppy. Some map animations allow users to control the duration, providing a tool that is customized to individual differences in perception and cognition (Slocum, Yoder, Kessler, & Sluter, 2000).

Rate of change describes the smoothness or variability of the animated map sequence. It is computed as the amount of change between maps divided by the duration. If the positions or attributes of features on the map change substantially during the animation, the animation has a high rate of change. One can reduce rate of change by increasing the duration of each map and thus smoothing the transition from map to map. Similarly, reducing the amount of change between maps gives a lower rate of change and a smoother animation. One way to accomplish this is by using overlapping time intervals for the maps rather than discrete intervals. For example, the first map might show disease incidence for weeks 1–4, the second for weeks 2–6, the third for weeks 4–8, and so on. The two-week overlap means that some of the data on one map also appears on the next map in the animated sequence.

Order defines the sequence of maps in the animation. Map animations typically rely on chronological order—a sensible choice for representing change over time. Using criteria other than time for ordering maps makes animations difficult to interpret. The viewer has to pay close attention to decipher the timing and sequencing of events.

Other visual variables mentioned in Chapters 2 and 4 such as hue, size, and shape can also be used effectively in map animation. Contrasting colors, changes in symbol size, "decay" images, and flashing symbols attract the eye to important data events (Harrower, 2003).

Animation is an effective way of displaying the spread of infectious diseases through time and space. The smoothly changing patterns of contagious and hierarchical diffusion and the intensity of epidemic spread are all apparent on map animations. As in the H1N1 Swine Influenza example, animated maps reveal the expanding contours of disease emergence and the rapidity of spread. Animations, however, are primarily visual tools. They clearly show regular patterns, but if events move or vary in intensity unpredictably over time, the animation will be difficult to comprehend. Viewers also have trouble analyzing information on animated maps. The images move by quickly and are difficult to compare. Although one can often stop the animation to focus on a particularly interesting map, the contrast with other maps may not be apparent. New animated mapping tools address many of these concerns by incorporating user needs and perceptual capabilities in tool design (Harrower, 2003).

Forecasting Spatial Diffusion of Communicable Diseases

Our discussion of animation and map sequences has primarily emphasized retrospective mapping; however, prospective mapping—forecasting future patterns of disease spread—is also of great importance as public health departments prepare for future epidemics. GIS can be used to implement spatiotemporal models for predicting patterns of disease spread and to manage the complex space–time data sets that underpin such predictions.

Efforts are under way to link epidemic models with spatial diffusion models to predict the movement of infectious diseases through time and space (Cliff & Haggett, 2006). Many of these approaches rely on graphs (networks) to represent the spatial interactions between places. The ***Spatiotemporal Epidemiological Modeller (STEM)*** is a graph-based system that combines an SIR model of disease transmission with a network representation of contacts/flows among places (Ford, Kaufman, & Eiron, 2006). Interactions among places can be expressed in different ways, based on proximity or contiguity, or by using data on flows of people from place to place. STEM is available for download (The Eclipse Foundation, 2011).

Epigrass is a similar graph-based, open-source platform for simulating the spread of disease through complex networks (Coelho, Cruz, & Codeço, 2008). In Epigrass, epidemic models of disease spread within and among cities interact with network models of the flows of infected and susceptible populations. Nodes (cities) and edges (links between cities) are dynamically updated to forecast the changing incidence of disease. Loosely coupled with GIS, Epigrass uses GIS data as input to the epidemic simulation and then outputs the results in a variety of formats including map animations (Figure 7.10).

Agent-Based Models of Disease Spread

An emerging trend in epidemic modeling is to focus on individuals rather than larger populations or nodes through the use of agent-based modeling. ***Agent-based models (ABMs)*** simulate the behaviors and interactions of individuals to generate insights about populations and communities. Each individual is represented by an ***agent***, defined as "a piece of computer code capable of autonomous, goal-directed actions" (O'Sullivan, 2008, p. 541). In agent-based models of disease spread, the agents are individuals who may be infected, susceptible, exposed, or recovered. Agents move from place to place and come in contact with other agents, providing opportunities for disease transmission. An early model developed by Bian (2004) considered people's interactions between and within two levels, homes and workplaces (Figure 7.11). Biological characteristics of the disease such as ***infectivity*** (the ability of the disease to cause infection) and ***virulence*** (the severity of the disease after infection occurs) are incorporated in the simulation (Yang & Atkinson, 2008). Individual differences in immune response can also be modeled effectively in the agent-based approach. Like graph models, typically ABMs are loosely coupled with GIS, relying on GIS primarily for data input and updating.

There are many good examples of agent-based epidemic modeling, and the ability of models to represent complex social and spatial interactions in diverse

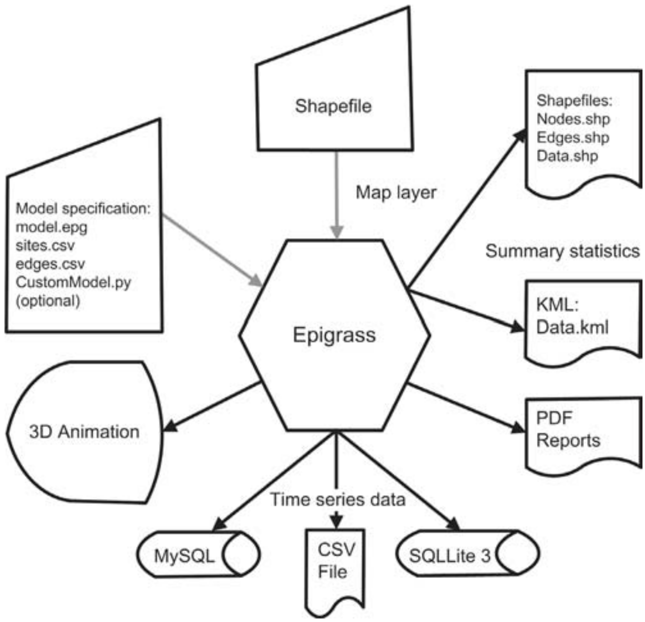


FIGURE 7.10. The Epigrass time–space epidemic simulation model imports geospatial data and exports simulation results in diverse formats including map animations. From Coelho, Cruz, & Codeço, 2008. Originally published by BioMed Central in *Source Code for Biology and Medicine*, Open Access.

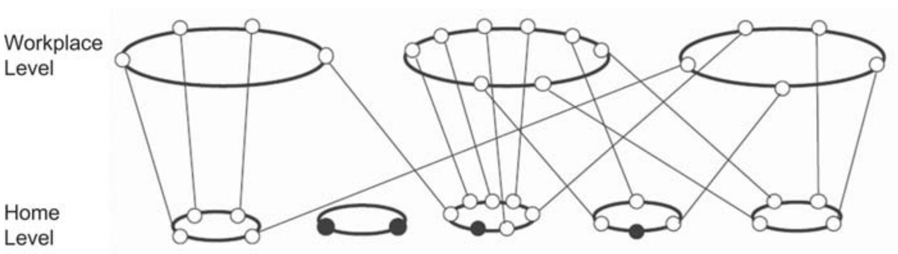


FIGURE 7.11. Schematic diagram of a two-layer interaction structure used in an agent-based model. The two layers are homes and workplaces. The filled circles represent people whose interactions occur only at home. The open circles represent people who interact at home and at work. The straight lines connect the home and work locations of these people. People's movements within and between homes and workplaces create opportunities for disease transmission. From Bian (2004). Copyright 2004 by Pion, Ltd. Reprinted by permission.

geographic settings has increased greatly. Yang and Atkinson (2008) devised an ABM to predict the spread of an airborne disease through a university campus. The model incorporates individual time–space activity patterns in simulating transmission dynamics. Lee, Bedford, Roberts, and Carley (2008) designed an agent-based simulation of an influenza epidemic in Norfolk, Virginia, that includes detailed geospatial data on schools, workplaces, social and recreational facilities, as well as models of social networks based on survey data. Public health responses to the epidemic such as school and workplace closings are also directly modeled. Although it is difficult to assess the accuracy and validity of these models, their effort to simulate real-world behaviors in complex environments is noteworthy.

Great progress has been made in efforts to simulate the spread of communicable diseases through space and time. Harnessing vast GIS data sets and advances in computing speed and power, researchers have moved away from the more aggregate, network approaches based in spatial diffusion modeling to the individual approaches typical of agent-based modeling. As this transition occurs, it is important to ask: What is the value of the added complexity of ABMs versus aggregate models for accurately predicting disease spread (Hupert, Xiong, & Mushlin, 2008)? The predictive power of ABMs rests on their ability to accurately model human social and spatial interactions at the individual scale, a daunting task. At the same time, aggregate models often rely on critical assumptions about population flows, interpersonal contacts, and disease transmissibility. Validation studies are sorely needed, and it may be that some kinds of models perform better for specific diseases and specific geographic contexts. We are also likely to see increasing reliance on real-time data on human movements and interactions gleaned from mobile, GPS-enabled devices.

Planning Public Health Interventions

In addition to modeling and analyzing communicable diseases, GIS can play a key role in planning policies to prevent and limit disease spread. Many such policies are distinctly geographical. They involve ***spatial targeting***—directing interventions to the places and people most in need—and geographically tailoring policies to reflect local circumstances, environments, and populations. Research shows that spatially targeted health policies are often much more effective than those that cover an entire population or region in a uniform manner (Keeling & White, 2011). There is also great value in geographically tailoring public health policies to reflect local environmental conditions and the socioeconomic and cultural circumstances of local populations. Knowing what exists where is critically important.

Policies to limit the spread of communicable diseases can be grouped into four broad categories—behavioral, environmental, medical, and mobility—as listed in Table 7.1 (McLafferty, 2010). ***Behavioral policies*** include efforts to reduce behaviors that place people at risk of infection. For HIV/AIDS, safe

TABLE 7.1. Strategies for Controlling Communicable Diseases

Strategy	Examples
Medical	Vaccination Treatment of infecteds
Environmental	Water supply Sanitation Housing improvements Neighborhood improvements
Mobility/contact	Social distancing Travel restrictions Isolation Quarantine
Behavioral	Education—for example, safe sex, clean needles Incentives to change behavior Legal restrictions on behavior

Note. Adapted from McLafferty (2010). Copyright 2010 by V. H. Winston & Son, Inc. Adpated by permission.

sex education, provision of free condoms, and syringe exchange programs are examples of behavioral policies. When risk behaviors vary from place to place, spatial targeting of policies and programs to places where people are most likely to engage in such behaviors can be effective in curbing disease transmission. A study in St. Petersburg, Russia, used GIS to map locations of HIV/AIDS risk behaviors among injection drug users and recommended that harm reduction efforts be focused in areas where behaviors were spatially clustered (Heimer, Barbour, Shaboltas, Hoffman, & Kozlov, 2008).

Environmental policies involve modifying natural and built environmental characteristics to reduce the spread of infections. Environmental modifications, including provision of clean drinking water and sewage systems, have long been and continue to be extraordinarily important in controlling communicable diseases. A team of researchers in Lusaka, Zambia, used GIS to map and analyze the locations of cholera cases in relation to the presence of water drainage networks in the city. GIS was used to determine the number of cholera cases within 500-square-meter grid cells and the corresponding length of drainage networks. Findings showed a strong inverse association between disease incidence and drainage, a correlation that reflected a town planning policy during the colonial period that segregated the city's native population in poorly drained residential areas outside the central city. The maps and spatial overlays highlighted the need for spatially targeted improvements in drainage infrastructure to reduce the risk of future cholera outbreaks (Sasaki, Suzuki, Fujino, Kimura, & Cheelo, 2009).

Environmental modifications may also be used to discourage risk behaviors that are important in infectious disease transmission. As noted earlier, recent studies suggest that certain kinds of place environments, "risky places," are conducive to risk behaviors such as substance use, whereas other types of settings appear to discourage risky behavior (Rhodes et al., 2005; Mennis & Mason, 2011). Modifying such environments by reducing or eliminating the characteristics that promote risk may be an effective way to curb the spread of infection. GIS can contribute to planning and implementing these environmental modifications. For example, GIS can be used to identify potential high-risk locations based on specific combinations of environmental characteristics, and to prioritize locations where environmental changes are most needed. The published literature contains few, if any, examples of these types of GIS applications, although their potential is evident.

The third broad category of public health strategies—***medical policies***—involves the use of biomedical interventions to limit spread of infections. These policies include, for example, treatment of the infected, a strategy that simultaneously reduces morbidity, mortality, and the number of carriers of infection, and vaccination of persons at risk of infection. Mapping is widely used to target testing and treatment programs for diseases such as gonorrhea, HIV/AIDS, and hepatitis C (Trooskin, Hadler, St. Louis, & Navarro, 2005; Jennings et al., 2006).

Vaccination campaigns can benefit in many ways from application of GIS. The history of smallpox eradication shows the importance to vaccination efforts of creating geographically based, systematic disease surveillance systems, and spatially targeting intensive containment vaccination to places and populations at highest risk (Henderson, 1980). Simulation studies confirm the value of spatial targeting. They reveal that during an epidemic, targeting the vaccine to regions that are most affected is efficient and effective in curtailing epidemic spread (Keeling & White, 2011). Although GIS was not a part of the smallpox eradication campaign, the technology can clearly support the rapid, targeted vaccination efforts that form the core of the public health response for many important communicable diseases. For example, GIS can be used to deploy vaccination teams to highly affected areas and to identify suitable locations for storing and administering vaccines (Khan et al., 2010). Efforts to eradicate polio in the Democratic Republic of Congo via an intensive vaccination and treatment campaign used Google Earth® to map the spatial distribution of wild poliovirus cases along the Congo River (Kamadjeu, 2009). Maps were used to dispatch vaccination teams to affected areas, to assist in planning social mobilization efforts, and to identify passage points along the river where mobile populations traveling on boats, canoes, and rafts could be screened and vaccinated.

In addition to targeting vaccination efforts, GIS can support the many research and planning activities that are needed to make vaccination a success. Khan et al. (2010) provide a comprehensive summary of the use of GIS in vaccine trials. The success of a trial rests on detailed knowledge of the population(s) and place(s) within which testing will occur. Population size, characteristics, and disease burden are very important in assessing vaccine efficacy. "GIS can help in investigating the spatial pattern of the disease burden and then any significant heterogeneity in this indicator can be addressed during data analysis" (Khan et al., 2010, p. 307). GIS can also be used during the trial itself to map the spatial locations of people who refuse to participate and to plan locations for health clinics providing essential services to participants. Finally, GIS provide a tool for estimating ecological variables that describe the local environment for each participant (Emch et al., 2006). Statistical adjustments for such variables are needed in assessing the protective efficacy of the candidate vaccine.

The fourth category of public health interventions is ***mobility policies***—interventions that aim to restrict the mobility of infected persons and limit their social interactions. These policies are especially important for diseases that spread rapidly from person to person such as influenza. ***Social distancing*** involves efforts to limit social interactions by closing schools, workplaces, and sport/entertainment events. Social distancing is widely credited with limiting the spread of H1N1 influenza during the 2009 pandemic (Franco-Paredes, Carrasco, & Preciado, 2009). Other mobility policies include travel restrictions such as banning certain people from entering a country and the use of thermal scanners at airports to screen entering visitors. Although studies show that travel restrictions are largely ineffective, they have been widely adopted by some countries in recent pandemics (McLafferty, 2010). The most extreme mobility policies are isolation and quarantine, efforts to separate infected persons from the rest of the population by placing them in a contained and restricted geographic setting.

Examples of the use of GIS for planning and implementing mobility strategies are not presently available in the research literature. However, insofar as a GIS provides a set of methods and technologies for visualizing and modeling people's movements and spatial interactions, it may be valuable in supporting public health agency efforts to develop and implement effective mobility strategies. Flow data from cell phones and other GPS-enabled devices can be used to describe the dynamics of people's space–time interactions and movements. Public health planners can use location–allocation models, discussed in Chapter 10, to identify critical locations where mobility restrictions would be expected to have the maximal impact on limiting disease transmission. Simulation and agent-based models, discussed earlier in this chapter, may be used for similar purposes. Incorporating dynamic understandings of human mobility in policies to limit the spread of communicable diseases poses an important challenge for future GIS research investigations.

Privacy and Confidentiality

Mapping communicable diseases at detailed geographic scales raises significant concerns about privacy and confidentiality. Just as maps of core areas can be used to stigmatize places, maps of case locations reveal personal information that can be used to stigmatize or discriminate against infected people. People's access to insurance or to health care and medical treatment may be denied because of geographical location. Curry (1999) calls this the “power of the visual,” the violation of privacy for individuals or groups that results from creating maps of social and spatial information.

Even if health information is not displayed on a map, the spatial data management capabilities of GIS raise additional privacy concerns. These include the rapid growth of unregulated and potentially inaccurate spatial data, and the ability to create large dossiers on individuals by linking bits of information about people and places (Curry, 1998). GIS are central to these activities as this technology can be used to join data from different sources based on a common geographical location. Now that many devices offer a variety of location-based services, there is concern that service users’ locations can be detected and disclosed. Research has shown that landscape or map information can be used by third parties to estimate a person’s location, even from spatially perturbed data (Ardagna, Cremonini, & Gianini, 2009).

Changing technology has made GIS tools widely available online, increasing the risk that disclosed data on locations can be readily mapped and disseminated. The development of online GIS services described in Chapter 1 means that a database of addresses can be easily geocoded and the lon/lat coordinates can be easily mapped and displayed on the web using KML, even by people who do not have access to GIS software. Software for converting a database in shapefile format to KML format for mapping online can also be downloaded.

How can the needs of researchers and policymakers for geographically detailed health information be reconciled with the important right to privacy? Chapter 3 discusses protections in the collecting and managing of health data. In preparing maps, several strategies can be used for avoiding the most obvious privacy violations. One is to aggregate health data to larger spatial units. Several of the studies mentioned earlier in this chapter began with residential addresses, but grouped the data by census tract for analysis and mapping (Becker et al., 1998). The address information is essentially discarded after assigning the cases to tracts. This is not a perfect solution because some of the tracts have such small populations that their tract totals might reveal personal information. To address this issue, we can omit tracts with small populations from analysis and mapping. In the Baltimore study, only tracts with more than 30 cases of gonorrhea were included in the GIS analysis (Becker et al., 1998).

Data from the 2001 Canadian census were used to model the relationship between geographic area population size and uniqueness for common demographic variables (El Emam, Brown, & AbdelMalik, 2009). The objective of the research was to estimate the minimum population size at which a geographic area’s population is sufficiently large so that no further data aggregation is necessary to present the data without suppression. The models were applied to a database of prescription records from retail and hospital pharmacies provided to commercial research and analysis firms.

Early studies of health using GIS frequently included maps of individual health events represented as points at residential address locations. Later research has demonstrated that, even when there is little detail and the resolution of the map image is low, the point locations can be reengineered or ***reverse geocoded*** to reveal the address (Brownstein, Cassa, & Mandl, 2006; Curtis, Mills, & Leitner, 2006).

It is possible to develop ***geographical masks*** that preserve the security of individual health records while retaining enough location information to make it possible to answer questions that can only be answered with some knowledge of the geography of health events. Both the validity of these masks and their relative security have been examined (Armstrong, Rushton, & Zimmerman, 1999). The development of methods for anonymizing individuals and the impact on detection of spatial clusters has been described (Cassa, Grannis, Overhage, & Mandl, 2006).

The extent to which disclosure of metadata about the mask and multiple releases of masked data affect the confidentiality of the masked data has also been investigated (Zimmerman & Pavlik, 2008). Researchers found that home addresses of spatial locations that were de-identified using a nondeterministic blurring algorithm could be re-identified from multiple anonymized data sets produced from the same set of addresses (Cassa, Wieland, & Mandl, 2008).

In addition to statistical methods of data anonymization, normative modeling techniques have been used (Wieland, Cassa, Mandl, & Berger, 2008). These methods assign health events from their actual locations to one of a set of locations to which health events can be moved. The set of locations can be configured so that no health events will be moved to places where no one resides, for example, the middle of Central Park in New York City. The assignment is made to meet the objective of minimizing the relocation distance needed to assure the desired probability of re-identification. The model was tested using data for New York County. Compared to aggregating data by ZIP Code, the mathematical programming model "moved" cases over much shorter distances. The impact of the method on the ability to detect clusters in the spatial pattern of cases was also evaluated using SaTScan.

None of these strategies deals with issues of stigmatization, data linkage, and profiling, issues that are social and political in nature and not easily amenable to technical solutions. Addressing these issues will require multifaceted and multidisciplinary approaches (Curry, 1999) that are sensitive to the needs and concerns of individuals and communities and that emerge out of a legal, ethical, and political framework that is open and participatory.

Maintaining privacy and confidentiality is one of the most significant challenges in health mapping, especially for nonvectored infectious diseases, but also for a wide array of health issues including mental health and cancer. There are no easy solutions. Any attempt to address privacy and confidentiality involves balancing the competing interests of individuals and groups defined by class, ethnicity, race, or place, with the interests of protecting public health and providing high-quality data for public health research. GIS researchers, through their work on privacy and confidentiality, are engaging with other stakeholders in the political, legal, and social arenas to advance our ability to use individual-level health data for geographical analysis in ways that protect people.

Conclusion

Infectious diseases that spread directly from person to person are significant health concerns nationally and globally. Caused by living organisms that change in response to our efforts to control them, such diseases present a continually shifting challenge to public health. GIS mapping and analysis have an important role in depicting the large disparities among communities in infectious disease burden, in understanding how such diseases emerge and evolve, in charting the dynamics of epidemics through space and time, and in directing interventions to promote health.

CHAPTER 8

Exploring the Ecology of Vector-Borne Diseases

Zoonoses are diseases that are naturally transmissible from vertebrate animals to humans. More than 200 zoonotic diseases have been described, some of them recognized for hundreds of years. Emerging zoonoses such as avian influenza have heightened global public awareness of the importance of diseases linking human and animal populations. Because these diseases often involve domestic or companion animals, there has been a growing emphasis on ***veterinary public health***, improving human health through the application of veterinary science.

A large percentage of zoonotic diseases are vector-borne. Vector-borne infectious disease involves a causative agent—usually some type of microorganism—that is the direct cause of the disease in the host. Vector-borne disease agents may be parasites, bacteria, or viruses. The ***vector*** is a living organism—usually an insect—involved in the transmission of the disease. For some vector-borne diseases, the transmission cycle also involves an ***intermediate host*** organism in which the agent develops or multiplies and a ***reservoir*** population of organisms that, in addition to human hosts, maintain the agent.

New technologies in production and transportation, human population pressure, and climate change are transforming ecological systems, even at the global scale. These transformations involve changes in land use, vegetation cover, species and species locations, and climate. Many of these factors, along with elimination or decreased support for public health programs to control vector-borne diseases, have been suggested as explanations for the resurgence of disease. Vector-borne diseases affect both animals and plants. In addition to the effects of vector-borne diseases on agricultural production, some diseases infect only wild animals and plants. The implications of the resulting changes to ecosystems extend, even if indirectly, to human health. ***Ecoepidemiology*** has emerged in response to a perceived need to broaden the scope of assessment of the impacts of environmental change (Hales, Weinstein, & Woodward, 1997, p. 191). This approach entails a shift in emphasis from the study of direct, or toxicological, mechanisms to the study of indirect, or ecological, mechanisms and a shift from the individual to the region.

GIS applications in the study of zoonotic and vector-borne diseases attempt to model some aspect of how people live with animals and vectors in a particular ecological system. Chapter 8 provides an overview of the global burden of zoonotic diseases and examples of GIS analyses of surveillance and detection, vector and host population distribution, land use and activity patterns that bring people into areas where vectors and hosts are present, outbreak prediction, and vector control.

The Global Burden of Zoonotic Diseases and the Challenge of Emerging Infectious Diseases

The impact of disease in populations has traditionally been measured in terms of mortality, an approach that does not account for the effects of illness or disabling injury. The first global burden of disease study, published in 1990 (Murray & Lopez, 1996), introduced the ***disability-adjusted life year (DALY)*** as a measure of disease burden. The DALY for a disease or injury cause is calculated as the sum of the years of life lost due to premature mortality in the population and the years lost due to disability for incident cases of the disease or injury. One DALY corresponds to one lost year of healthy life. The ***burden of disease*** in a population is the gap between the current years of healthy life and the years of healthy life that would be lived if everyone in the population survived the full life span free of disease and disability. Although this approach to the measurement of health status in populations has been criticized and methods for measuring disease burden continue to be refined (Sundby, 1999; World Health Organization, 2008), the global burden of disease studies provide a framework for placing vector-borne diseases in context.

Based on data for 2004, zoonotic diseases are not major contributors to the global burden of disease (World Health Organization, 2008). Malaria, caused by one of several species of parasites of the genus *Plasmodium* transmitted to humans by *Anopheles* mosquitoes, ranks twelfth on the list of the leading causes of global burden of disease in people of all ages. This ranking is primarily due to the high incidence of malaria in low-income countries in Africa. On both the list for low-income countries and the list for all countries in Africa, malaria ranks fourth in the leading causes of the global burden of disease for people of all ages. Worldwide, noncommunicable diseases account for nearly half of the burden of disease in people of all ages. Among adults, even in low- and middle-income countries, noncommunicable diseases account for 45% of the disease burden.

Given these statistics, what explains the growing concern about zoonotic and vector-borne diseases? In any given region, some vector-borne diseases are ***endemic*** or permanently present even if they are controlled. The reemergence of diseases in areas where they have been present, the importation of cases of diseases in areas where they have not been present, and the establishment of disease transmission in areas where transmission has not occurred before are all causes of concern.

The resurgence of diseases like epidemic yellow fever and dengue fever over the last 30 years has contributed to a renewed focus on vector-borne diseases (Gubler, 2004). Dengue fever, also known as "breakbone fever," is caused by a virus that is a member of the *Flavirus* genus, like the West Nile virus and the virus causing Japanese encephalitis. This virus is also known as an ***arbovirus,*** shortened from arthropod-borne virus, because it requires a blood-sucking arthropod to complete its life cycle. Dengue is different from most flaviruses, which depend on animals other than humans to maintain the natural transmission cycle in which humans are only incidental (Mackenzie, Gubler, & Petersen, 2004). Although animal reservoirs where dengue is maintained still exist in tropical forests, the virus is adapted to humans and is maintained in urban areas in tropical regions. There are four serotypes of the dengue virus. ***Serotypes*** are groups of closely related microorganisms distinguished by characteristic sets of antigens. A case of dengue confers immunity only for that serotype. A person could contract the disease four times, once with each serotype, during a lifetime.

In the late 18th century, a disease like dengue was causing epidemics in Asia and the Americas (Holmes & Twiddy, 2003). Shortly after World War II, an outbreak of hemorrhagic fever in children in Manila was recognized as dengue. A small proportion of infected people develop dengue hemorrhagic fever or dengue shock syndrome, more lethal than dengue fever, which is a self-limiting illness. In 1970, multiple serotypes of the dengue virus were found primarily in Southeast Asia where the disease was ***hyperendemic,*** exhibiting high and continued incidence (Mackenzie et al., 2004). In other regions, dengue was ***hypoendemic,*** affecting only a small proportion of the population at risk, or not endemic. By 2007, tropical regions in both hemispheres were hyperendemic with multiple serotypes found (Figure 8.1). Disease is ***holoendemic*** when almost every person in a population is affected.

Transportation systems have made it possible to find cases of diseases in countries where these diseases are not usually found. More than 1,000 confirmed cases of dengue virus infection were observed in Germany from 2002 through 2007, making it one of the most common travel-related infections (Jansen, Frank, Koch, & Stark, 2008). Only one case was acquired in Germany, the result of a needle stick injury in a hospital. Such ***nosocomial infections,*** resulting from medical care treatment but secondary to the person's original condition, laboratory-acquired infections, and infections resulting from blood transfusion or organ donation can be factors in vector-borne diseases (Kotton, 2007). The reported countries where dengue infections observed in Germany were acquired are spread around the globe (Figure 8.2).

Travelers can be thought of as "interactive biological units that pick up, process and drop off microbial genetic material at different times and in different places," acting as sentinels, couriers, or transmitters of disease (Wilson, 2003). Travel was a factor in an outbreak of chikungunya documented in Italy, a temper-

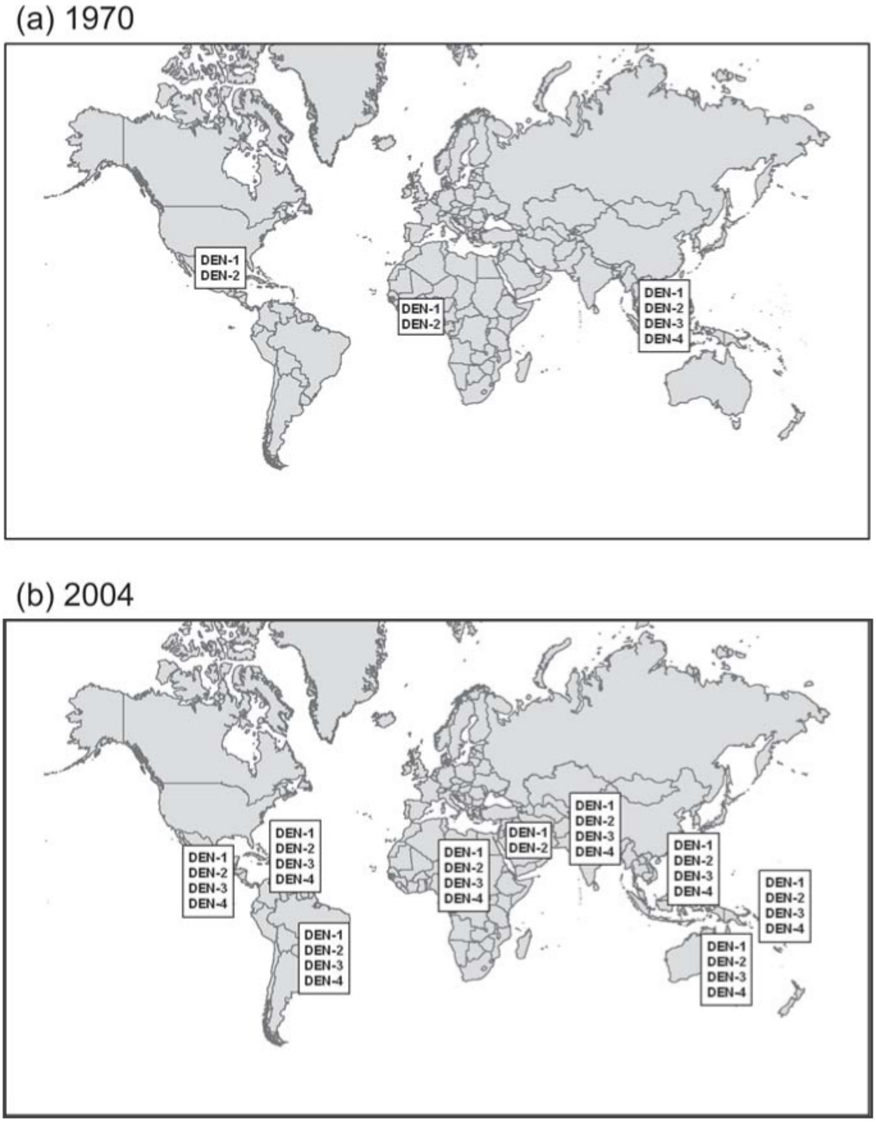


FIGURE 8.1. The global distribution of dengue virus serotypes has changed over the last 30 years. By 2004, all four serotypes were present in most tropical regions. Mackenzie, Gubler, and Petersen (2004). Reprinted by permission from Macmillan Publishers Ltd.: Nature Medicine, copyright (2004).

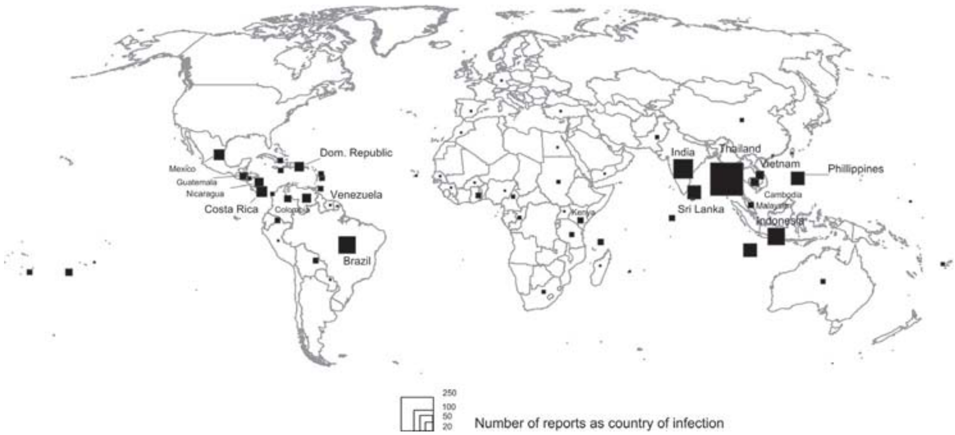


FIGURE 8.2. The number of dengue cases reported in Germany from 2002 to 2007 mapped by the country where the infection was acquired. From Jensen, Frank, Koch, and Stark (2008). Copyright 2008 by Springer. With kind permission from Springer Science+Business Media.

ate region, in 2007 (Rezza et al., 2007). Chikungunya is a virus transmitted to humans by *Aedes* mosquitoes, and it was first isolated in Tanzania in 1953. The vector in the outbreak in Italy, *A. Albopictus*, was first documented in the country in 1990, and a breeding population was established by 1991. The introduction of this competent vector was traced to the warehouse of a tire retreading plant that had imported used tires infested with mosquito eggs from a supplier in Georgia in the United States. The virus was introduced in 2007 by a man traveling from a region in India where a large outbreak had occurred. He was visiting relatives in Italy, was probably high viremic at the time of his visit, and developed a fever after several days in Italy. The introduction of the agent was then sustained by the local vector population. In addition to this man, the ***index case*** or first case, more than 200 additional cases were identified. Laboratory analysis showed a high similarity between the strains found in Italy and those found in an earlier outbreak on islands in the Indian Ocean. It is possible that the vector *A. albopictus* is also establishing itself in the Upper Rhine Valley in Germany (Jansen et al., 2008). While some cases of chikungunya occurring in Europe are imported due to the return of travelers from regions where the disease is endemic and was acquired, other European cases of chikungunya may now be ***autochthonous,*** resulting from transmission in the place where the disease is found.

Along with the resurgence of vector-borne diseases and their introduction to different areas, new diseases have emerged. A study of 335 emerging infectious diseases affecting humans and originating between 1940 and 2004 found a significant rise in these diseases over time (Jones et al., 2008). These origin events were dominated by zoonoses (60.3%), with the majority of these occurring in wildlife (71.2%). Vector-borne diseases accounted for 22.8% of all emerging infectious diseases studied, and this percentage was higher among diseases emerging in the last decade of the study period. "Once introduced and established, it is unlikely that zoonotic disease agents can be eliminated from an area" (Gubler, 2008, p. 63).

The threat that an emerging or reemerging disease might lead to a ***pandemic,*** an outbreak of many cases in a large geographic area, has led to calls for improved surveillance of vector-borne diseases and for strengthened local public health infrastructure to prevent and contain disease outbreaks. GIS analyses have been critical to the prevention and control efforts for many vector-borne diseases.

Surveillance and Mapping of Vector-Borne Diseases

Designing effective surveillance programs for vector-borne diseases is a challenging task. Monitoring these diseases requires understanding the distribution of disease vectors and the strains of disease they may be carrying, the distribution of animal and human host populations and their levels of susceptibility, and the distribution of animal and human cases. Because vector-borne diseases are emerging in areas where they have not been found in the past, surveillance systems need to be designed to detect the introduction of vectors and diseases and to report relevant information so that effective control measures can be undertaken and evaluated for their effectiveness.

Surveillance of Human Cases

CASE DEFINITION

The distribution of vector-borne infectious diseases is often assessed using human case reports, in part because good baseline data on the distribution of vectors has sometimes been unavailable. Like the infectious diseases discussed in Chapter 7, vector-borne diseases are often complex and difficult to diagnose. Lyme disease, the most common tick-borne disease in the United States, illustrates this problem.

Lyme disease is a multistage, multisystem disease caused in North America by *Borrelia burgdorferi*, a spirochete transmitted from mammal to mammal by ticks of the genus *Ixodes*. Humans "are inadvertent hosts of the spirochete" (Walker et al., 1996, p. 463). In 60 to 80% of patients, early Lyme disease is indicated by erythema migrans, a characteristic skin rash appearing around the site of the tick bite after a week. Late-stage disease affecting the musculoskeletal and neurological systems can be more difficult to diagnose, particularly if there is no history of erythema migrans in a person who lives in or has visited an area where Lyme disease is endemic, or permanently present, and depends on laboratory confirmation.

In 1990, the Centers for Disease Control (CDC) and the Council of State and Territorial Epidemiologists published, for the first time, uniform criteria for reporting cases (Wharton, Chorba, Vogt, Morse, & Buehler, 1990). At the same time, the CDC made Lyme disease nationally notifiable and developed a national case definition. In the years after Lyme disease was initially reported in southeastern Connecticut (Steere, Broderick, & Malawista, 1978), state criteria for defining a Lyme disease case varied (Vogt, 1992). The development of a standard definition is important in the design and implementation of a reporting system so that the time and expense of reporting will result in consistent information on valid cases. Since 1990, the case definition for Lyme disease in the United States has been revised twice, in 1996 and in 2008 (Centers for Disease Control and Prevention, 2010a). The most recent revision modified language on laboratory evidence for diagnosis and case classification, including changing the definition for a confirmed case and adding definitions for probable and suspected cases.

The complexity of the case definition reflects the difficulty of diagnosing cases of many vector-borne infectious diseases. A case of Lyme disease is confirmed if the person has the skin lesion erythema migrans or if at least one late manifestation of disease is present and the case is laboratory-confirmed. Erythema migrans must be diagnosed by a physician, and laboratory evidence of infection is required for persons with no known exposure and with persons with at least one late manifestation of the disease. Exposure is defined as having been in a county where Lyme disease is endemic within 30 days before the onset of erythema migrans. The Lyme disease case definition contains a geographic standard for determining whether disease is endemic to a county: if at least two confirmed cases have been previously acquired or established populations of a known tick vector are infected with *B. burgdorferi.* This standard has, however, been problematic in some regions like the southeastern United States where erythema migrans and other Lyme disease symptoms have been observed in patients who live in areas where there is no documentation of transmission of *B. burgdorferi* to humans. Also, a "count by county" standard is not uniform across the country because counties vary considerably in land area, habitat area, and population size. The county of San Bernardino, California, for example, is larger in area than the state of Connecticut (about four times as large, or 20,064 square miles compared to 4,872) and Connecticut has eight counties.

Case definitions of vector-borne diseases, though necessary for surveillance, are sometimes controversial. Surveillance case definitions developed for reporting may differ from definitions of clinical cases. The actual burden of disease for many vector-borne diseases is believed to be much higher than reported due to surveillance issues. Researchers have argued, for example, that the current case definition for dengue hemorrhagic fever contributes to a misperception of low disease burden from dengue in the Western Hemisphere (Rigau-Pérez, 2008). They have proposed defining additional disease endpoints besides dengue hemorrhagic fever and improving the sensitivity of tests for dengue hemorrhagic fever so that the tests can be applied even in regions of the world where laboratory services are limited.

MAPPING HUMAN CASES

For some emerging infectious diseases, temporal and spatial concentration of human health problems with rapid onset may have been the clue that the disease was infectious or vector-borne in the first place. For infectious diseases of known etiology, the identification of areas of high and low incidence is relevant to the design of intervention strategies because high-incidence areas may indicate places where many people are being exposed to the disease agent, particularly if the time between exposure and onset is short. Areas of low incidence are either places where the agent–vector–host cycle is not established, or where people are not present, or both.

Case data from a surveillance system can be geocoded using the methods described in Chapter 3. Surveillance systems usually report residential address. Distribution by residential location is almost always relevant for medical services planning, but the residence may not be the relevant location for identifying disease clusters. Questions about the relevant locations to map and the impacts of aggregating cases or populations at risk to geographic areas have been raised for decades (Maxcy, 1926). Spot maps of typhus cases in southeastern U.S. cities showed no particular concentration by residential neighborhood other than a tendency toward localization in the central portions of the cities (Figure 8.3). Spot maps by place of employment, however, revealed focal centers. Because exposure to Lyme disease in Connecticut is believed to occur primarily on a peridomestic basis, looking at residential locations makes sense. In other regions of the country or in other countries where exposure to infected ticks may be more likely in recreational settings, the residential distribution of cases may not be useful in identifying areas where people are at risk for acquiring the disease.

Publishing maps of cases based on actual individual residential location may disclose the identities of individuals, as discussed in Chapter 7. In addition, studies of zoonoses are increasingly conducted at the global or regional scale, and maps of individual case locations at these scales are not useful. For these reasons, case data from surveillance systems are usually aggregated for reporting purposes. When mapping count data reported for areas, choropleth maps with areas shaded based on the count of cases, as discussed in Chapter 4, are not always effective in communicating spatial patterns, especially if the geographic areas of the map units vary considerably in size (Figure 8.4).

The distribution of cases needs to be understood in the context of the distribution of the denominator population. One approach for accomplishing this is to calculate rates for predetermined geographic areas, like census tracts. Some drawbacks to this approach are that the geographic areas used for reporting often arbitrarily partition the numerator and the denominator. Places with low rates may have large numbers of cases, places with high rates may have small numbers of cases, and some places may have no cases because no one lives there. As in documenting core areas for communicable disease transmission, GIS analysis enables us to look at the joint distributions of cases and population so that we can better classify areas.

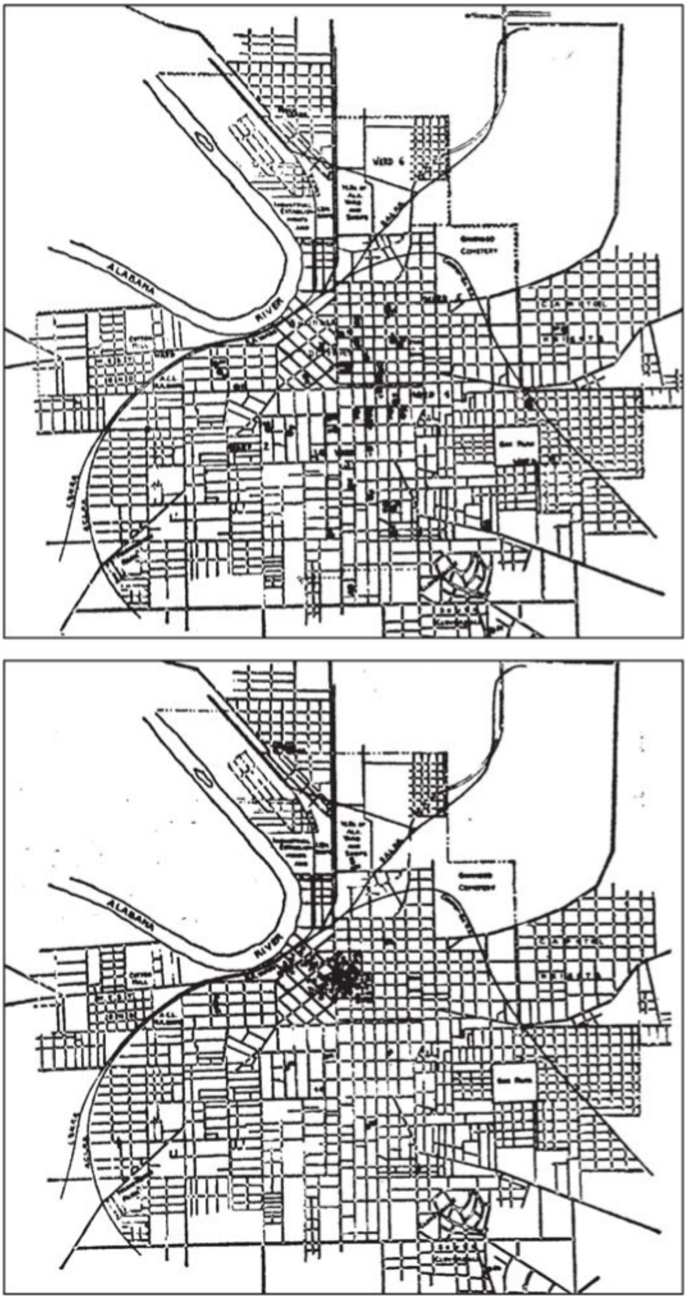


FIGURE 8.3. The top map of cases of mild typhus in Montgomery, Alabama, 1922–1925, according to residence shows no clustering. The bottom map of cases according to place of employment shows a concentration of cases near the center of the city. From Maxcy (1926).

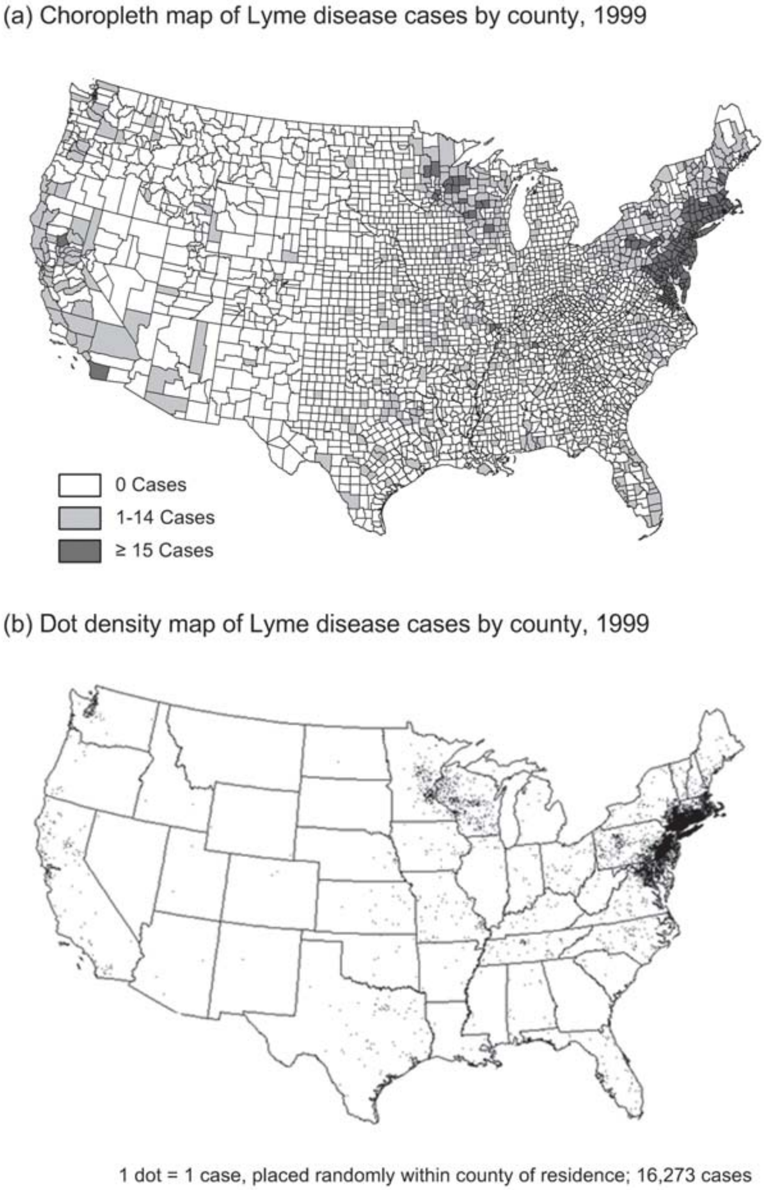


FIGURE 8.4. Mapping human cases of Lyme disease. The maps report data at the county level in the United States for the same year, 1999. Both maps display counts by county. Because counties in some states, especially in the West, are large, the map in Figure 8.4a overstates the level of disease in large counties. The dot density map in Figure 8.4b offers a more visually accurate picture of the distribution of cases. Figure 8.4a is from Centers for Disease Control and Prevention (2001). Figure 8.4b is from Centers for Disease Control and Prevention (2004).

Spatial databases can be used in GIS to show various representations of population distribution (Figure 8.5): the choropleth map of population counts by census tract, the street network representation that shows where structures are likely to be located, the land use/land cover representation derived from remote sensing data that shows areas of residential development of different density, and the grid representation that interpolates population counts for political-administrative units to a grid based on ancillary data. The street network view helps the public health analyst better interpret the meaning of no cases or few cases reported from an area by providing a picture of whether or not residential

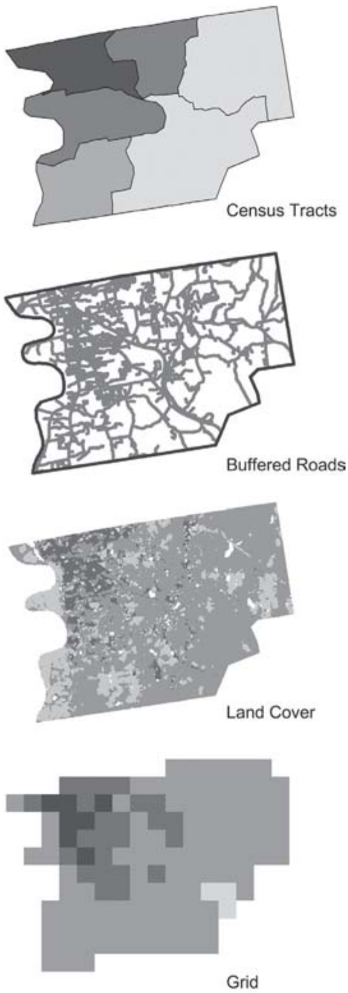


FIGURE 8.5. Spatial databases in a GIS provide different kinds of information about the population distribution within a town.

or high-density residential development exists in the area. The land use/land cover representation depicts the settings of residential areas—for example, adjacent to commercial strips or surrounded by forested zones. As noted in Chapter 6, grid representations of population are increasingly being used in health studies to model population density at a variety of scales.

Although we will not observe a case in a place where there is no residence, it is not always true that disease patterns "follow" the distribution of residences or population. If it were true, there would be no geographical variations in disease unexplained by the geographical distribution in population. The same case distributions can be embedded in different underlying population distributions (Figure 8.6). This highlights the need to view the joint distribution of numerator

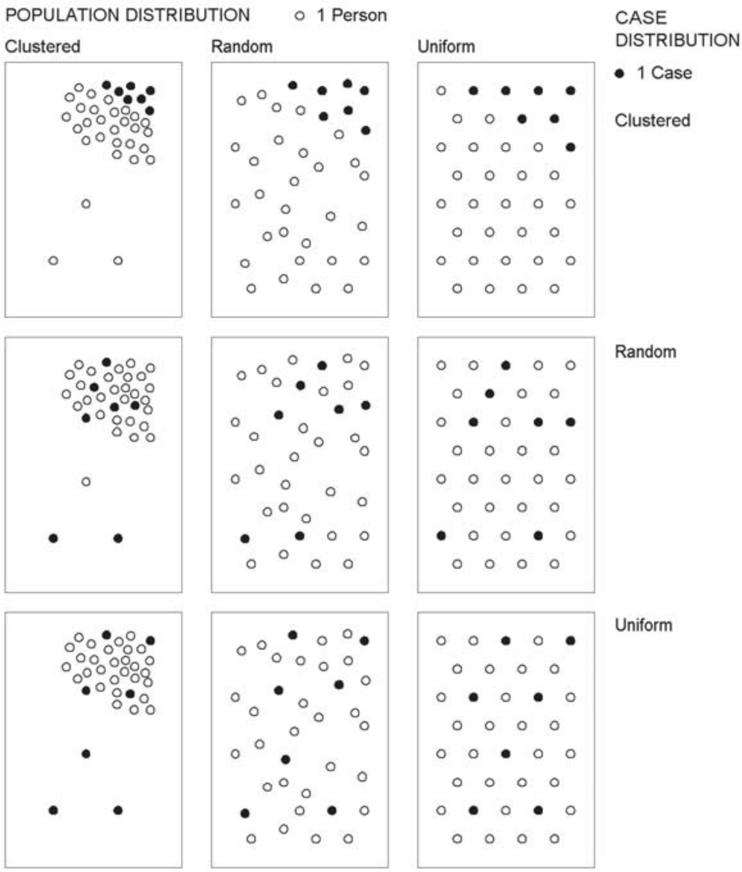


FIGURE 8.6. The number of cases of disease and the population size are the same in each study area shown above, but similar case distributions can be embedded in different underlying population distributions.

and denominator before calculating rates for areas that might arbitrarily partition the numerator and denominator or before performing a clustering analysis with an arbitrary distance criterion. Maxcy (1926, pp. 2975–2976) recognized this problem in his study of typhus: "The question arises whether this apparent concentration is merely the result of a greater density of population in that part of the city. . . . The division of the city is peculiarly unfavorable for the purposes in mind [rate calculation], inasmuch as the wards are arranged radially in such manner that all except one include portions of the central part of the city."

Maps of cases rather than rates also provide useful map sequences showing the spread of vector-borne diseases. Maps of West Nile virus cases in the United States reported at the state level document the westward spread of the disease since the first case was observed in 1999 (Figure 8.7). During the summer and fall of 1999, an outbreak of West Nile encephalitis occurred in New York. Although West Nile virus had been used in laboratories in the United States and shipped to other countries before 1999 (Cromley, 2003), this was the first outbreak of the virus in the Western Hemisphere.

West Nile fever is a mosquito-borne flavivirus infection endemic in Africa and Asia (Tsai, Popovici, Cernescu, Campbell, & Nedelcu, 1998). Outbreaks occurred in southern France in the early 1960s and in Romania in 1996 (Lundstrom, 1999). The outbreak in Romania resulted in a high fatality/case ratio (Tsai et al., 1998). Mosquitoes in the home and, for apartment dwellers, flooded basements, in addition to spending time outdoors, were confirmed as risk factors in two case–control studies conducted in Romania (Han et al., 1999).

In New York, 56 cases including seven deaths had occurred by late fall of 1999. One of the fatal cases was an international case involving a Canadian citizen who had visited New York late in the summer and had onset of encephalitis several weeks later. The outbreak of West Nile virus was accompanied by intense research efforts to determine the particular strain involved (Anderson et al., 1999; Jia et al., 1999; Lanciotti et al., 1999). Along with the human cases during the outbreak in New York, West Nile virus contributed to increased deaths in wildlife in the area and the early response to the outbreak included surveillance of deaths in animal populations (Figure 8.8).

Animal Surveillance

Given the key role of animal populations in zoonotic and vector-borne diseases, surveillance of animal host populations is also important. For example, studies have attempted to model the distribution of infected deer ticks by collecting ticks from deer killed by hunters and hunter-reported location of kill (Kitron, Jones, Bouseman, Nelson, & Baumgartner, 1992; Amerasinghe et al., 1992; Glass, Amerasinghe, Morgan, & Scott, 1994). These studies are obviously limited by spatial biases in where deer are killed. Nevertheless, GIS applications have been developed to map disease in wild animals, domesticated animals, and companion animals.

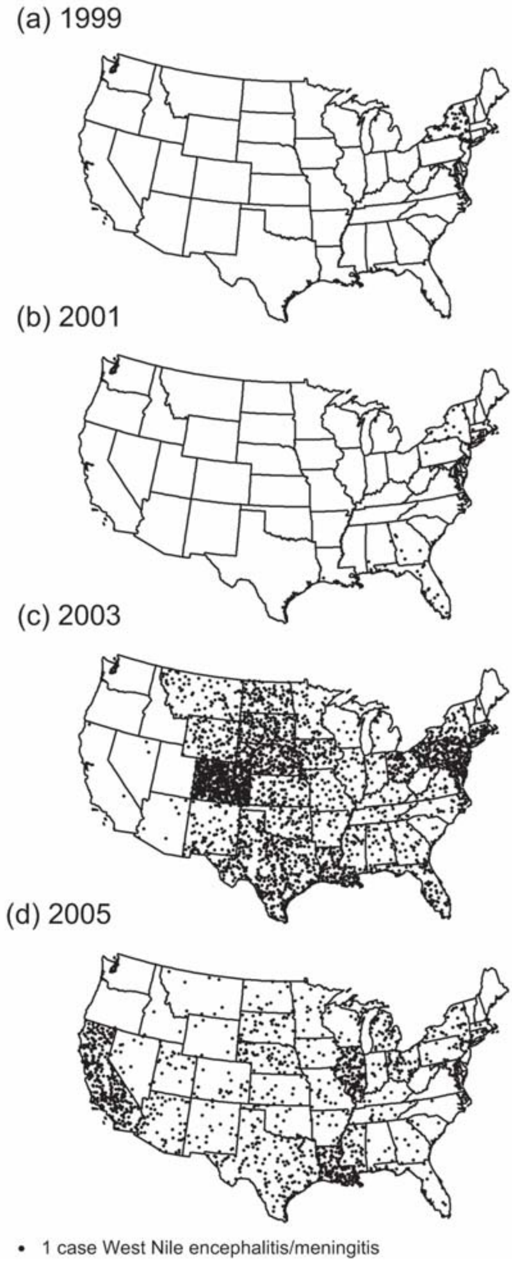


FIGURE 8.7. A map sequence of the spread of West Nile cases by state shows the original outbreak in New York in 1999. By 2005, the virus had reached the West Coast and was endemic in most states. The cases shown are cases of neuroinvasive disease because reporting is more complete for the more serious forms of the disease than for West Nile fever. Data from Centers for Disease Control and Prevention (2010b).

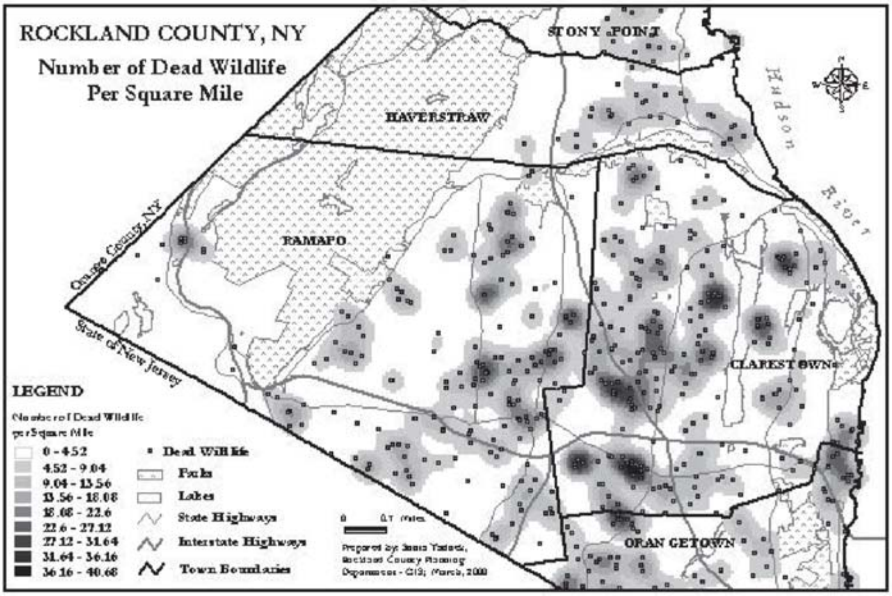


FIGURE 8.8. The number of dead wildlife per square mile in Rockland County, New York, mapped as part of the surveillance effort for West Nile virus after the 1999 outbreak. From Rockland County Planning Department GIS (2000). Reproduced by permission of Rockland County, New York. We hold Rockland County harmless regarding the accuracy of the map.

WILDLIFE SURVEILLANCE FOR RABIES

Rabies is a zoonotic disease caused by an RNA virus (Krebs, Wilson, & Childs, 1995). Animal hosts maintain and transmit the disease to humans, usually by bites, although nonbite exposures have been documented. Human infections are not important in maintaining the virus because humans do not contribute to the transmission cycle. Mammalian carnivores are the essential hosts, primarily dogs in the developing world and wildlife in developed countries where rabies in domesticated dogs has been brought under control through vaccination programs. In the United States, wildlife has been the principal reservoir since the 1960s (Krebs, Strine, & Childs, 1993), with different wild terrestrial reservoirs in different regions of the country (Blanton, Palmer, Christian, & Rupprecht, 2008). In humans, the virus affects the central nervous system and results in death unless effective postexposure treatment is provided.

Human rabies cases are rare, but the virus that causes rabies is widespread globally and many people are exposed annually to animals with suspected rabies (Warrell & Warrell, 2004). From 2000 to 2007, only 25 cases were reported in the United States (Centers for Disease Control and Prevention, 2008c). Bats are recognized as an important wildlife reservoir for variants of the rabies virus that are transmitted to humans, accounting for 58% of cases diagnosed in the United States from 1980 to 1998 (Centers for Disease Control and Prevention, 1998). Bat bites are of particular concern because injury from a bat bite is often more limited than injury from the bite of a terrestrial carnivore like a raccoon, and people may not be aware that they have been bitten. Because "all bites by carnivores (especially raccoons, skunks, and foxes) and bats must be considered possible exposures" (Krebs et al., 1995, p. 689), understanding the geographical patterns of places where human contact with these animals has occurred is important. GIS can be used to portray these patterns.

Bretsky (1995) used GIS to map the spread of rabies in Connecticut from 1991 through 1994, on the basis of reports of human contact with animals. Connecticut's experience during this time was part of the most intense outbreak of wildlife rabies ever to occur in the United States (Rupprecht, Smith, Fekadu, & Childs, 1995). From the time the index case was reported in West Virginia in 1977, it took approximately 14 years for the first case to arrive in the town of Ridgefield, Connecticut, on the New York border. Over the next 4 years, the disease spread throughout the state from southwest to northeast (Figure 8.9), and a second wave was reemerging in the southwestern part of the state. The initial epizootic wave had advanced approximately 30 kilometers per year (Wilson et al., 1997).

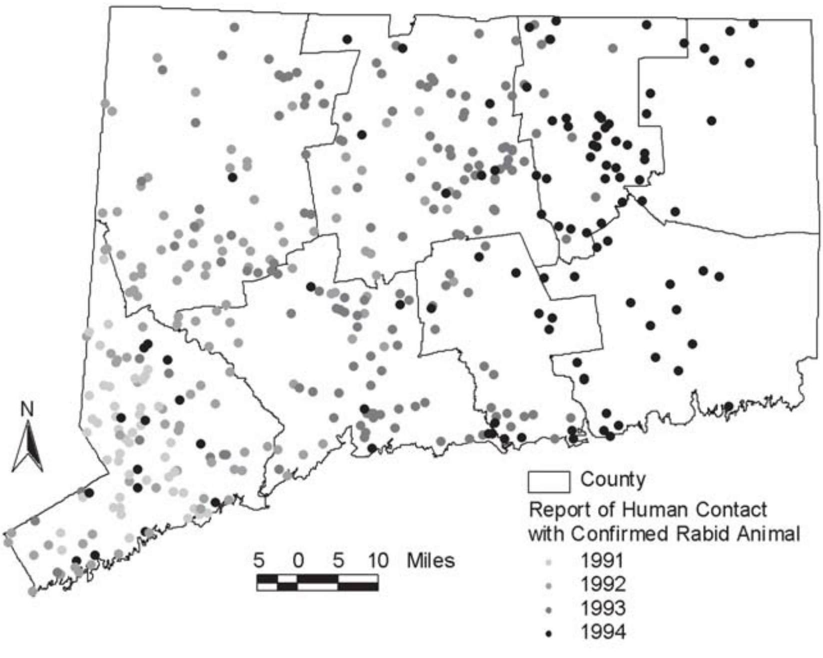


FIGURE 8.9. Reports of human contact with a confirmed rabid animal during the 1991–1994 epizootic in Connecticut. Data from Connecticut Department of Environmental Protection.

LIVESTOCK SURVEILLANCE FOR RIFT VALLEY FEVER

Rift Valley fever is a zoonosis caused by a virus in the *Phlebovirus* genus. Although it primarily affects animals, humans can acquire the disease, in the vast majority of cases through direct and indirect contact with organs and blood of infected animals. First identified during an outbreak of disease in sheep in the Rift Valley of Kenya in 1931, repeated outbreaks have since occurred in North Africa and sub-Saharan Africa. In 2000, the first cases were confirmed outside of Africa, in Saudi Arabia and Yemen. The disease has not been observed in urban areas, and, among humans, agricultural workers, meat processors, and veterinarians are most likely to be affected by the disease.

Because of the economic consequences for livestock herds, Saudi Arabia imposed import bans on Somali livestock (Soumare et al., 2007). The Somaliland government had a strong interest in developing a surveillance system that would comply with World Organisation for Animal Health requirements for random sampling. In the study region, pastoral herds of sheep and goats are nomadic.

A GIS application was developed to implement a two-stage random sampling process. First, a set of random sampling sites are identified, and, second, the required number of animals are randomly selected from the nearest herds to the sites. A GIS extension that generates random coordinates for the required number of sample sites within an area was used for the first stage. Then, circular buffers were created around each of the sample sites to identify the geographical area where sampling of animals was carried out. The buffer distance was adapted to conditions; a radius less than 10 kilometers in dry areas might result in too few animals for sampling but a shorter radius might be appropriate in more densely populated areas around permanent watering points. A surplus of random points was generated in case target sites could not be accessed due to political conflict, the presence of landmines, or other factors.

The sampling scheme was used in Somaliland in 2001 and in Puntland in 2003. A follow-up survey was conducted in 2004. The results showed the highest antibody activity in the Nugal Valley. The research suggests the possible usefulness of establishing a resident sentinel herd to monitor future disease outbreaks.

CANINE SURVEILLANCE FOR LYME BORRELIOSIS

B. burgdorferi causes Lyme disease not only in humans but also in domestic animals, especially dogs. Dogs are competent reservoirs of the disease, and a number of studies have attempted to assess the distribution of Lyme disease based on prevalence of the agent in dogs. One study conducted in Wisconsin and northern Illinois used a canine survey to assess seroprevalence of antibodies to the Lyme disease agent in pet dogs (Guerra, Walker, & Kitron, 2001). Samples were obtained from veterinarians in counties with a history of Lyme disease or presence of the *Ixodes scapularis* vector and in adjacent counties.

The residential locations of the dogs were mapped using GIS. Dogs were classified as seropositive or seronegative, taking into account whether or not they had been vaccinated. In general, the spatial distribution of canine seroprevalence matched the reported distributions of human incidence and vector ticks in the study area, but canine seroprevalence was also observed in a county where people were not previously considered at risk for Lyme disease. The study demonstrated the potential use of an active surveillance program for pet dogs, especially for monitoring the spread of disease. A review of animal sentinel studies concluded that the use of animal data to predict human risk has been limited to date and suggested that increasing attention be given to the use of animal data as sentinel information of relevance to human health (Scotch, Odofin, & Rabinowitz, 2009).

Vector Surveillance

The resurgence of vector-borne diseases has been associated with renewed efforts to monitor vector distribution directly and to analyze vector infection rates. As with other infectious disease surveillance programs, support for vector surveillance eroded during the 1970s and 1980s when vector-borne infectious disease was not at the forefront of public health concern. Vector surveillance is time consuming and expensive, and good baseline data sets are often not available. Many tick collection studies undertaken in response to the emergence of Lyme disease were based on the first reported collection from the site where the study took place (Ginsberg, 1993).

SAMPLING FOR VECTOR SURVEILLANCE

To develop a scheme for surveying tick populations on a statewide basis in Rhode Island, a GIS analysis partitioned the state into 42 zones 10 kilometers square (Nicholson & Mather, 1996). Road, land use, vegetation, and hydrography data were included. Forested habitats were identified as areas where more than 50% of the cover was tree canopy. Land use/land cover data were derived from remote sensing data. Eighty tick collection sites of approximately 4 hectares were selected based on location in the state, type and amount of forested habitat, and road accessibility. Each large zone had from one to three tick sampling sites located within it. After 18 samples were taken from each site, mean tick density and mean infection rate were calculated. An entomologic risk index was computed as the product of tick abundance and the local infection rate.

In a study conducted in the Middle Atlantic states, another region of high Lyme disease incidence in humans, tick collection sites were randomly selected in state parks, state forests, and other large open areas providing access (Bunnell, Price, Das, Shields, & Glass, 2003). Flags one square meter in area were used to collect ticks, and GPS was used to record lon/lat coordinates of the data collection sites. Over 2 years, a total of 320 sites were sampled, including 24 sites sampled in both years. GIS analyses were performed to assess the habitat characteristics associated with adult tick density patterns.

A transect sampling approach was used in an entomological survey of sandflies near Mount Vesuvius, an area of intense transmission of leishmaniasis in southern Italy (Rossi et al., 2007). Leishmaniasis is caused by protozoan parasites transmitted by sandflies of the *Phlebotominae* subfamily. GIS was used to design the sampling procedures to identify 49 sites split approximately evenly on the coastal and Apennine sides of the volcano. GPS was used to record the locations of the sample sites. Sandfly densities were higher on the coastal side, which is also less urban. In addition to research conducted in the field to collect vectors and test them for infection, passive surveillance of vectors has also been used.

Passive Surveillance of Vectors

In the late 1980s, following documentation of the presence of *I. scapularis* ticks in Maine and outbreaks of Lyme disease in Massachusetts, the state of Maine initiated a statewide tick identification service for the general public (Rand et al., 2007). Veterinary clinics have also been important contributors to the system. During the period 1989–2006, more than 24,000 ticks were submitted to the program, representing 14 species. About half of the ticks were *I. scapularis*.

The availability of these data has made it possible to compare tick abundance and human cases of Lyme disease over multiple years and regions in the state during a period when the number of cases increased dramatically. The ecological impacts of the advance of *I. scapularis* into previously uninfested areas, including decreases in the population of the common mouse tick *I. muris* perhaps due to displacement, have also been investigated. Despite the limitations of passive surveillance, the system has yielded a wealth of data relevant to vector-borne diseases.

Temporal and Spatial Integration of Surveillance Data

One of the most important functions of a GIS is data integration. In the case of infectious disease epidemiology, GIS can be used effectively to integrate data temporally and spatially. In the case of zoonotic diseases, there is a recognized need to integrate data on vectors, hosts, and human cases in time and place (Brooker & Utzinger, 2007). These efforts include designing integrated surveillance systems mapping human case data in relation to vector and host populations.

REGIONAL VARIATION IN VECTORS AND DISEASES

Many vector-borne infectious diseases involve the same agent but different vectors and hosts in different regions. In the northeastern United States, the principal Lyme disease vector is *I. scapularis*. Lyme disease transmission in California is very different from the cycle in the northeastern United States. In that state, the pattern has two cycles involving two different ticks (Brown & Lane, 1992).

I. neotomae, with higher observed infection rates, was identified as responsible for maintaining the disease in the woodrat population but not involved in the human disease because it does not bite humans. *I. pacificus* ticks were infected at lower rates, not enough to maintain endemic disease, but enough to transmit the disease to humans. The detection of infected woodrats and *I. pacificus* ticks in the mountains near Los Angeles suggests a Lyme disease cycle maintained in wildlife in a main recreational area for one of the largest metropolitan areas in the country. Subsequent research has identified more than 100 verterbrate species, including mammals, birds, and reptiles that serve as host species for at least one stage in the *I. pacificus* life cycle (Castro & Wright, 2007).

While the same disease may be transmitted by different vectors, the same vector may transmit more than one disease. Human anaplasmosis, a disease formerly known as human granulocytic ehrlichiosis (HGE) and later as human granulocytic anaplasmosis (HGA), is a tick-borne disease caused by *Anaplasma phagocytophilum*. The ticks that transmit *B. burgdoreferi*, the agent for Lyme disease (*I. scapularis* in the eastern and upper midwestern United States, *I. pacificus* in the western United States, and *I. ricinus* in Europe), also transmit *A. phagocytophilum*.

At the time Lyme disease became a notifiable disease in the United States, HGE was recognized and had an established case definition but it was not a reportable disease. Over the last decade, the CDC has broadened its surveillance of human cases of rickettsial vector-borne diseases (Waxman, 2009). There has also been additional research to assess the rate of coinfection in vectors (Nadelman et al., 1997; Holman et al., 2004), including infection with *Babesia microti*, microscopic parasites that infect red blood cells and cause babesiosis. The observed rates of coinfection in vectors have ranged from 2 to 26%. Because the diseases may be transmitted by the same tick, testing humans for coinfections is warranted in patients suspected of having acquired other tick-borne diseases. Interventions to control tick populations would obviously have implications for all of the diseases. The value of monitoring human cases, vectors, and host populations together is demonstrated in the ArboNET surveillance system.

THE ARBONET SURVEILLANCE SYSTEM

To limit the impact of West Nile virus in the United States after the outbreak in 1999, the CDC and the U.S. Department of Agriculture cosponsored a meeting in November 1999 to develop programs to monitor virus activity and to prevent future outbreaks of disease (Centers for Disease Control and Prevention, 2000). The New York outbreak had resulted in extensive mortality in crows. Because of bird migration patterns, CDC efforts included surveillance of areas from Louisiana and Alabama along the Gulf Coast to Massachusetts and Maine. The surveillance guidelines called for active bird surveillance in both wild and sentinel populations, active mosquito surveillance to monitor virus activity and to identify potential vectors, active veterinary surveillance—particularly for horses

(Nolen, 2000), and enhanced passive human surveillance for reporting viral encephalitis. The guidelines also described the minimal laboratory diagnostic support required and emphasized mosquito control methods for preventing further outbreaks of the disease.

The ArboNET was introduced in 2000 as a national electronic surveillance system for arboviral diseases in the United States (Institute of Medicine and National Research Council, 2008). Although developed in response to the West Nile outbreak, more arboviral diseases were added beginning in 2003 and 14 diseases are monitored under the system as of 2008. ArboNET surveillance is broad in scope, covering human cases, presumptive viremic blood donors, veterinary cases, avian cases, sentinel chickens, and mosquitoes. Health care providers, laboratories, and veterinarians report data to state and local health departments where the surveillance data are entered into an electronic database. Data are reported at the county level. Reports are uploaded to the CDC database, usually on a weekly basis. The U.S. Geological Survey prepares weekly updates of maps with cases by state and county (U.S. Geological Survey, 2010c).

This system enables analysts to integrate data on human cases with data on West Nile activity in host and vector populations (Figure 8.10). An advantage of the system is that human cases are reported using standard case definitions for neuroinvasive West Nile disease (causing meningitis and encephalitis) and for nonneuroinvasive West Nile disease (causing fever and other symptoms). As noted earlier in the discussion of case definition, there is bias in reported West Nile cases. Testing and reporting is more complete and representative for the more severe neuroinvasive disease than for West Nile fever. Nevertheless, reports of West Nile fever show the presence of disease in areas where no cases of the more severe form have been reported. Although only human cases are nationally notifiable in the ArboNET system and all surveillance is passive, the system has been an important advance in vector-borne diseases surveillance in the United States.

EMERGING DISEASES IN A CHANGING EUROPEAN ENVIRONMENT

Integrative work in vector-borne diseases has also proceeded in Eurpe. The European Commission funded a program on ***Emerging Diseases in a Changing European Environment (EDEN)*** as part of its Sixth Framework. The EDEN program links 49 partners from 24 countries in Europe and the Mediterranean basin in a network of 80 teams (Lancelot, Poncon, Hendrickx, & Fontenille, 2009). The interdisciplinary teams support close integration of biology, ecology, geography, and modeling in the study of vector-borne diseases. By identifying ecosystems with a high risk of vector-borne diseases emergence and by modeling the associated epidemiological processes, EDEN is intended to support the creation of early warning and disease-monitoring systems by public health agencies to prevent and control vector-borne diseases. Because of the difficulties of producing a map of infected vectors (or even vectors or hosts), many GIS applications have been developed to model habitat (Beck, Lobitz, & Wood, 2000).

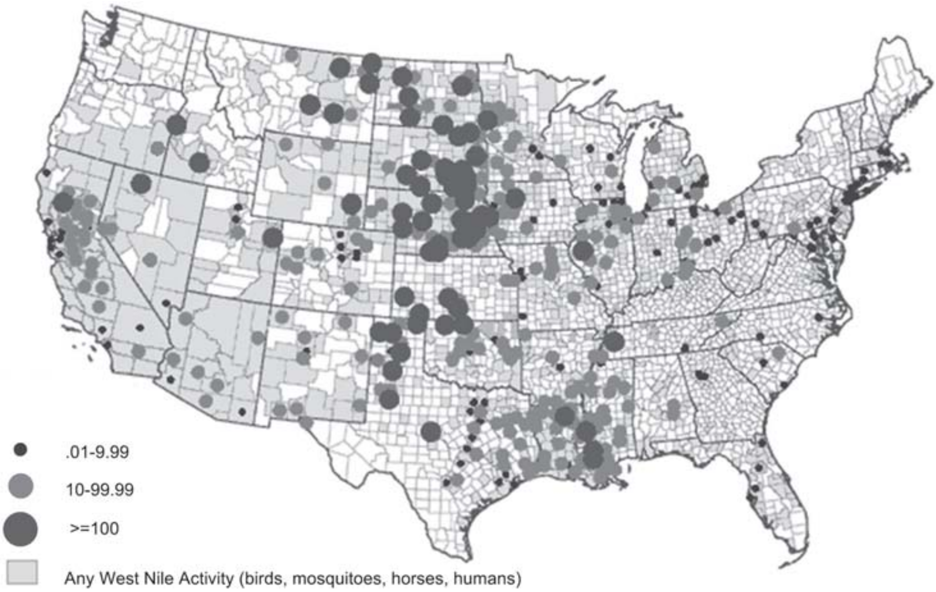


FIGURE 8.10. Data on the incidence of West Nile human neuroinvasive disease per 1,000,000 population by county in 2005 are mapped with data on any West Nile activity from birds, mosquitoes, horses, or humans drawn from the ArboNET surveillance system. From Centers for Disease Control and Prevention (2007b).

Modeling of Vector-Borne Diseases

Environmental modeling of vector-borne diseases is a response, in part, to the difficulties of vector and host surveillance, but environmental and ecological studies are also necessary because they help us to understand the ecological processes involved in vector-borne diseases transmission. Modeling has been used to identify areas where disease is endemic. Understanding these processes and patterns is essential to intervening in effective ways to address vector-borne diseases, including designing meaningful surveillance systems.

GIS analyses have been used to model habitat for mapping the likely distribution of vectors and hosts, evaluating the environmental characteristics of places where human cases, vectors, and hosts are observed, and assessing the likely impacts of global climate change. These efforts are important in helping public health analysts to predict where disease might occur.

Modeling Habitat and Ecological Processes

Using GIS to model habitat can make an important contribution to understanding current and to predicting future patterns of vector-borne diseases. Habitats are significantly affected by land cover change. Residential development or land cover change associated with climate change in an area alters the environment in ways that will either increase or decrease exposure.

An ***ecological niche*** is the entirety of environmental factors in a microhabitat that enables an organism to survive, including the status of the organism within that environment, which affects its survival. Environmental niche models have been developed to predict the distribution and potential spread of disease and invasive species (Guo, Kelly, & Graham, 2005). A study of the potential spread of *A. Albopictus*, which is present in tropical areas, into the United Kingdom relied on studies of environmental factors affecting the distribution and seasonal activity of the vector as described in other northern regions (Medlock, Avenell, Barrass, & Leach, 2006). Strains found in temperate regions are affected by the seasonal change in daylight and are adapted to cold temperatures by eggs that experience a suspension of development. Relevant databases were integrated in a GIS to describe environmental conditions in each 1-kilometer cell of a grid covering the United Kingdom over a 52-week period. The model maps zones where overwintering is possible and where spring hatching can occur and starts. The GIS analysis made can run different scenarios based on the effects of temperature and daylight. The highest chance of establishment was identified as southern England and Wales, especially around London and southern coastal port cities.

Ecological studies are also valuable because they may give insight into factors that affect temporal cycles in infectious disease risk. Acorns from oak trees are an important source of food for the white-footed mouse *Peromyscus leucopus*, which is the principal reservoir in the Lyme disease cycle in the northeastern United States (Jones, Ostfeld, Richard, Schauber, & Wolff, 1998). A large autumn crop of acorns also draws the white-tailed deer into oak forests. Large crops are not produced every autumn, however. Instead, large crops are produced every 2 to 5 years, with few or no acorns produced during the intervening years. Experimental addition of acorns resulted in increased density of mice and of ticks. These experiments suggest that it "may be feasible to predict the risk of contracting Lyme disease from infected nymphal ticks in oak forests on the basis of masting events, with the risk being greater 2 years after an abundant acorn crop" (Jones et al., 1998, p. 1025). Acorn masting was identified as a factor influencing the outcome of control programs of targeted acaricide applications to white-tailed deer (Stafford, Denicola, Pound, Miller, & George, 2009).

Modeling environmental risk based on habitat alone without follow-up data on the distribution of vectors and human cases is problematic. Statistical methods to predict the distribution of species generally require both presence and absence data for calibration, and collection of absence data is often infeasible (Guo, Kelly, & Graham, 2005). To test the adequacy of environmental risk measures, the geographical distribution of risk areas needs to be compared either to the distribution of vectors and hosts or to the distribution of human cases, and preferably both (Mather, Nicholson, Donnelly, & Matyas, 1996; Nicholson & Mather, 1996). GIS analyses have been used to evaluate the environmental characteristics of case locations.

Evaluating the Environmental Characteristics of Case Locations

A number of GIS applications have evaluated the local and neighborhood environmental characteristics of case locations, generally through point-in-polygon analysis (discussed in Chapter 4). Through this type of analysis, it is possible to take a point in one data layer and compare its location to a set of areas in another data layer to determine which area the point lies within (Figure 8.11). Once this determination is made, the attributes of the area can be associated with the point. Points can then be selected on the basis of the attributes of the areas they lie within. Also, once the polygon in which the point lies is determined, the characteristics of adjacent areas can be evaluated and associated with the point (Figure 8.12), including reporting distance to adjacent areas.

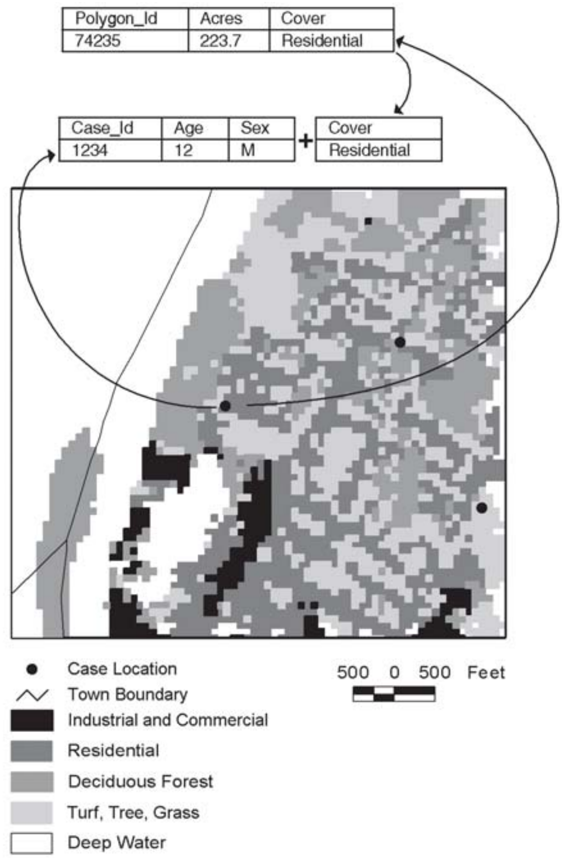


FIGURE 8.11. A point-in-polygon analysis identifies the land cover of the polygon where the case is located and assigns that land cover type as an attribute of the case through a spatial join procedure.

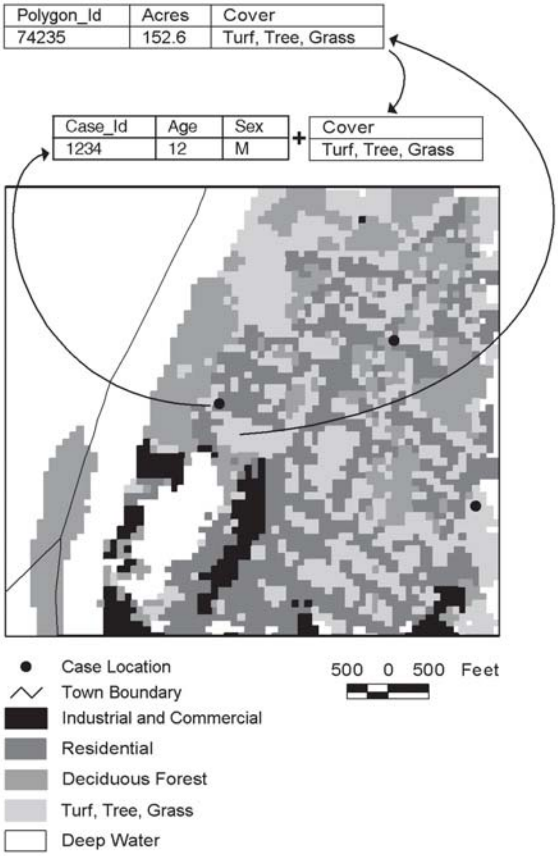


FIGURE 8.12. The land cover of the adjacent polygon is assigned as an attribute of the case.

In this way, the local and neighborhood characteristics of places with a high frequency of cases can be assessed. Once the characteristics of areas where many cases are observed have been determined, it is possible to develop maps showing regions with similar environmental conditions regardless of disease incidence. Maps of this type can be important for suggesting areas where disease may be underreported, where disease incidence may increase if people move in and local environmental conditions remain unchanged or where disease may spread or emerge in the future. Evaluating the environmental characteristics of locations where cases are observed and then comparing them with the characteristics of places where people live or engage in activities but do not acquire the disease can offer additional insights into the disease process.

This approach has been used not only to investigate environmental characteristics of human cases, but also to assess the environments where exposure occurs and the environments where infected vectors may be found. Studies of this type are also important for validating risk indices based on habitat modeling.

ENVIRONMENTAL CHARACTERISTICS OF HUMAN CASES

In the analysis of Lyme disease, case databases have been integrated with other GIS data layers describing environmental characteristics like elevation, vegetation, and soils that may influence tick distribution. The relations between 127 environmental variables and Lyme disease incidence were explored in Baltimore County, Maryland (Glass et al., 1992). Empirical data on tick density and tick infection rates were not included. These data are difficult to collect for large geographic areas. Based on an overlay analysis of the distribution of cases and controls to determine the environmental characteristics of their residential locations, several environmental variables were found to be associated with increased risk, including location in one of two watersheds, on loamy soils, and in forested areas. Residence in highly developed areas was associated with a decreased risk. Only 14% of the land area in the county had environmental characteristics associated with increased risk for Lyme disease; approximately 8% of the county's population resided in these high-risk areas.

A 1992 case–control study of Lyme disease in southeastern Connecticut found that the only variable significantly associated with the incidence of Lyme disease was self-reported residence in a "village" or higher density residential setting. In the study area, these settings are usually surrounded by wooded zones. Villages do not correspond to census tracts or other political administrative units for reporting aggregate population. The hypothesis was tested using GIS (Cromley, Cartter, Mrozinski, & Ertel, 1998). Rather than calculating rates for administrative units, the analysts calculated rates for regions defined by residential density. Geocoded cases acquired by active surveillance were classified as being located inside or outside of a village setting. "Village setting" was operationalized as any area of contiguous medium- or high-density residential development that was at least 30 acres in size. The study area population residing inside and outside villages was also estimated. The analysis confirmed a lower relative risk for people living in villages. A review of research on spatial patterns and human correlates of Lyme disease in the United States found that "the only environmental variable consistently associated with increased LD risk and incidence was the presence of forests" (Killea, Swei, Lane, Briggs, & Ostfeld, 2008, p. 167). Settlement density has also been identified as a risk factor for vector-borne diseases in tropical settings (Laveissiere & Meda, 1999).

Studies of the environmental contexts of human cases have also been conducted in urban settings. Variables measuring natural and anthropogenic landscape characteristics conducive to the transmission of West Nile virus were analyzed for the urban areas of Chicago and Detroit (Ruiz, Walker, Foster, Hara-

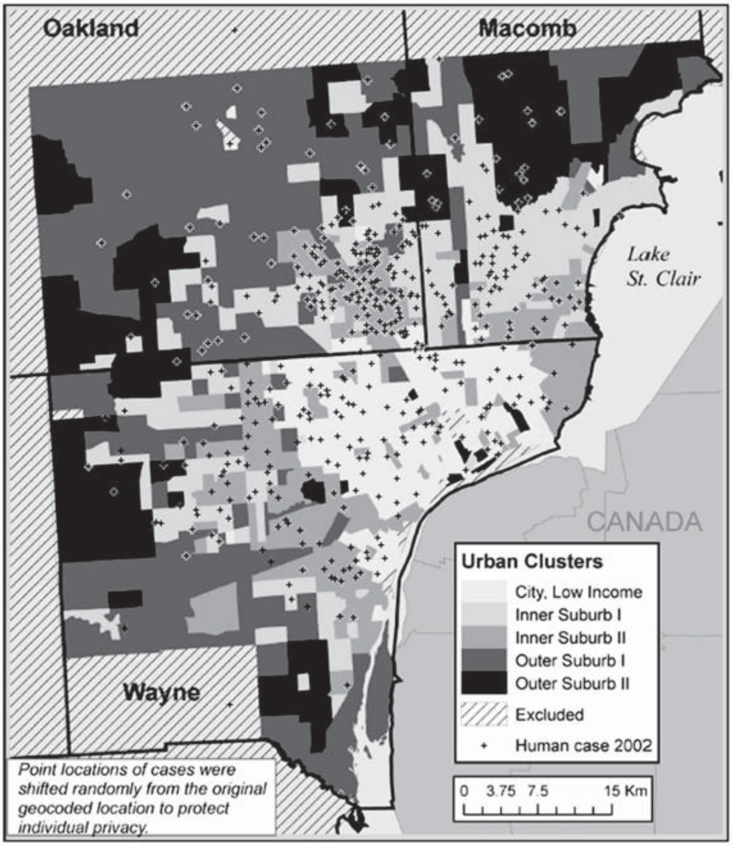


FIGURE 8.13. Human cases of West Nile neuroinvasive disease and West Nile fever reported by the Michigan Department of Community Health, mapped with five urban landscape classes based on factor analysis of housing, vegetation, elevation, land use, and socioeconomic characteristics of census tracts. Incidence rates were highest in inner suburbs in areas of relatively low relief and elevation with less vegetation than in other areas where the housing stock dates to the 1940s and 1950s. From Ruiz, Walker, Forster, Haramis, and Kitron (2007). Originally published by BioMed Central in the *International Journal of Health Geographics*. Open Access.

mis, & Kitron, 2007). The analysis yielded a set of five urban landscape classes mapped for both cities. The distribution of human cases in each city was mapped against the landscape classes (Figure 8.13), and the rate of West Nile per 100 people was assessed by zone. These analyses contributed to the development of hypotheses about West Nile transmission in urban centers in the Midwest, and they also provide a rigorous basis for selecting field sites for avian and mosquito collection.

Environmental Characteristics of Exposure

A similar conceptual approach was used to investigate the epidemiology of rodent bite and the distribution of the rat population in New York City (Childs et al., 1998). Rat bite fever, unlike Lyme disease, is a relatively rare but potentially fatal infection. Few epidemiological studies of rat bite have been conducted. Moreover, there have been very few attempts to describe the distribution of rat or mouse populations in urban centers by direct field observation. Like the data on rabies discussed in this chapter which was based on cases of human contact with animals that might have rabies, data on where people come into contact with rats can be analyzed to investigate the environmental contexts of contact. In the New York study, a GIS was developed to determine the city block where the bite occurred and the environmental and social characteristics of the blocks (Figure 8.14). A set of control blocks where no bite had been reported was selected for comparison.

GIS analysis was used to produce maps of the distribution of blocks with different probabilities for rodent bite in each of the five boroughs. The predictive power of the maps was evaluated by comparing reports of rodent bite from a later year with the maps and by environmental sampling of randomly selected blocks in Manhattan and Brooklyn for evidence of rat infestation. Blocks where rodent bites occurred tended to be closer to subways, railroads, and parks. These places

Bite Characteristics
Bite ID
118

+

Census Block Characteristics				
Block ID	Total Pop	Pop Density	% Black	% Hispanic
202	1259	135866.4	45.0	74.0

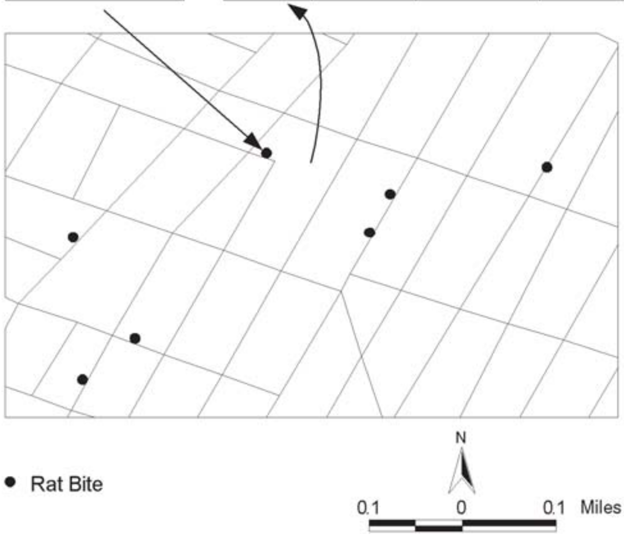


FIGURE 8.14. A point-in-polygon analysis to identify the local census block characteristics of a rat bite case and assign them as attributes of a case through a spatial join procedure.

are potential sources of exposed ground, providing rats with burrow sites and sites where human refuse accumulates and where rats might forage.

For diseases like rabies or rat bite fever, an individual is likely to be aware of the circumstances of exposure. Even for tick-borne diseases, individuals may find a tick on their bodies and be able to report when and where they acquired it. For other diseases, like those borne by mosquitoes or acquired through contact with water, exposure is more difficult to pinpoint, and geospatial technologies are being used to investigate exposure.

Schistosomiasis or snail fever, also known as bilharzia or bilharziosis, is a parasitic disease caused by several species of worms of the genus *Schistosoma*. Snails are natural reservoirs of the agent, and humans acquire the disease by wading or swimming in bodies of water that are infested with infected snails. When infected humans urinate or defecate in freshwater, the water body is contaminated with *Schistosoma* eggs. When the eggs hatch, they release miracidia, worms in the larval stage. The miracidia penetrate the snails where they form sporocysts. Germ cells from the sporocysts divide to produce cercariae, the larvae capable of infecting humans. Cercariae emerge from infected snails on a daily basis. When they come into contact with human skin in the water, they can attach to and penetrate the skin. The parasites mature in 6 to 8 weeks and then begin producing eggs that are shed by humans into freshwater sources by urination and defecation.

The mortality rate is low for schistosomiasis. It is primarily a chronic disease that affects the growth and development of children and causes abdominal pain, cough, diarrhea, and fatigue in adults. The disease can also cause genital sores that may persist even after successful treatment of the schistosomiasis infection and may increase HIV transmission (Hotez, Fenwick, & Kjetland, 2009). Although schistosomiasis is not found in the United States, it is endemic to more than 70 countries, and people who travel can import the disease to areas where it has not been or once was endemic (Meltzer et al., 2006).

Assessing human exposure to infectious cercariae has traditionally been accomplished by direct observation or through self-reports. In the same way that researchers are using personal monitoring devices to assess exposure to environmental contaminants (Elgethun, Fenske, Yost, & Palcisko, 2003) as discussed in Chapter 6, devices for monitoring activity patterns in time and space are also being used to investigate exposure to vector-borne diseases. A study carried out in a village in Sichuan province in China, where schistosomiasis is endemic, provided GPS receivers—worn in vests borrowed from a children's pesticide exposure study—to a random sample of 12 people who provided informed consent to monitoring of their activities (Seto, Knapp, Zhong, & Yang, 2007). The individuals wore the vests for 98 hours during the day on two separate days. The GPS data could only show where and when individuals spent time during the day, so the number and locations of water contacts could only be determined from interviews. Maps of the individuals' activity patterns can be an aid in recall of these patterns of contact. Furthermore, the study yielded important insights into the participants' high level of mobility. It suggested that agricultural labor was shared across villages and that water contact probably occurs on a regional and not just a local basis. These exchanges could sustain transmission of the disease and make control efforts more difficult.

Environmental Characteristics of Infected Hosts

Hantavirus pulmonary syndrome is a severe illness affecting the cardiovascular system and resulting in death in slightly less than half the cases (Centers for Disease Control and Prevention, 1999). The most frequently seen agent of the four different hantaviruses in North America is Sin Nombre virus, transmitted to humans from the deer mouse *Peromyscus maniculatus* by direct contact with infected mice, their droppings, or their nests, or by inhalation of virus particles from mouse excrement. Residence in a dwelling with substantial rodent infestations has been identified as an important source of exposure.

Environmental data from remote sensing and GIS analysis have been used to predict host infection status for the Sin Nombre virus (Boone et al., 2000). Vegetation type and density, elevation, slope, and hydrology were characterized for 144 field observation sites in the Walker River basin along the Nevada–California border. Deer mice were trapped at field sites, and blood samples were taken to determine current or past infection with the virus. Field sites were classified as positive or negative based on the infection status of the mice captured. Discriminant analysis was then used to examine relationships between the environmental variables characterizing each site and the site's infection status. Combinations of environmental variables were found that could correctly predict the infection status of deer mice with 80% accuracy.

In reviewing the literature on studies looking at spatial patterns in the environmental correlates of disease, specifically Lyme disease, Killea et al. (2008) found that studies of the environmental correlates of disease do not always produce consistent results. This is also true in studies of the influences of neighborhood characteristics on health in the health disparities literature discussed in Chapter 11. These inconsistencies are not necessarily the result of differences in data collection and methods. Disease processes themselves are likely to be ***spatially varying processes***, in which relationships are conditioned on parameters defined by spatial dependencies. It is not just a question of documenting geographical variation in soil moisture, for example, as it explains the geographical distribution of a vector. The relationship between soil moisture and vector populations may itself be spatially variable. Progress in understanding the determinants of spatial variation in disease risk and incidence depends on incorporating knowledge of the biology of individual components of regionally variable disease systems, collecting data over longer periods of time, achieving greater standardization of data collection and analysis across regions, and testing the effect of the same environmental variables at multiple spatial scales (Killea et al., 2008). The importance of working across multiple scales in studies of vector-borne diseases is highlighted by two studies, one of Chagas disease in Argentina and another of schistosomiasis in Kenya (Kitron et al., 2006). "Spatial heterogeneity on the micro-scale may not be detected using coarse spatial resolution, and conversely, general patterns on the macro-scale may not be detected using fine spatial resolution" (p. 49).

Modeling and Mapping Prevalence: The Malaria Atlas Project

The Malaria Atlas Project was initiated in 2006 to advance the science of malaria cartography (Hay & Snow, 2006). Mapping malaria is challenging because transmission intensity is geographically heterogeneous. From the outset, the intent was to develop a global map of malaria, to apply strict inclusion criteria for prevalence rate reports (reports of the proportion of a sampled population that is confirmed positive for malaria parasites), to collect data on both *Plasmodium vivax* and *P. falciparum*, and to place the final peer-reviewed database into the public domain (Malaria Atlas Project, 2011).

The project focused on developing a global continuous *P. falciparum* malaria endemicity surface for the year 2007 (Hay et al., 2009). Four main steps were involved in meeting this objective. First, analysis searched for and processed prevalence rate reports (Guerra et al., 2007). These reports were georeferenced. Second, a prevalence rate database was used to make a continuous, age-standardized, urban-corrected malaria prevalence surface using geostatistical modeling techniques. Third, validation procedures were applied to assess the accuracy of the endemicity predictions and the uncertainty associated with them. Fourth, populations at risk were estimated.

Geostatistical models, including models that incorporate time, are being used to predict and map vector-borne diseases prevalence. Expanding on a modeling approach developed and applied by Diggle and others (Diggle, Tawn, & Moyeed, 1998; Diggle, Moyeed, Rowlingson, & Thomson, 2002), the MAP researchers worked with a global database of almost 8,000 survey reports to model a continuous endemicity surface on a 5×5 kilometers grid based on initial work to define the global spatial limits of malaria transmission (Hay et al., 2009). The methodology allowed for the use of Bayesian methods in statistical inference. Model-based geostatistics are useful because they allow analysts to address uncertainty in different stages of the modeling process. In addition to predicting the endemicity at each location, the method provides a measure of confidence that can be associated with each prediction.

In assembling the prevalence reports for the mapping project, the researchers noted an increasing tendency for national surveys to be conducted so that they would be representative of all areas within a country, not just areas of high prevalence. There were also many zero prevalence values recorded in the reports analyzed (Hay et al., 2009). The prevalence rate is acknowledged to be a less direct measure of malaria transmission than other measures like the ***entomological inoculation rate,*** which is the number of infective bites per capita over a time period, or the ***basic reproduction number*** or ***basic reproductive rate,*** R_0, which is the mean number of secondary cases that a single typical infected case will cause in a population with no immunity in the absence of interventions to control the disease.

Researchers used a combination of data provided by the source of reports, an online database of geographic names, online gazetteers, and paper maps to create lon/lat point references for the reports. Surveys that could not be georeferenced or that could be georeferenced only to larger areas (greater than 25 km^2) were excluded (Guerra et al., 2007). The data were recorded using decimal degrees with a precision of no less than 2 decimal places. The administrative regions in which the points were located were also coded. Surveys were also classified as urban, peri-urban, or rural depending on their location with respect to areas of population density based on data from the ***Gridded Population of the World*** version 3 (Center for International Earth Science Information Network, Socioeconomic Data and Applications Center, 2011).

The georeferenced, validated, and age-standardized reports were used to model the endemicity surface (Hay et al., 2009). The prediction of the prevalence rate at a given unsampled point is dependent on the spatial distribution of survey points around the unsampled point, the spatial variation evident in the values for sampled points, and the number of people sampled in each survey referenced to a sampled point. Heterogeneity in space was modeled using semivariograms as discussed in Chapter 6. The temporal structure of the data was also taken into account. The reports covered different times throughout a period from 1985 to 2009. A report was referenced temporally by the midpoint in decimal years between the start and end months of the report. The final database was stratified into three global regions: the Americas, Africa (including Yemen and Saudi Arabia), and Central, South, and East Asia.

For each region, a Bayesian geostatistical model was developed to predict the value of the prevalence rate in the 2–10 cohort in 2007. Distances were measured as spherical distances, described in Chapter 9. At the scale of the analysis, the earth's curvature needs to be taken into account in distance calculations. The predictions were made at points on a regular 5 × 5 kilometer grid of the areas earlier modeled to be within the spatial limits of stable *P. falciparum* transmission.

Once the prevalence rate map was created and the model was validated by assessing its ability to predict known values of the prevalence rate at reported locations, the map of prevalence rates was compared with a population density surface to assess population at risk. Data from the ***Gridded Rural–Urban Mapping Project*** (GRUMP) (Center for International Earth Science Information Network, Socioeconomic Data and Applications Center, 2011) for 2000, adjusted to the United Nations' national population estimates, were projected to 2007. To assess risk to children, the national population counts were stratified by age to obtain 0–4, 5–14, and less than 15 years population cohorts. Areas with different levels of endemicity from the model of prevalence rates were overlaid with the GRUMP data. Approximately 1.4 billion people were estimated to live in stable risk areas for *P. falciparum* transmission. In the American and Asian regions, a third of children 0–4 and 5–14 years of age were exposed; in the African region, 43% of children were exposed.

Although the mapping effort was based on prevalence reports, data on the entomological infection rates and basic reproductive rate can be used to inform the modeling process (Hay & Snow, 2006). These measures are metrics of the ***force of infection,*** the rate at which susceptible individuals acquire an infection. A study of the entomological inoculation rate and *P. falciparum* infection in children in Africa found evidence of heterogeneous biting and heterogeneous susceptibility (Smith, Dushoff, Snow, & Hay, 2005). Some people are bitten more than others, and some people are more susceptible to infection each time they are bitten. The study found that 20% of people received 80% of all infections. This heterogeneity has implications for studying disease transmission and for assessing the effectiveness of vaccines.

By mapping the survey report data and providing open access to the methods used to collect, analyze, and map the data, the project is creating a framework for ongoing collection and analysis of malaria on a global scale. Because the researchers have provided so many people with a way of viewing these data in time and space, there is an increased incentive to improve surveillance and reporting. The 2007 global malaria endemicity map is the first of a planned series designed to make it possible to monitor and evaluate the progress of efforts to control and eliminate the disease.

Modeling Immunity to Disease

An emerging area of research in disease surveillance is modeling immunity to disease. The level of immunity within a population is an important factor affecting disease transmission, and it must also be taken into account in designing programs to control the spread of disease. Researchers reanalyzed data from a killed cholera oral vaccine trial conducted at a research site in Bangladesh in 1985 (Ali, Emch, Yunus, & Clemens, 2009). Outbreaks of cholera occur when *Vibrio cholerae*, the bacterium that causes the disease, is sufficiently present in drinking water or in food to provide an infective dose if ingested. Cholera causes acute diarrheal illness which, if of sufficient severity, can be fatal. The source of contamination of water bodies is feces from an infected person. Sewage and water treatment systems are effective in controlling cholera, along with measures for preventing food-borne disease.

The analysis conducted in Bangladesh integrated the database from the vaccine trial with a longitudinal health and demographic database of the study area population, including data on hospitalization for cholera and a spatial database of the study area. Specifically, the analysis focused on whether disease transmission changed after the mass vaccination campaign and whether the vaccine had a greater impact in areas with a lower force of infection.

Spatial autoregressive lag models of the cholera hospitalization rate were estimated using GeoDA. One model was for the prevaccination period and one for the postvaccination period. Although cholera hospitalization in the study region increased after the vaccination period due to temporal fluctuations in cholera in Bangladesh, villages with higher vaccine coverage had significantly lower hospitalization rates for cholera. Changes in the locations of clusters of high cholera hospitalization between pre- and postvaccination periods support the hypothesis that spatial differences in vaccine coverage can change the spatial structure of the disease. The impact of the vaccine was affected by the force of infection, with a steeper descending trend in cholera hospitalization in areas with low force of infection than in areas with high force of infection in the post-vaccination period. Efforts to control the spread of disease through methods like vaccination are occurring at a time of global climate change.

Global Climate Change and Vector-Borne Diseases

GLOBAL CLIMATE CHANGE

In addition to biological and ecological determinants of vector-borne diseases, climate factors have been studied as influences on the worldwide spread of infectious diseases. Scientists have identified upward trends in global mean temperatures, sea level, and ocean heat content over the last 50 years (Patz & Olson, 2008). Although a wide range of health outcomes are sensitive to climate including heat-related morbidity and mortality (Curriero et al., 2002), there is concern that climate change could directly influence vector-borne diseases transmission by affecting the vector's geographic range, increasing rates of reproduction, affecting biting behavior, and shortening incubation periods of the pathogen. Equally significant indirect effects of climate could occur through land cover changes affecting microclimates. The effects of climate change on the occurrence and prevalence of disease in livestock have also been considered (Gale, Drew, Phipps, David, & Wooldridge, 2009). A growing body of literature is investigating the connections between climate and infectious disease using climate change models, remote sensing data, and GIS.

GEOGRAPHIC RANGE OF VECTORS

Due to the expansion of the range of *I. scapularis*, Lyme disease is an emerging disease in central and eastern Canada (Ogden et al., 2008). Risk maps for range expansion of the vector based on current conditions and with climate change were developed for census subdivisions in the region. Ambient air temperature, habitat, and numbers of ticks immigrating on migratory birds were selected as components of a risk model. Remote sensing data from the ***Advanced Very High Resolution Radiometer (AVHRR)*** of the National Oceanic and Atmospheric Administration's satellite were used to model forest cover as a measure of habitat. A map of U.S. counties with populations of *I. scapularis* was used with knowledge of distances that passerines fly per day to develop a measure of larval tick immigration. The index of tick immigration was validated in the field at sites in southern Quebec. Predicted increases in temperature are expected to increase bird migration speed in North America, and the algorithm was used to predict the expansion of the range under changing temperature conditions.

TRANSMISSION AND BITING

The dynamics of vector-borne diseases transmission associated with specific climate events like the El Niño Southern Oscillation (ENSO) have also been investigated using climate modeling techniques (Lemon et al., 2008). The ENSO of 1991–1992 has been suggested as a major climate factor in the southwestern United States, creating favorable conditions for an increase in the rodent population, thereby leading to the outbreak of hantavirus pulmonary syndrome. Springtime precipitation in 1992 and 1993 at 28 sites with confirmed human cases of hantavirus pulmonary syndrome and 170 control sites was estimated and compared to precipitation during the previous 6 years (Glass et al., 2000). Elevation and Landsat Thematic Mapper data collected the year before the outbreak were also used to estimate disease risk. The study showed an association between elevation and satellite data and hantavirus pulmonary syndrome risk the following year. There is a clear need for long-term studies to identify trends resulting from periodic phenomena like ENSO (Calisher et al., 2005).

Many vector-borne diseases processes show seasonal variation along with large within-year variation of incidence. Variability in weather patterns has been investigated for possible associations with variability in entomological parameters like biting rates. A soil moisture model of surface water availability combined with land cover and soil features improved prediction of biting rates for two *Anopheles* mosquitoes associated with malaria outbreaks in an endemic region of Kenya (Patz et al., 1998). Modeling soil moisture and lagged soil moisture substantially improved prediction of variability in bite rates over the predictions made from modeling rainfall alone.

Environmental Impacts of Controlling Vector-Borne Diseases

Vector-borne infectious disease transmission depends on ecological systems that may be complex—involving more than one agent, vector, or host—and regionally variable. This complexity is reflected in efforts to control vector-borne diseases. In some cases, efforts are made to prevent the introduction of disease into a new area by restricting migration and trade. In the case of West Nile virus, for example, international restrictions on the movement of horses were put into effect. Attempts to control a disease once it has become established in an area can conflict with efforts to protect the environment.

Failure to control disease can directly affect the health of wild animal populations. During the rabies epizootic in the 1980s and 1990s in the northeastern United States, for example, more than 20,000 rabies cases were reported in raccoons (Bretsky, 1995). For some diseases, like Lyme disease, direct effects of the disease in wildlife populations have not been reported, and animal populations are significantly impacted not by the disease itself, but by control measures that affect their habitat. Conflicts also arise when wildlife hosts seed epidemics of disease in domesticated animals like cattle that can then pose a threat to human health. An example is bovine tuberculosis (TB) in Great Britain. The likeliest source of infection of cattle is the badger, although this has not been fully proved (Krebs et al., 1998). Natural habitat for badgers is often found in or around cattle pasture areas. Different strategies ranging from severe culling to only partial removal of badgers have been implemented for the last two decades to control bovine TB. Because these strategies were implemented in succession rather than in parallel, it has been difficult to assess their relative effectiveness.

Controlling disease in livestock is also challenging from a public health perspective. Global changes in poultry production and transportation, especially in Asia, contributed to the outbreak of H5N1 avian influenza in southern China in 2003 (Sims, 2007). Backyard poultry production was also identified as an important but overlooked factor in the outbreaks in Nigeria in 2006, the first outbreaks reported in Africa (Cecchi, Ilemobade, Le Brun, Hogerwerf, & Slingenbergh, 2008).

TABLE 8.1. Intervention Options for Vector-Borne Disease Control and Potential Environmental Impacts

Control method	Environmental effects
Self-protection precautions	Negligible; possible health effects of vaccines, repellents on user
Habitat manipulation	Powerful effects in areas where habitat disrupted (can be limited to areas with high human presence)
Manipulation of host populations	Powerful effects on host species and associates (efficacy not always established)
Manipulation of vector population genetics	Vector competence may contribute little to force of transmission; new pest species might emerge
Biological control	Depends on species utilized (efficacy not always established)
Broadcast pesticide applications	Powerful effects on nontarget species in application areas
Targeted pesticide applications	Main effects confined to nest associates of targeted species (can be limited to areas with high human presence)

Note. Adapted from Ginsberg (1994, p. 347). Copyright 1994, reprinted by permission of John Wiley & Sons.

Seven control methods for dealing with vector-borne infectious disease problems have been identified (Table 8.1). Only self-protection measures can be associated with minimal environmental impacts. Efforts to control vector-borne infectious disease by intervening in an ecological system are bound to have impacts on other aspects of the environment and may produce unintended, undesirable consequences. The ethical, legal, and social implications of genetically modifying vectors to make them less competent are of great concern (Spielman, 1994; Macer, 2005). This approach seems out of step with the rich and broadly focused research on vector-borne diseases being carried out at a variety of scales around the world and the efforts at disease control this type of research is beginning to support (Childs, 2009).

A major unanswered question in vector-borne diseases control is how much reservoir or vector populations have to be reduced before there is an impact on human health. Total elimination is usually out of the question, but partial elimination may be ineffective: lowering vector abundance or agent prevalence may not produce equivalent declines in human exposure risk (Ginsberg, 1993) or may instead exacerbate the infectious disease problem by disrupting territorial systems and decreasing diversity (Krebs et al., 1998). When ecosystem diversity decreases, the disease transmission cycle may actually become more efficient (Ostfeld, 2009).

A Syndemic Perspective on Disease

The resurgence of interest in vector-borne diseases no doubt reflects their emergence or reemergence in populations and areas of the world where it was believed these problems had been conquered. Vector-borne diseases have remained "common among populations lacking basic human rights such as control over their land, political rights and access to water and sanitation" (Winch, 1998, p. 47). The global distribution of many of the same diseases—cholera, malaria, encephalitis—that are now the subject of interest was studied in the late 1940s as part of a disease atlas project supervised by Dr. Jacques May under the auspices of the American Geographical Society (American Geographical Society, 1944).

The research on climate change and vector-borne diseases does not suggest that increases in vector-borne diseases risk can be attributed to climate trends alone. Although vector populations are sensitive to climate trends, there are multiple factors underlying the emergence and reemergence of vector-borne diseases. Housing conditions, public health resources, and access to medical care are among the factors that are likely to influence the emergence of vector-borne diseases even in areas where the potential for transmission may be increasing in response to climate change (Martens, 2000).

By examining temporal and spatial patterns, a study of the increase in tick-borne encephalitis in the Baltic region concluded that climate change alone could not account for the rapid increase in the incidence of the disease over the last two decades (Sumilo et al., 2007). Tick-borne encephalitis is caused by a virus transmitted in the region principally by *I. ricinus* and *I. persulcatus* among mice of the genus *Apodemus*. Tick-borne encephalitis and Lyme disease, vectored by the same species, are among the most significant vector-borne diseases in Europe. Systematic surveillance of tick-borne encephalitis has been conducted for 30 to 50 years within the region.

Changing patterns of incidence at the county level were investigated for the period 1970–2006. Dekadal (10-day) mean daily minimum and maximum temperatures were analyzed because this time scale is appropriate to onset of tick activity in the spring. No spatial pattern in the overall pattern of temperature change during spring, or any other time of the year, was found that could be related to the variable patterns of tick-borne encephalitis incidence observed in each county in each time period. The analysis also investigated spatial and temporal patterns in rainfall and snow cover. These could not be related to spatial and temporal patterns of change in tick-borne encephalitis incidence.

The results indicated that climate change observed over the last 35 years has contributed to a more permissive background for transmission of the disease. Important human-induced changes in the environment also took place during the study period. The collapse of industrial production in the region in the post-communist period significantly reduced environmental pollution and increased solar radiation input. At the same time, dislocations in the agricultural sector led to changes in land cover from fields and pasture to woods, with probable effects on rodent abundance.

Human exposure to infected ticks and acquisition of disease depend on a wide range of environmental factors. These factors may be abiotic (climate and landscape), biotic (tick and host distributions and abundance), and socioeconomic (human behavior and access to vaccines). Biological factors determining virus transmission potential are the "submerged bulk" of a zoonotic iceberg (Sumilo et al., 2007). Socioeconomic factors cause an increase in relative exposure and work with the underlying factors in the spread of disease. These factors are linked in a hypothetical model of tick-borne encephalitis epidemiology in the study region (Figure 8.15).

This research and other studies referenced in this chapter are representative of an emerging approach to the study of health: the syndemic perspective (Baer & Singer, 2009). The ***syndemic*** perspective is distinguished by three premises: the complex interactions among comorbid health conditions in a population, especially the biological linkages; the social and environmental conditions that interact to promote the emergence and transmission of disease, especially as they lead to disease clustering; and the environment as an encompassing biocultural environment, wherein the natural environment is not separate or independent from human action. In this view, "diseases do not simply coexist with other diseases in overlapped space within the same population, the bodies of their individual members, or the organs of individual sufferers. Rather, diseases interact synergistically in substantial ways that impact the health of the individuals and populations they infect" (Baer & Singer, 2009, p. 135).

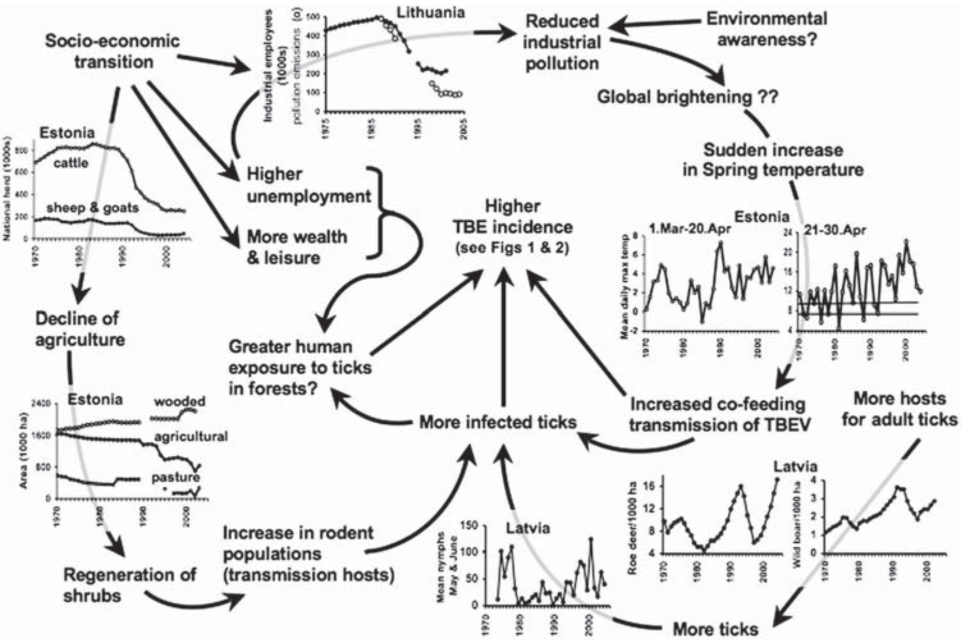


FIGURE 8.15. A hypothetical model of the epidemiology of tick-borne encephalitis in the Baltic region showing factors contributing to the emergence of disease. From Sumilo et al. (2007). Originally published in *PLOS One*. Open Access.

Conclusion

Emerging and reemerging vector-borne infectious diseases are challenges requiring new responses from public health and medical care systems. These diseases are often undiagnosed, untreated, and unreported, a situation of special concern because delays in diagnosis and treatment often result in severe chronic health problems or death. Ecological studies of agent–vector–host relationships and improved surveillance methods have been cited as important priorities for addressing these infectious disease problems.

A view of key tasks necessary to reduce the burden of vector-borne diseases is emerging (Scott & Morrison, 2008). Research is needed to develop methods for assessing risk of disease transmission that are operationally feasible and epidemiologically effective and to set goals for disease prevention. Sensitive, specific, and inexpensive ways to estimate herd immunity—to specific serotypes depending on the disease—are needed because immunity affects epidemic transmission. The use of vaccines to elevate immunity artificially can be coordinated with vector control so that effective vector control efforts can be sustained over time. Methods for reducing vectors and reducing human contact with vectors that are specific to the needs and possibilities in particular community settings are needed (Spielman, 1994). Field-based prospective longitudinal cohort studies in endemic locations are needed to develop and assess locally adaptive interventions based on entomologic and epidemiologic risk and disease incidence and severity.

GIS analysis is playing an important role in the renewal of efforts to view the problems of vector-borne diseases at a variety of geographic scales, including the global scale. As many of these studies point out, however, the important issue from a human health perspective is how our better understanding of the disease process leads to better prevention and intervention and improves access to health services.

CHAPTER 9

Analyzing Access to Health Services

Fiscal and administrative pressures are transforming health care delivery in the United States. Changes in technology, shifts in medical practice, and the ever-present pressure to contain health care costs are reshaping how health care is provided, where, and for whom. Millions of Americans lack health insurance (an estimated 43.8 million in 2008), including 9% of all children (Cohen & Martinez, 2009), and the population is becoming more diverse in terms of class, culture, and ethnic background. These changes are having profound effects on access to health services. Some health care facilities are closing their doors, others are relocating or expanding, and most are offering different types of services, in different settings. Moreover, despite the rhetoric of choice, health care access is increasingly regulated by health insurers and managed care providers and constrained by lack of insurance coverage. This chapter discusses the use of GIS to analyze access to health services in this dynamic context. We consider the role of GIS in providing and managing information about health service locations, the measurement of geographical access to services, and the analysis of changing service distribution patterns.

The aim of health services is to improve health and well-being. Although we typically think of biomedical health service providers such as physicians and hospitals, a much broader array of activities contribute to health, including education services, water supply and sanitation facilities, mental health care, and social services. There are two general ways of providing health care. ***Informal health care*** is care provided by families and communities in a home or community setting (Moon & Gillespie, 1995). The vast majority of health care is provided informally. Informal care is neither monetized nor assigned a value through market mechanisms or budgeting processes. Women provide well over half of informal health care delivered in the United States and worldwide (Timyan, Griffey Brechin, Measham, & Ogunleye, 1993; Navaie-Waliser, Spriggs, & Feldman 2002).

In contrast, ***formal health care*** is care provided by public, private, and voluntary organizations, through providers such as hospitals or physicians. Formal care takes place in a variety of settings, including clinics, workplaces, schools, and, increasingly, individuals' homes. In the formal sector, carers typically receive a monetary wage for their services, and government regulation of services is common. This chapter focuses primarily on formal health services; however, there are important links between the two types of health care that can be examined geographically. Changes in the intensity and structures of formal health care affect the need for informal health services, and vice versa. In many countries, hospitals are sending patients home earlier, changing the locus of patient care from formal to informal settings.

Access

Access is a multidimensional concept that describes people's ability to use health services when and where they are needed (Aday & Anderson, 1981). It describes the relationship between attributes of service need and the characteristics of service delivery systems. Penchansky and Thomas (1981) identify five important dimensions of access. ***Availability*** defines the supply of services in relation to needs: Are the capacity and types of services adequate to meet health care needs? ***Accessibility*** describes geographical barriers including distance, transportation, travel time, and cost. It highlights the geographical location of services in relation to population in need. ***Accommodation*** identifies the degree to which services are organized to meet clients' needs, including hours of operation, application procedures, and waiting times. ***Affordability*** refers to the price of services in regards to people's ability to pay. Income levels and insurance coverage are critical aspects of affordability. Finally, ***acceptability*** describes clients' views of health services and how service providers interact with clients. Acceptability encompasses barriers linked to gender, culture, ethnicity, and sexual orientation that affect willingness to use particular health services and the sense of comfort and satisfaction in receiving services. Services are acceptable if clients are well treated and satisfied, if providers and clients communicate openly, and if clients are confident about the quality of care delivered.

Geographical Accessibility

GIS necessarily emphasize ***geographical*** or ***spatial accessibility***, the geographical dimensions of access. People's access to health services is rooted in their daily activity patterns in time and space. The framework of time geography, discussed in the Introduction, offers important insights into individual health care decision making. With the home as a base, people move about in space to conduct various activities including work, school, shopping, and the care of children or elderly dependents. These activities form an ***activity space***, a geographically defined zone typically centered on the home, within which everyday life unfolds. Access to health care is both an important component of and constrained by activity spaces. Such spaces emerge from people's socially defined roles and responsibilities that reflect the needs of the individual and the household and community context of which she or he is a part. For many women, the "double day" of paid work outside the home and domestic work in the home reduces the time available for taking care of personal health care needs. Women with limited time, resources, and access to transportation may choose to neglect their own health care while prioritizing the needs of others around them (Young, 1999). These time–space constraints clearly have a central role in shaping geographical access to health care.

After making a decision to utilize health care services when a perceived need exists, an individual must choose a health care provider. In making that choice, the person weighs the advantages and disadvantages of alternative providers, who typically are located in or near his or her activity space. Provider characteristics including quality, availability, and cultural appropriateness of care also come into play, as do barriers such as insurance coverage (Rosenberg & Hanlon 1996). The alternative that best satisfies perceived health care needs, within the time–space constraints of daily life, is often chosen. When aggregated together, these individual choices form spatial patterns of health care utilization—the flows of people over space to health services.

A fundamental aspect of health care utilization patterns is ***distance decay***, or the tendency for interaction with service facilities to decrease with increasing distance (Figure 9.1). For a wide range of services, including many types of

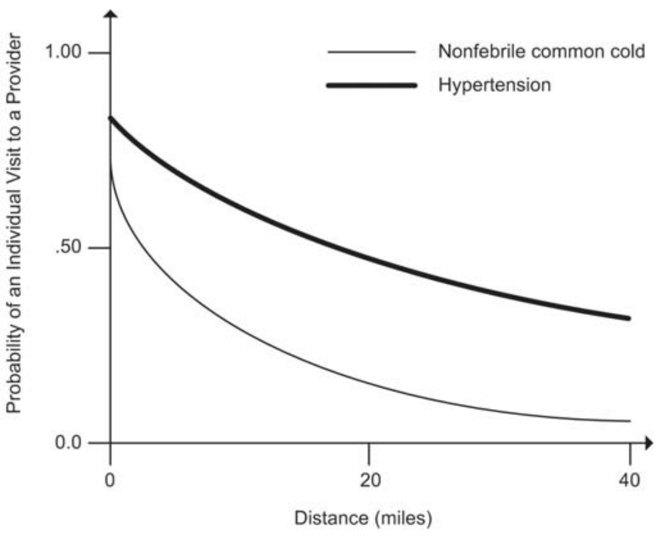


FIGURE 9.1. Distance decay in the utilization of health services. The frictional effect of distance varies, depending on the severity of the health concern for which an individual seeks care.

health services, we find that utilization decreases as distance increases. Studies in a variety of contexts, for different types of health services, confirm the significant effect of distance on utilization and its persistence after controlling for age, illness, and other known risk factors (Joseph & Phillips, 1984; Seidel et al., 2006; Hiscock, Pearce, Blakely, & Witten, 2008). Distance decay is a consequence of the added time, cost, and effort of traveling long distances. As costs increase, the ability and willingness to travel decrease. People's knowledge of and familiarity with service opportunities also decline with distance, exacerbating the pattern of distance decay. In addition, physicians may be less willing to refer patients who live far from health facilities for treatment (Lin, Allan, & Penning, 2002).

The frictional effect of distance varies among health services. Studies reveal a pronounced decline in utilization with distance for hospital-based elective and psychiatric procedures, even after controlling for medical need. For example, a recent study in Britain found that eligible women living far from a plastic surgery center were less likely to undergo breast reduction surgery than were eligible women living nearby (Nair, Richardson, Thompson, Shortt, & Stewart, 2009). In contrast, acute emergency procedures show little or no distance decay (Joseph & Phillips, 1984; Haynes, Bentham, Lovett, & Gale, 1999). Goodman, Fisher, Stukel, and Chang (1997) found no decrease in utilization with increasing travel time for conditions in which there is strong medical consensus on the need for hospitalization, but significant decreases with distance for conditions in which outpatient treatment is a reasonable alternative. Thus, the severity and urgency of the health episode and medical practice decisions about how and where such episodes should be treated all play a role in distance decay. The policy implications of distance-related differences in health care utilization are complex. Are rates of utilization excessive among those living near health facilities, or do those living far away forgo the use of essential services? Do individuals distant from services rely more on informal care or on formal home-based care? Regardless, geographical access has significant and varying effects on health care utilization patterns.

The role of geographical accessibility in service utilization also depends on population characteristics. People differ in their abilities to overcome distance and in how locational constraints affect service use. Travel for health care is strongly affected by demographic and socioeconomic characteristics such as income, occupation, age, and gender. Research indicates that people whose mobility is limited by low incomes, age, or poor access to transportation are more sensitive to distance and thus more likely to use the nearest health care provider or to forgo care altogether (Haynes & Bentham, 1982; Arcury, Preisser, Gesler, & Powers, 2005). A study in Savannah, Georgia, found that distance is a more important factor in health care-seeking behavior for inner-city residents than for those living in the suburbs or on the urban fringe (Gesler & Meade, 1988). Insurance coverage confounds these relationships. Uninsured patients may bypass nearby hospitals or physicians because the services are not affordable or do not accept people without insurance.

Socially and culturally defined roles and activities complicate the ties between geographical proximity and health care utilization. Proximity to workplaces and other sites that are important in people's daily lives plays an important role in health care decisions. These time–space activity patterns vary by gender, stage in life cycle, and socioeconomic position (Kwan, 1999). Complex and geographically dispersed daily routines have been observed even among highly disadvantaged populations (Matthews, Detwiler, & Burton, 2005). Cultural differences come into play as well. For immigrant groups, language and cultural barriers inhibit utilization of health services, even when those services are geographically accessible (Dyck, 1990). Chinese immigrants in Toronto had a strong preference for Chinese-speaking family physicians regardless of location (Wang, Rosenberg, & Lo, 2008). Perceptions of place and location, and the meanings attached to them, vary through time and space, affecting peoples' willingness to seek out particular kinds of services and health care providers (Kearns, 1993). Thus, social and geographical dimensions of accessibility are closely intertwined.

Characteristics of the local environment also affect the role of distance in people's use of health services. Crime, lack of safety, and environmental pollution can inhibit people's use of health services even when services are located nearby. Tarlov and others (Tarlov, Zenk, Campbell, Warnecke, & Block, 2009) used GIS to integrate data on distance to mammography screening facilities, local crime rates, and stage at diagnosis for breast cancer patients in Chicago. Statistical analysis showed that the number of homicides in the area near the closest mammography facility was associated with a higher risk of late-stage diagnosis, whereas distance to mammography had no effect. They hypothesized that fear of crime deterred women from receiving timely mammography screenings.

This brief discussion of accessibility highlights the interrelationships among the many dimensions of access to health care. Location and distance have significant effects on people's willingness and ability to use services, but these geographical effects vary in importance and meaning among places, populations, times, and individuals. In emphasizing the spatial aspects of accessibility, geographic information systems can easily hide or ignore the important social dimensions. This means that particular care needs to be taken in structuring GIS-based studies of access and interpreting the findings.

Mapping Service Locations

Preparing maps of health service locations is an important application area for GIS. Such maps can be used to display service location patterns, to provide information to residents about service locations and availability, and to visualize the spatial match between service needs and resources. Information about health services typically exists in tabular form, as lists of service providers and their addresses. Local, state, and federal agencies often maintain separate lists of their own services, as do private and voluntary organizations, leading to multiple lists that must be linked and collated for assessing service availability and accessibility. From this tabular information, finding out what types of services exist in an area, and where they are located, is a difficult task.

Many health agencies now use GIS to manage spatial information about services. Addresses of health services are geocoded and then displayed on a map. Storing service data in a GIS can be beneficial for both service providers and the people who need services. Providers can quickly view their service portfolio to visualize areas of over- and undersupply. They can also query the GIS to examine types of services offered, utilization levels, staffing, and financial performance. For service clients, GIS offers a tool for identifying nearby service providers and finding out their attributes—services offered, hours of operation, and so on. Easy-to-use web-based systems are being developed to facilitate these querying and mapping functions.

Often health planners want to know not just about one type of service, but about an array of services that support health and well-being—education, childcare, jobs, mental health, substance abuse, and social welfare. GIS is being used to map the uneven spatial distributions of health-related community resources and services and to examine the associations between service needs and resources (Pearce, Witten, Hiscock, & Blakely, 2007; Macintyre, Macdonald, & Ellaway, 2008).

To better integrate diverse health-related services, planners advocate the creation of ***service hubs*** (Wolch, 1996)—sites with health and social service agencies located in close proximity. The geographical concentration of interrelated agencies maximizes accessibility for service clients and promotes coordination among service providers. To analyze accessibility to service hubs and the full set of services that enhance health, one can overlay geocoded data for diverse services, displaying multiple layers of access to multiple service agencies. Morrison, Alexander, Fisk, and McGuire (1999) developed a GIS to allow welfare recipients to pinpoint the locations of essential health and social services, including job centers, childcare facilities, and primary health care centers. Bus routes were also displayed, along with the residential locations of employable welfare recipients.

Mapping Health Care Needs and Services

Service location information is particularly relevant when analyzed along with data on health care needs. Fundamentally, ***need*** describes the prevalence of health conditions that should be addressed by health care services. It can be measured and analyzed in many different ways. Typically, health planners use a combination of demographic, socioeconomic, and health outcome indicators, both quantitative and qualitative, in defining need. For example, in describing the need for prenatal care services, one would want to consider the number of pregnant women or number of women in the childbearing age groups. In addition, because women who have high-risk pregnancies require more intensive and frequent prenatal services, indicators of pregnancy outcome or risk, such as low birthweight or maternal age, are also relevant (McLafferty & Grady, 2004). When examining need for particular health care services, characteristics of the service may be important, as services may be targeted to particular population groups or restricted to individuals who meet certain eligibility criteria.

In North Carolina, Hanchette (1999) used GIS to identify communities in need of universal lead screening. CDC guidelines recommended universal screening in areas where either a large fraction of young children had elevated blood lead, or more than 27% of the housing stock was built before 1950. Using ZIP Code areas as a base, data layers were created depicting the age of housing and the prevalence of elevated blood lead in earlier screening tests. ZIP Codes that met the CDC guidelines were selected by querying the data layers and then targeted for universal screening.

Researchers are increasingly relying on statistical methods to create multidimensional indicators of health care need and socioeconomic deprivation. Analyzing heart disease in Kentucky, Barcus and Hare (2007) constructed a socioeconomic deprivation index based on census indicators including poverty, education, housing quality, and unemployment. The statistical method of factor analysis yielded a single important factor representing socioeconomic deprivation. GIS maps of factor scores show a concentration of disadvantage in southeastern Kentucky (Figure 9.2), an area with high heart disease incidence and mortality.

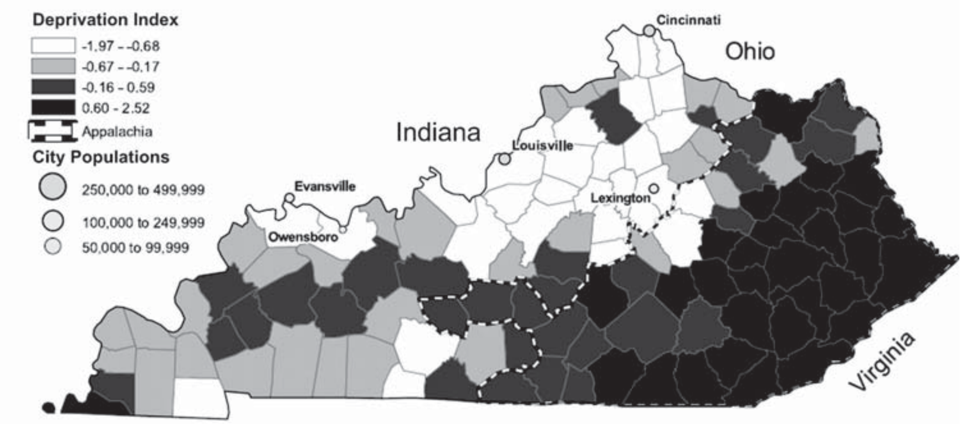


FIGURE 9.2. Geographic variation in socioeconomic deprivation, an important indicator of health care need, by county, in Kentucky. From *Southeastern Geographer*, Volume 47, no. 2. Copyright © 2007 by Southeastern Division, Association of American Geographers. Published by the University of North Carolina Press. Used by permission of the publisher.

Creating multidimensional indicators of need for services often involves linking variables measured over different geographies—for example, census tracts, point locations, and ZIP Codes. The linking process requires that each variable be estimated for a set of ***consistent* zones** or other common geographical units. Spatial interpolation methods, such as areal interpolation and dasymetric mapping discussed in Chapters 4 and 6, are invaluable in this estimation process.

Need also has qualitative dimensions, as described by "perceived" need. When available, data from health surveys can be incorporated to capture individuals' views of their health care needs and their perceptions of health-related resources (Macintyre, Ellaway, & Cummins, 2002). In examining access to care at a community health care center in Missouri, data from a needs assessment survey were geocoded and mapped to better understand spatial variation in health-related behaviors and perceived barriers to care (Phillips, Kinman, Schnitzer, Lindblom, & Ewigman, 2000).

Analyzing service needs in a GIS environment poses many challenges. For most health services the dimensions of need are not well defined; they may vary from person to person or by population group and may be challenging to measure. Combining and comparing indicators of need across individuals, groups, or areas are challenging tasks. Here the visualization and data layering capabilities of GIS can be exploited to view and analyze different dimensions of need among different population groups.

Assessing Potential Access to Health Services

Which communities and populations have poor geographical access to health services? Since the early 1900s, health planners in the United States have been concerned about this question, especially in rural areas. Efforts have been made to identify communities with poor access (shortage areas) and to implement policies for improving service availability. These efforts focus on ***potential accessibility***, the geographical match between people and essential health care services. At its core, the concept refers to the separation between services and population—how much distance, cost, time and effort are involved in reaching service facilities. It may also incorporate ***service capacity constraints,*** or restrictions on the number of people who can be served at each facility. There are many ways of characterizing potential access, and most can easily be implemented in GIS (Higgs, 2004).

Distance and Travel Time Measures

One of the simplest ways of measuring potential access is to calculate the distance from the population in need of service to the nearest service provider. Assume that we have a data set containing the point locations of all service providers—for example, the locations of all hospitals in the state of Wyoming. We also have a spatial data set that describes the population in need of service. This population data set can be either a point data set that contains residential locations of people in need of service or an area data set with counts of population by area. The latter must be converted to a point data set by finding the centroid of each populated area (zone), as described in Chapter 5, or by allocating the population to point locations within each area as discussed later in this chapter. To estimate distance, we first identify the service provider closest to each population point/centroid and then calculate the distance to that service provider. The average of those distances provides a widely used measure of geographical access to services. The larger the ***average distance***, the farther people must travel, on average, to their closest service facility, and the poorer the geographical accessibility. When population data are aggregated to geographic zones, the ***population-weighted average distance***, in which distances are weighted according to zonal populations, is an appropriate measure of access.

Differences in average distance to health facilities highlight inequalities in geographical access to care. In most countries, rural residents travel significantly farther for care than their urban counterparts (Probst, Laditka, Wang, & Johnson, 2007). To examine geographical access to pediatric medical services in the United States, Mayer (2006) used GIS to geocode provider locations by subspecialty. Distances from each ZIP Code centroid to the nearest provider were calculated and weighted by the under-18 population. Comparisons of population-weighted average distance across pediatric subspecialties showed substantial variations in access to care.

In addition to average distance, examining the ***frequency distribution of distances*** can shed light on geographical access to services. The frequency distribution of distances is a graph that shows the number of people living within a distance range (e.g., 5–10 kilometers) of their nearest service facility. Because people and health services typically cluster in urban areas, the frequency distribution of distances is often skewed, with a large proportion of the population close to services, and a significant minority quite distant from services. Figure 9.3 shows the frequency distribution of distances for the adult female population in Illinois to the closest mammography facility. Nearly 37% of adult women live within 2 kilometers of a mammography facility, a reflection of the high concentrations of people and facilities in the Chicago area. Seven percent of women live more than 20 kilometers from a facility and are thus disadvantaged by long travel distances.

These kinds of analyses are important for providing population-based evaluations of geographical access to health services at the national and regional scales. Onega et al. (2008) used GIS to estimate travel times to the nearest cancer center in the United States, and they compared travel time statistics among sociodemographic population groups. Excessive travel burdens for Native Americans and nonurban populations were highlighted. To compare spatial access to health care among population groups in Wales, Christie and Fone (2003) created travel-time frequency distributions for each group for three travel speed scenarios. Geographical access problems were most acute among rural residents and, under certain scenarios, among the elderly and residents of deprived areas.

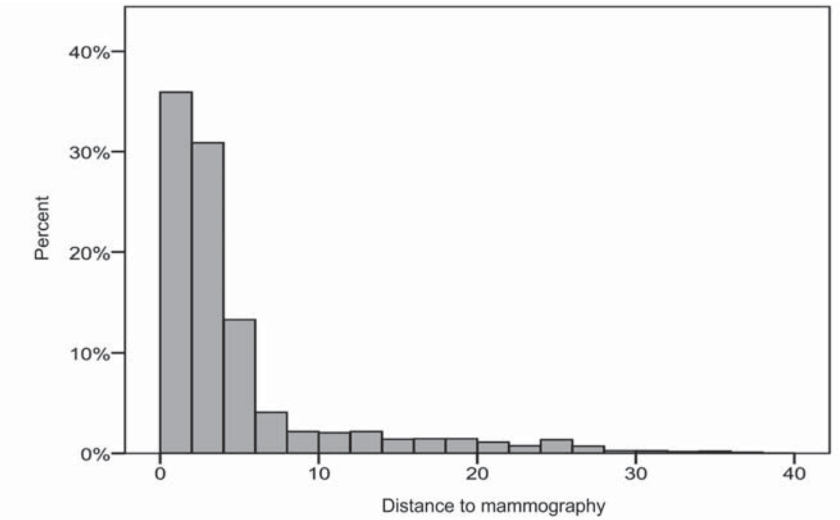


FIGURE 9.3. Frequency distribution of travel distances (in kilometers) to the nearest mammography facility for the adult female population in Illinois. The *y*-axis shows the percent of adult women who live within a particular distance range of a mammography facility.

Analyzing the distribution of distances can offer potent insights into the equality and inequality of geographical access among population groups.

Distance measurements can also be used in defining catchment areas (discussed later in this chapter) for health service providers. A maximum distance or travel time threshold delineates the area in which people have reasonable access to care. Schuurman, Fiedler, Gryzbowski, and Grund (2006) employed a 1-hour travel time threshold to identify catchment areas for hospitals in British Columbia.

MEASURING DISTANCE

Inherent in any assessment of geographical access is a measure of distance that represents the geographical separation, in distance, time, or cost, between people and services. There are many ways of measuring distance. For small-scale investigations—those at a national or regional scale—or when using unprojected coordinates such as lat/lon, distance is calculated along the curved surface of the earth. This is referred to as ***spherical distance***, and it measures the distance along a great circle connecting two points. The spherical distance in kilometers between points i and j can be estimated as:

$$d_{ij} = \arccos(z) \times 6371.11$$
$$z = \sin(Y_i)\sin(Y_j) + \cos(Y_i)\cos(Y_j)\cos(X_i - X_j)$$

where Y is the latitude (in radians) and X is the longitude (in radians).

In analyzing geographical access to services in relatively small areas, such as cities, metropolitan areas, and small states, the earth's curvature does not present a major problem.

A more common measure of spatial separation is the ***Euclidean (straight-line) distance***. If the projected coordinate locations of points A and B are (X_A, Y_A) and (X_B, Y_B) respectively, the Euclidean distance between A and B is

$$\sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$$

Euclidean distances are appropriate if one is working with projected geographical coordinates, as in the state plane or UTM coordinate systems; however, it is important to keep in mind that Euclidean distances do not take into account the curvature of the earth's surface. Most GIS use the Euclidean distance metric as the default in all distance calculations (i.e., in computing buffers and interpoint distances). In many situations, however, the Euclidean metric poorly represents travel patterns and travel potential.

One weakness of Euclidean distance is that it fails to take into account transportation routes and barriers to movement. Only rarely can people move from place to place along straight lines. In areas where the road network follows a grid pattern, one can approximate network distances by using the ***Manhattan metric*** to calculate distance. The Manhattan metric measures distance along the axes of a coordinate grid

$$|X_A - X_B| + |Y_A - Y_B|$$

Since the distance measurements vary depending on the orientation of the grid, it is important that the grid be oriented along the main axes of the road network (Figure 9.4). Although Manhattan distances are not as accurate as distances measured along a transportation network, they can be computed very efficiently and work well as a surrogate for network distances in places where streets follow a grid pattern.

Most GIS include tools for calculating ***network distances*** that follow a street, bus, or rail network. Given a starting node and an ending node, the GIS will compute the length of the shortest path (see Chapter 10) along the transport network, and the result can be used as the network distance. Such distances offer a better approximation of the actual distances people must travel to obtain health services.

Convenient tools for estimating network distances and driving times (discussed below) are also available in Internet-based systems such as MapQuest®, Google Maps®, and Bing®. In these systems, a pair of origin and destination locations is input, and network distance and estimated driving time between the two locations are output. For public health analysts, the main disadvantage of using these systems for distance/time estimation is that the systems are set up to handle one origin-destination query at a time. For multiple origins and

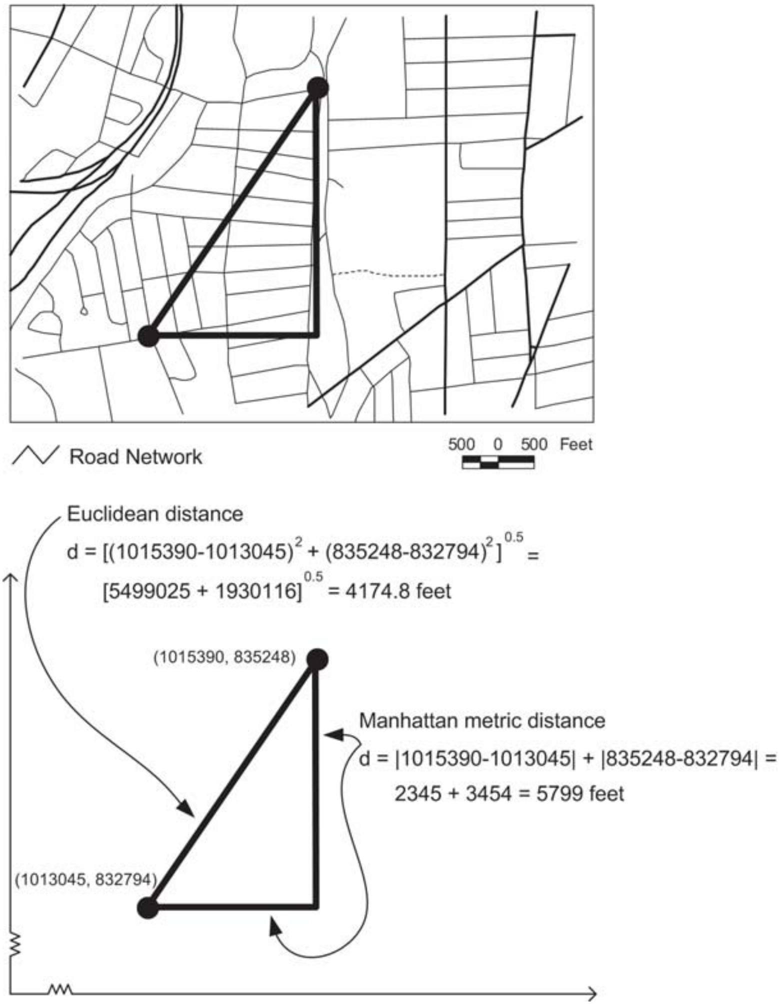


FIGURE 9.4. The calculation of Manhattan metric distance between an origin and a destination.

destinations, it is necessary to write customized computer code to automate the querying process.

MEASURING TRAVEL TIME

Although distance is a fundamental indicator of geographical access, travel time, cost, transportation access, and perceived distance are often much more relevant to health care utilization. Using GIS, one can estimate ***travel time*** along road networks, taking into account average speeds and speed limits on different classes of roads and physical barriers to travel (Schuurman et al., 2006; Apparicio, Abdelmajid, Riva, & Shearmur, 2007). To determine the travel time between two points, we identify the route connecting those points and sum the estimated travel times along each road segment in the route (Figure 9.5). Travel time provides a better indication of geographical barriers to health services than does travel distance, since by definition travel times incorporate access to transportation.

Mode of transportation is also important in estimating travel times. Vulnerable populations—people with low incomes, the aged, and others—often must rely on public transportation, walking, and taxis to access health services. Lovett, Haynes, Sunnenberg, and Gale (2002) analyzed access to general practitioner services in East Anglia, England, an area where travel by both bus and car is common. For those traveling by car, travel times were estimated along the road network from each postcode area to the nearest service provider. Car travel

ID	Name	Type	Length ft	Travel Speed mph	Travel Time min
544	Country Club	Road	424.544	45	0.11
565	Country Club	Road	3094.514	45	0.78
580	Country Club	Road	830.618	45	0.21
584	Country Club	Road	519.34	45	0.13
557	Winding	Lane	443.895	25	0.20
611	Winding	Lane	1154.725	25	0.52
638	Stony	Way	519.055	25	0.21
655	Stony Corners	Cir	338.462	25	0.15
666	Stony Corners	Cir	399.896	25	0.18
695	Cotswold	Way	697.009	25	0.32
			8422.028		2.84

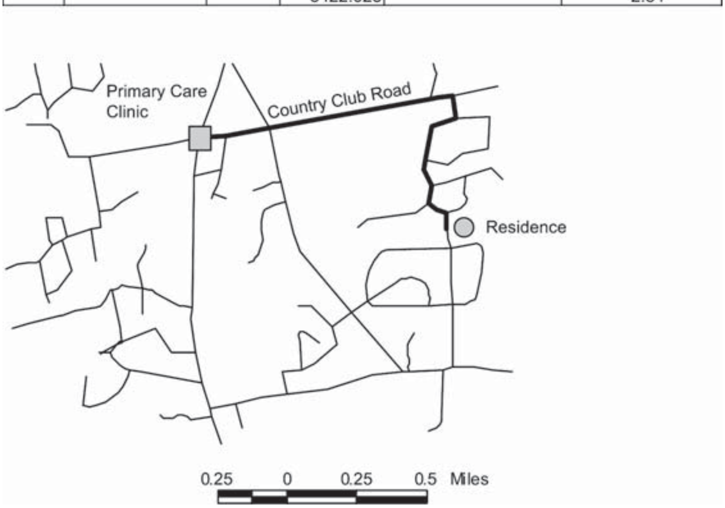


FIGURE 9.5. The measurement of travel time between an origin and a destination can be implemented in a GIS provided that data are available on the amount of time it takes to traverse a segment along the route of travel. For an automobile user traveling the speed limit, the 1.6-mile trip from the primary care clinic to the residence using the highlighted route would take about 3 minutes.

times were mapped as a continuous field by generating a triangulated irregular network (TIN) of the nodal travel time values. To evaluate accessibility by bus, researchers focused on the frequency of bus service and whether or not residents could walk to a bus route that went to a particular provider. Spatial buffer functions in GIS were used to identify areas within walking distance of bus routes (Figure 9.6). Analyzing combinations of car and bus access, Lovett and others found pockets of rural deprivation characterized by high health care need and low transportation mobility, and these areas were targeted for bus service improvements.

Walking and use of public transportation are especially important in developing countries where access to private transportation is limited. Speed of walking varies with slope, topography, quality of road or track and other environmental and social factors. In analyzing accessibility to primary health care clinics in rural South Africa, Tanser, Gijbertsen, and Herbst (2006) used GIS to estimate walking times from rural homesteads to the nearest clinic. Different speeds were assigned to different paths based on the quality and distribution of the road/path network and the presence of natural barriers.

Density Measures

Rather than focusing only on distance or travel time to the nearest service provider, one can compute ***density measures*** that describe the full range of providers in an area. Density refers to the number of providers available in relation to local population or geographic area. One can define density in relation to predefined geographic zones such as states or counties—the ***container approach***—or within a fixed buffer distance of a point of origin—the ***coverage approach*** (Higgs, 2004). A well-known container measure is the physician to population ratio, which is widely used in analyzing accessibility to care.

Although container approaches have traditionally been used for computing density measures of access to health care, such approaches have major limitations. The geographic zones that underpin density calculations are often arbitrarily defined, political units that differ in size, shape, and socioenvironmental characteristics. All people living within a zone are assumed to have equal access to care; this is a weak assumption, especially when zones are large (Hewko, Smoyer-Tomic, & Hodgson, 2002). Moreover, container measures ignore the availability of services across area boundaries. To address these limitations, it is better to use coverage approaches that estimate density within overlapping, analyst-defined coverage zones. Coverage measures are easily computed in GIS.

Meersman, Breen, Pickle, Meissner, and Simon (2009) used a GIS-based coverage measure in investigating variations in mammography use among women in Los Angeles County. The project involved geocoding residential locations for women surveyed in the California Health Interview Survey and geocoding locations of mammography facilities. For each woman, researchers computed the number of mammography facilities within spatial buffers of varying sizes. The density of facilities within a 2-mile buffer was positively associated with mam-

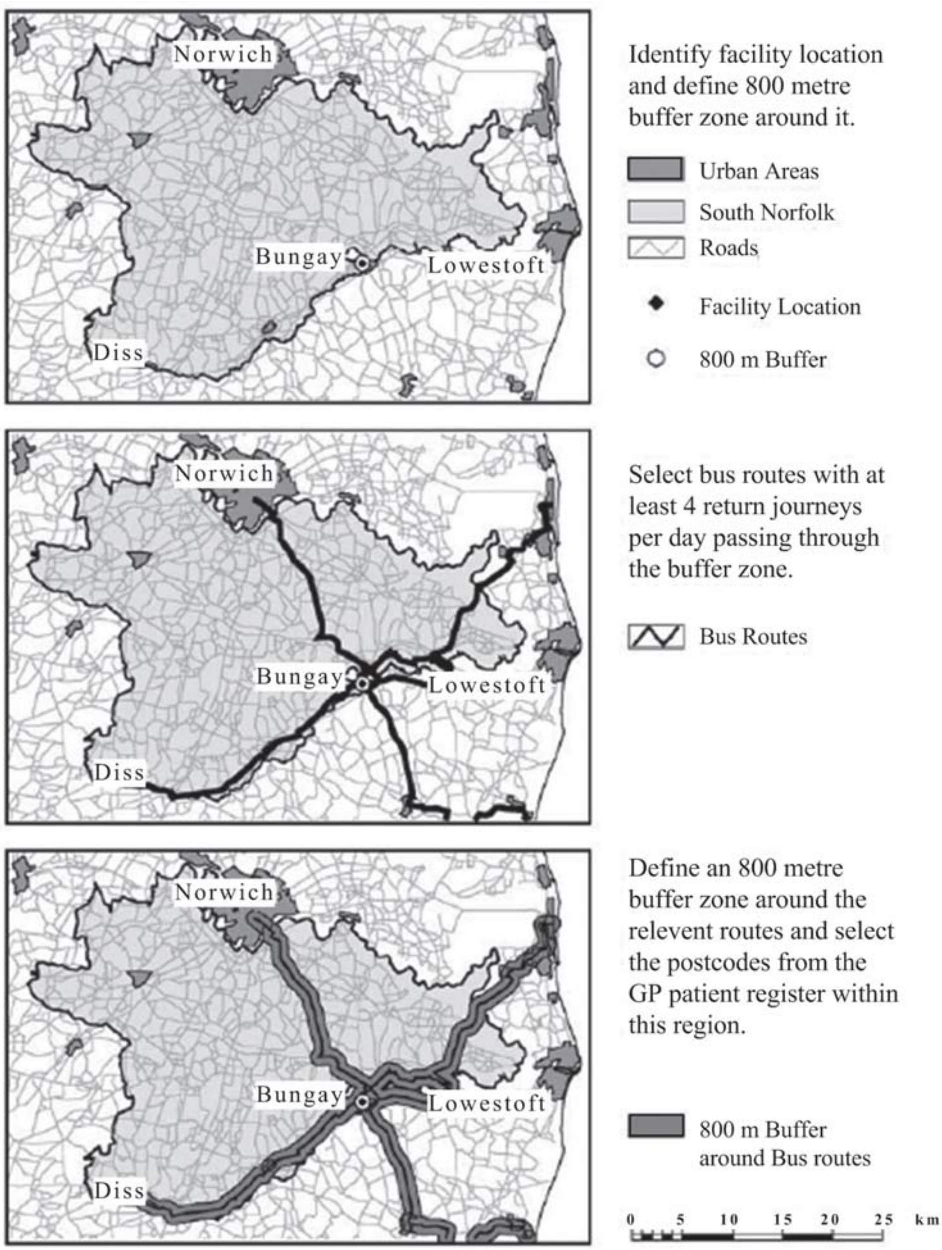


FIGURE 9.6. GIS procedures for evaluating accessibility to general practitioners' offices on the basis of the use of car and bus transportation. This figure shows the procedures used in modeling travel by bus. Reprinted from *Social Science and Medicine*, 55(1), Lovett, A., Haynes, R., Sunnenberg, G., & Gale, S., Car travel time and accessibility by bus to general practitioner services: A study using patient registers and GIS, 97–111, copyright 2002, with permission from Elsevier.

mography use, suggesting that geographical access is a barrier to mammography screening.

KERNEL ESTIMATES OF SERVICE DENSITY

When analyzing geographical access to health care in densely populated urban areas, kernel estimation, discussed in Chapter 5, can be employed in estimating service density (McLafferty & Grady, 2004). We begin by geocoding service providers to point locations. A circular window scans the map, and the kernel density of service providers is computed within each circular window. Kernel estimation depicts the density of service providers (number per unit area) as a continuous spatial variable, with peaks representing areas of high geographical access to providers and valleys indicating areas of poor access. Guagliardo and others used kernel density estimation to investigate children's geographical access to primary care physicians in Washington, D.C. (Guagliardo, Ronzio, Cheung, Chacko, & Joseph, 2004). The map of physician density revealed huge social and spatial inequalities in access to primary care physicians (Figure 9.7).

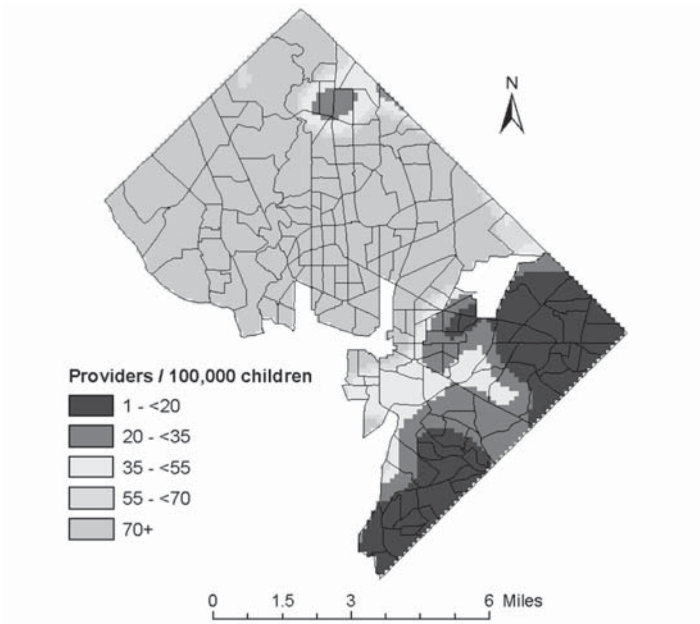


FIGURE 9.7. Ratio of density of pediatric service providers to density of children in Washington, D.C. Densities were calculated using kernel estimation. Note the absence of providers in southeast Washington, a low-income area. From Guagliardo (2004). Originally published by BioMed Central in the *International Journal of Health Geographics.* Open Access.

Physician density was highest in the affluent neighborhoods of northwest Washington and extremely low in the low-income neighborhoods of southeast Washington.

Kernel estimation has several advantages. The kernel weighting function incorporates a decline in access with increasing distance consistent with distance decay. Density estimates can be assigned to individuals and aggregated to show differences in access among population groups, or they can be used in individual-level statistical models (McLafferty & Grady, 2004). By computing a kernel density surface to represent population, and dividing provider density by population density, one can investigate the availability of services in relation to population—a provider to population ratio (Bailey & Gatrell, 1995). Despite these advantages, kernel methods have important limitations. The mathematical form of the kernel weighting function is somewhat arbitrary. Kernel methods perform poorly in rural areas where few providers exist, and one can debate whether having more providers (a higher density) nearby is necessarily better.

TWO-STEP FLOATING CATCHMENT AREA METHOD

To assess the local availability of services in relation to population need, Luo and Wang (2003) developed the ***two-step floating catchment area method (2SFCA)***. The method requires two point data sets: data on the locations and capacities of service providers and data on the locations of population in need of services. The 2SFCA moves between these two data sets in a two-step process. First, we construct a threshold distance or travel time window around each service provider j and compute the provider to population ratio R_j within that window (Figure 9.8). Second, for each population location i, we search all provider locations within a threshold travel distance/time and sum up the R_j values for those providers. The resulting value is the accessibility score for the population at place i. Higher values indicate better spatial access to health care providers. The 2SFCA method can easily be implemented in GIS using join and sum functions (Wang, 2006). Although the original 2SFCA method assumed that access is constant within the threshold distance/time window, a recent enhanced version incorporates distance decay (Luo & Qi, 2009). Other recent enhancements include incorporating differences in health needs and permitting the threshold distance/time to vary among population areas (McGrail & Humphreys, 2009).

McGrail and Humphreys (2009) used an enhanced 2SFCA method to evaluate the spatial accessibility of primary care physicians in Victoria, Australia. All general practitioners in Victoria and within 1 hour of the provincial border were geocoded to their practice locations. Resident population was geocoded to collection districts (CDs), small, census-defined zones averaging 500 in population size. GIS was used to compute travel times between CD centroids and physician locations, and the modified 2SFCA method was implemented. The map of spatial access to primary care (Figure 9.9) reveals large regions of poor spatial access in rural areas, particularly in areas with high-need populations.

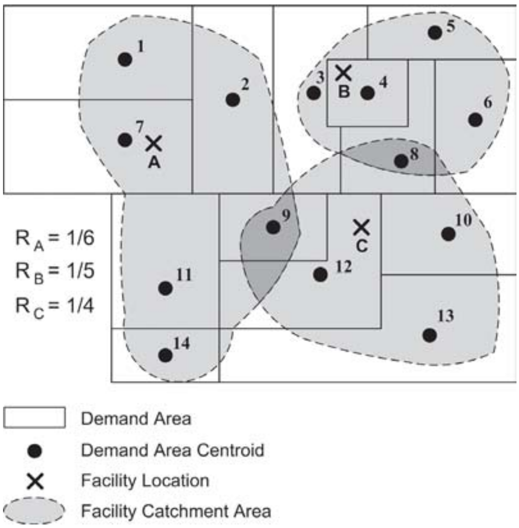


FIGURE 9.8. Schematic diagram of the first step of the two-step floating catchment area method. Locations A, B, and C represent health care facilities; numbers 1–14 identify demand area centroids. A catchment area around each facility is defined on the basis of a maximum travel time. The supply/population ratios (R) for each of the three facility catchment areas are shown. Facility A's catchment area contains one facility and six population demand points. In the second step of the two-step FCA, travel time windows are constructed around each demand point. R values within each catchment are summed, providing an index of spatial access for the population residing at that demand point.

Potential Accessibility Measures

Accessibility measures based on travel time, distance, or density offer only a partial view of access to services. In reality, people trade off geographical and nongeographical factors in making decisions about health service use. The widely used ***gravity model*** and ***potential model*** offer a method for modeling these tradeoffs in defining service access. The gravity model is based on an analogy with Newtonian physics in which the interaction between places is directly related to their relative sizes or attractiveness, and inversely proportional to the distance between them. People are willing to travel farther to obtain better (more "attractive") health care services. Attractiveness depends on price, quality of services, accommodation, cultural appropriateness, and a host of service-related factors. Different population groups typically evaluate service attractiveness differently, depending on the service qualities that are most relevant to their own needs. Gravity models belong to a more general class of ***spatial interaction models***, tools for modeling interactions between places. We discuss the models later in this chapter as methods for predicting flows of people to health services.

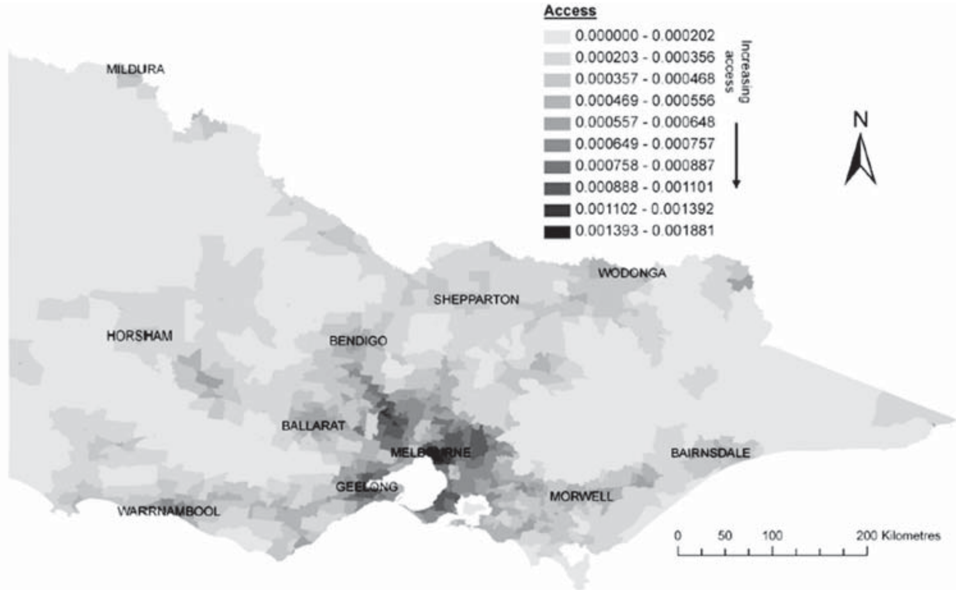


FIGURE 9.9. Spatial access to general practitioners in Victoria, Australia, was calculated using an improved two-step floating catchment area method. The map of spatial access shows strong urban-rural inequalities in access and disparities based on health care needs. From McGrail and Humphreys (2009). Originally published by BioMed Central in *BMC Health Services Research*. Open Access.

The ***potential model*** uses gravity concepts to describe patterns of accessibility to services. Potential access is calculated for a particular individual or area, i, and it measures the area's overall accessibility to services. Defining A_j as the attractiveness of health service j, and d_{ij} as the distance (or travel time or cost) from i to j, we can compute the potential accessibility of individual or neighborhood i to health services as

$$\sum_j A_j / d_{ij}{}^{\beta}$$

Higher values reflect higher levels of potential accessibility, which occurs when people live close to high-quality service facilities.

The distance exponent, β, describes the frictional effect of distance on service accessibility. When $\beta = 0$, distance has no impact on service utilization or access, and access depends only on the attractiveness of service providers. Conversely, when β is large, distance has a strong frictional effect, and access depends only on the distance to service facilities. Large values clearly give more weight to nearby services in computing potential accessibility.

Estimating potential accessibility is greatly simplified in a GIS environment. Distances between people or communities and service facilities are easily computed, and service attractiveness data can be tied to geographic locations. For each person or community *i*, one can calculate the potential accessibility to health services using GIS spreadsheet operations (Wang, 2006).

Figure 9.10 shows a map of potential accessibility to hospitals. Accessibility values were calculated for census tracts, with several simplifying assumptions. First, straight-line Euclidean distances from the centroid of the census tract to each hospital were used to represent geographical separation, and second, hospital size (number of beds) was employed as a surrogate for attractiveness. The differential shading of the census tracts reflects their varying levels of potential accessibility. Note the high levels of accessibility for census tracts located near the cluster of large hospitals in the center of the region.

In the classic gravity model, accessibility is inversely related to distance squared ($\beta = 2$)—a direct analogy with Newtonian physics. However, there is no reason to assume that the Newtonian distance exponent necessarily applies in modeling access to health care. Research shows that the appropriate distance exponent varies among populations and research contexts (Kwan, 1998; Lin, Allen, & Penning, 2002). Furthermore, the power function ($d^{-\beta}$) may be inappropriate (Thill & Kim, 2005). A more general expression for potential accessibility is

$$\sum_j A_j f(T_{ij})$$

where (T_{ij}) is distance or travel time and $f(T_{ij})$ is an ***impedance function*** that represents the decrease in access with increasing distance or travel time. Impedance functions can have different functional forms including power, exponential, Gaussian, and log-logistic (s-shaped) (Thill & Kim, 2005; DeVries, Nijkamp, & Rietveld, 2009). Kwan (1998) argues that a Gaussian impedance function fits real-world travel patterns. The Gaussian function depicts a gradual decline in access with distance close to the facility, and a steeper decline far from the facility (Figure 9.11). Note that the Gaussian function is also frequently used in kernel estimation. As shown in Figure 9.11, for a given problem, the appropriate distance impedance parameter β will depend on the impedance function used.

Researchers have used statistical methods to find an appropriate impedance function (Thill & Kim, 2005; Wang, 2007; Scott & Horner, 2008). Using empirical data on patient travel to health services, one can determine the exponent that best fits actual travel patterns. Analyzing Chinese immigrants' spatial access to physicians in Toronto, Wang (2007) used data from a survey of Chinese immigrants to determine an appropriate impedance function. Differences in transportation and mobility can also be incorporated. The research on physician accessibility in East Anglia, England, discussed earlier in this chapter (see Figure 9.6) calculated potential accessibility based on network travel times.

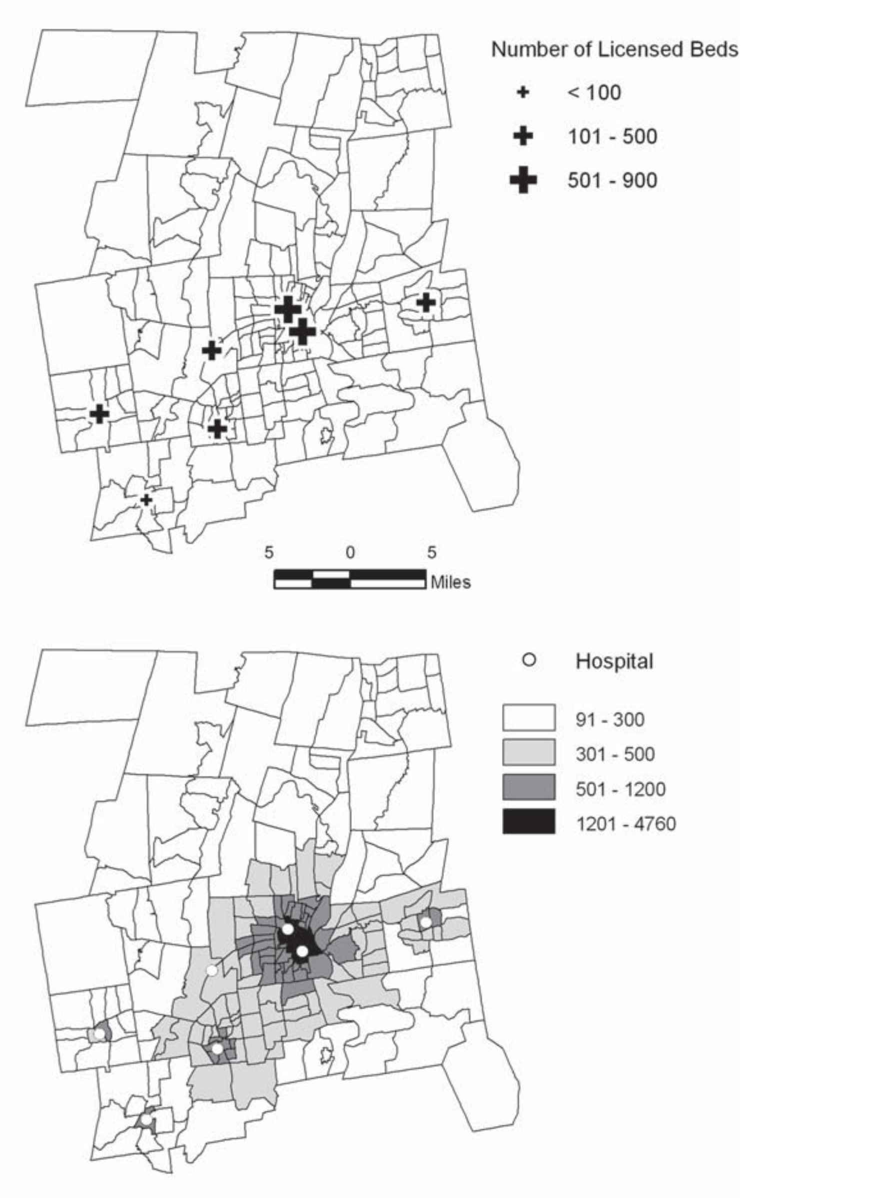


FIGURE 9.10. The top map shows the locations of community hospitals in a region. The bottom map shows potential accessibility to hospitals based on the number of licensed beds as a measure of hospital attractiveness and distance from hospital to census tract centroid. Census tracts in the center of the study region close to large hospitals have the highest potential accessibility. Census tracts close to small hospitals and census tracts located far from hospitals have lower potential accessibility.

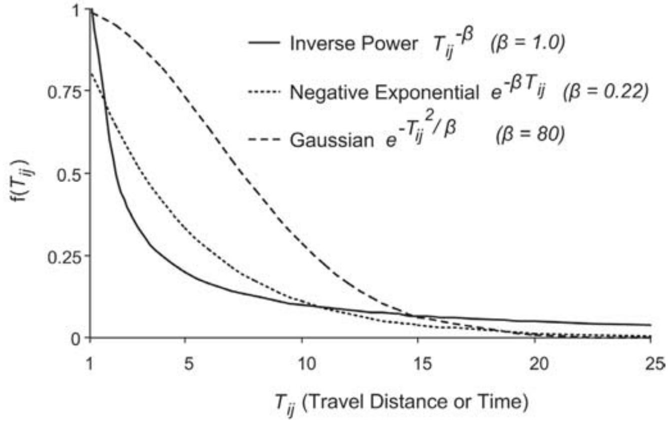


FIGURE 9.11. Different spatial impedance functions. T_{ij} is a measure of travel time or distance.

Another approach is to use an impedance function calibrated for one study area to predict potential accessibility in another study area. Knox (1978) did this in estimating intraurban patterns of potential accessibility to general practitioner services. The precise form of the distance function ($e^{-1.52d_{ij}}$) came from an earlier study of general practitioner use. Applying impedance functions from one study area to another has certain limitations, however. The effects of distance can vary over time and space, leading to errors in estimating potential accessibility. Similarly, the frictional effect of distance can differ substantially among population groups, reflecting differences in income, access to transportation and sociocultural factors.

Another problem is that the distance exponent depends in part on the spatial configuration of service opportunities (Haynes & Fotheringham, 1984). Research indicates that the distance exponent tends to be closer to zero for centrally located zones that are accessible to a large number of service facilities than for peripheral zones located distant from service opportunities. If this is the case, the distance exponent will not be transferable from one study area to another unless the two areas contain similar geographical arrangements of service opportunities and population groups—a highly unlikely situation. To address this problem, one can calculate potential accessibility over a range of exponent values and explore the stability of the observed accessibility patterns.

Defining and measuring the attractiveness term is also an important issue in applying potential models. Attractiveness is a multidimensional concept that describes the range and number of services offered, appropriateness, price, and quality of treatment. To define attractiveness, researchers have used service capacity—number of physicians, number of hospital beds—as a surrogate measure (Morrill & Earickson, 1968), but clearly this approach is limited. A better approach is to use a set of variables describing attributes of the health care provider and the range and quality of care delivered (Hyndman & Holman, 2001). Cultural characteristics of providers such as languages spoken can also be incorporated (Wang, 2007).

Visualizing Accessibility

Regardless of how it is measured, potential accessibility to health services is distributed unevenly over space, reflecting the way most health services are provided—at fixed sites, serving a dispersed population. Some individuals will always live closer to the service sites than others. GIS provides a tool for viewing and summarizing geographical inequalities in accessibility and seeing if differences in accessibility stem from obvious gaps in service coverage or are structured along class, ethnic, or racial lines (Talen, 1998; Pearce, Witten, Hiscock, & Blakely, 2007; Macintyre, Macdonald, & Ellaway, 2008). Talen (1998) describes a GIS for "visualizing fairness" in service distribution patterns. The system incorporates a variety of accessibility measures including average travel distance and population coverage. GIS maps of accessibility are produced that can be viewed individually and related to maps that show the distributions of populations groups, housing values, and environmental features. Maps and graphs reveal the differential patterning of accessibility. Figure 9.12 presents an example of an ***equity map*** that describes patterns of accessibility to health care services for residents of public housing in Montreal, Quebec (Apparicio & Seguin, 2006).

Representing Population in Potential Accessibility Modeling

An important issue in accessibility modeling is the underlying geographic representation of population in need of service. Ideally one would know the residential location of each person or household in the population, but such point data are rarely available. Instead population data are aggregated to predefined zones such as blocks or census tracts. Doing so leads to ***spatial aggregation error*** when distance or travel time measures are calculated. Research indicates that spatial aggregation error has significant effects on measurement of spatial accessibility, especially when using large geographic zones (Hewko, Smoyer-Tomic, & Hodgson, 2002; Langford & Higgs, 2006). This error is also important in determining optimal locations for new health care services, as discussed in Chapter 10. There are several strategies to reduce the impacts of spatial aggregation error, and GIS is critical to their implementation. A naïve approach is to redistribute population evenly within each zone; a much better method is dasymetric mapping, which redistributes the population unevenly based on ancillary data such as road or building locations (Langford & Higgs, 2006). Another recent method involves constructing travel time zones around service facilities and distributing population among those zones based on ancillary data (Shi & Berke, 2009).

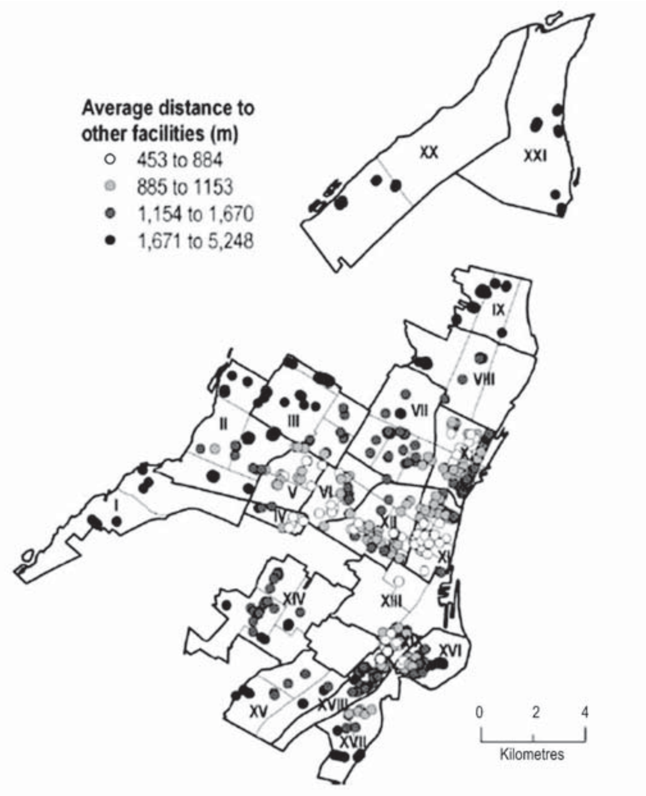


FIGURE 9.12. An equity map showing average network distances to health services for residents of public housing in different rental areas and rental districts in Montreal, Quebec. Each dot represents a public housing building. The dark-shaded dots indicate poor geographical access to health care. From Apparicio and Sequin (2006). Copyright 2006 by Urban Studies Journal Limited. Reprinted by permission of Sage.

Accessibility and Activity Spaces

Accessibility can also be assessed in relation to activity spaces, or the spaces that enclose daily travel patterns. As noted earlier, an individual's activity space consists of the set of locations that are visited regularly in everyday life—workplaces, schools, and shopping centers. It is the space that an individual travels within on a daily or weekly basis (Cromley & Shannon, 1986). Health care services located in or near an individual's activity space are more accessible and conveniently reached than those located far away. The accessibility measures discussed up to this point in this chapter all measure access with respect to residential (home) locations; yet people may travel to health care services from the workplace and other daily activity sites. Research shows that activity spaces extend well beyond the home location, even among disadvantaged populations. For example, in a study of low-income families in three U.S. cities, Matthews, Detwiler, and Burton (2005) found that 90% of daily activities occurred outside the census tract of residence. Varying in size and location among people and places, activity spaces are important in understanding spatial access to health care.

Using travel diary data to describe activity spaces, Sherman, Spencer, Preisser, Gesler, and Arcury (2005) devised a suite of GIS-based indicators of accessibility to health care providers. Respondents in a rural region of North Carolina were asked about routine activity and health care provider locations. Household locations were geocoded via GPS and merged with activity locations and transportation data. For each individual respondent, the authors measured an activity space using three alternative methods: a standard deviational ellipse, a road network buffer, and a standard travel time polygon. Figure 9.13 shows an example of a road network buffer activity space, created by buffering the roads that connect the person's home and activity locations.

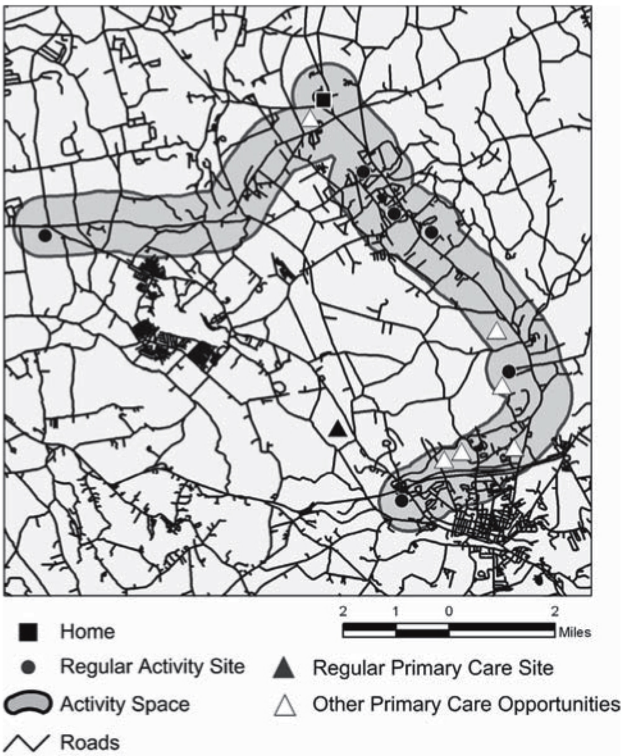


FIGURE 9.13. GIS representation of an individual activity space based on a road network buffer. The activity space includes the home location and other regular activity sites. For this person, the regular primary care site lies outside the regular activity space, although there are a number of other primary care opportunities located within the person's activity space.

The Global Positioning System is increasingly being used to record people's everday movement patterns in time and space (Troped, Wilson, Matthews, Cromley, & Melly, 2010). GPS identifies individual locations at regular time intervals, recording a dense set of space–time location points. These points can be used in creating three-dimensional ***space–time aquaria*** to visualize people's activity locations throughout the course of a typical day or week (Kwan, 2000). Simpler spatial measures of activity space can also be constructed from GPS-recorded activity points (Rainham, McDowell, Krewski, & Sawada, 2010). Although innovative and efficient for representing activity patterns, GPS data also present a series of methodological challenges, including difficulties in analyzing the very large number of track points and signal gaps caused by human or device error. Also, methods for aggregating individual activity spaces to generate population-level estimates are not well developed.

After delineating activity spaces based on travel diary or GPS data, one can compute indicators of accessibility to health care such as the number of providers within the activity space, the travel time from the activity space to the nearest provider, the overall size of the space, or the amount of time available during the day for accessing care. The advantage of these approaches is that accessibility is defined in relation to the routine patterns of everyday life, instead of focusing solely on the home location.

Analyzing Service Utilization

GIS are also valuable tools for analyzing ***revealed accessibility*** to health care services—that is, patterns of health service utilization. Such patterns are the result of individual choices about when and where to use services, the geographical configuration of health care opportunities, and the mediating effects of medical referrals and regulations. This section examines GIS-based methods for investigating several key questions relating to utilization: What is the service area for a health care facility? How will changes in health care delivery, for instance, the closing of hospitals, affect market areas and utilization? Are services and procedures over- or underutilized in particular areas?

Identifying Service Areas

The ***service area*** or ***catchment area*** for a health care provider is the geographic area that contains the bulk of population served. For a health care provider, the service area ties the client population to a geographic area, a neighborhood or community or set of communities. Some health facilities have ***mandated service areas*** in that they are required to serve the population living within a particular region, say, a county or set of ZIP Codes. Public schools often have mandated catchment areas, as all children who live within a given area are required to attend a particular school. Such mandated areas are less common in the case of health services in the United States, although some publicly provided services have mandated catchment areas. Furthermore, many managed care plans restrict health care choices to a given set of providers, resulting in mandated service areas for those who belong to a managed care plan.

When choice of providers is not mandated, health care services have ***natural service areas*** that arise through individual decisions and medical referral patterns. The service areas for different health care providers typically overlap, reflecting the diversity of health care needs and choices among people living in the same area. Patient origin information is essential for identifying “natural” service areas. The GIS analyst geocodes the addresses of patients who use the health care facility, and plots those address locations on a map (Parker & Campbell, 1998). The resulting geographical distribution of addresses defines the natural service area for the health care facility (Figure 9.14). The map of addresses may reveal outlier patients who reside very far from the service facility. To focus on the primary service area, we can plot the 80 or 90% of clients who live closest to the facility and identify the service area this way (Figure 9.14).

If client data are not available by residential address, but only by area, one can construct service area maps in several different ways. One is to rank the areas based on their respective shares of the health facility’s clients and define the service area as those areas that make up a prespecified percentage of the facility’s clients (Figure 9.15). Alternatively, one can use a plurality rule that defines the service area as the set of places in which a plurality of patients utilize the particular health care facility (Wennberg, 1999).

An understanding of service areas is important for health care providers because it ties the client population to a particular area or set of communities. This area can be examined in its own right to see if all populations are being adequately served and to assess the diversity of population health needs. Analyzing the social and demographic characteristics of service areas may reveal populations with unmet needs. Providers who want to expand their client base can use the service area map to identify places and populations that are not being well served and to chart out areas for future expansion.

Spatial Interaction Models of Health Care Utilization

Although maps of service areas are useful descriptive tools, they do not address the determinants of service utilization patterns and thus have limited value for forecasting and planning. What are the effects of distance, facility size, and service level on utilization? Spatial interaction models provide an essential tool for examining this question. The models describe and explain the movements or interactions between places as a function of distance and other factors. As noted earlier, spatial interaction models were first developed based on an analogy to Newtonian physics. However, they have been extended and enhanced greatly over the past decades, and there is now a suite of methods that can be applied to a variety of health planning problems (Lowe & Sen, 1996).

A particular form of spatial interaction model—the ***origin-constrained model***—has been widely used in the United States for health care planning

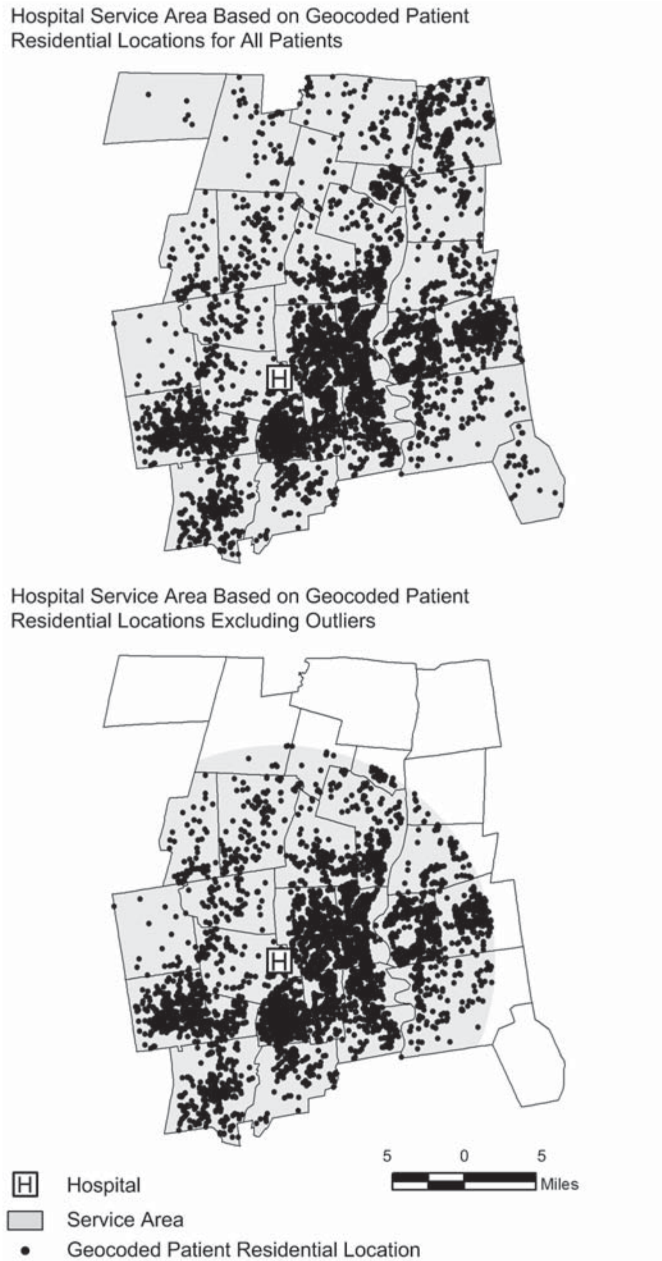


FIGURE 9.14. The primary service area of a health care facility identified by mapping patient residential locations in the region. When the locations of all patients are considered, the entire study region will be included in the service area. When we define the service area in terms of the closest 90% of patients, the primary service area is smaller. Boundaries of towns within the study area are shown for reference.

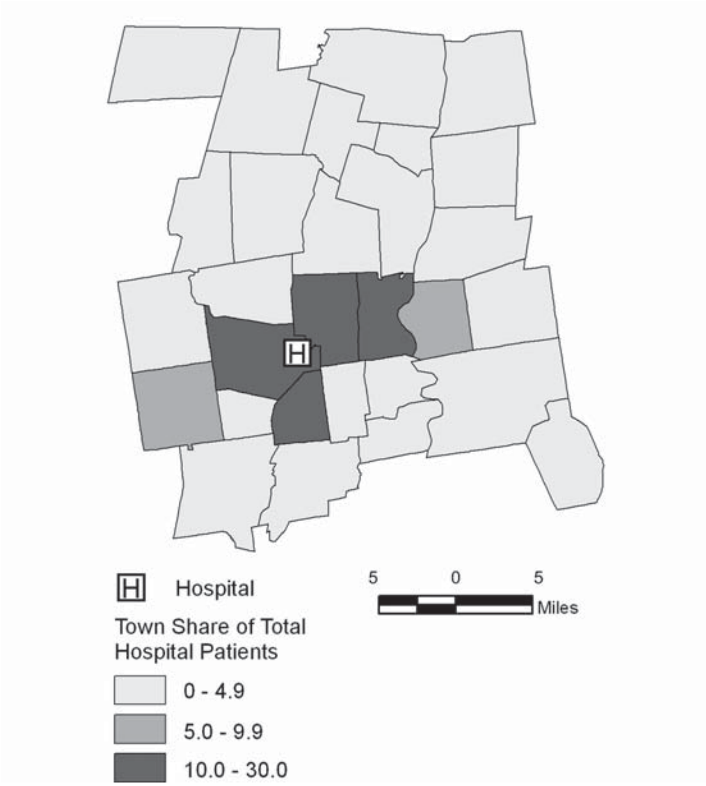


FIGURE 9.15. The primary service area of a health care facility identified by mapping the areas that account for the largest shares of hospital patients. In this case, 23 towns in the region each accounted for less than 5% of the total number of patients served by the hospital. About 60% of the patients served by the hospital resided in one of four towns.

(Martin & Williams, 1992). This model assumes that the number of trips from an origin area—for example, a town, ZIP Code, or census tract—is known and fixed. Interaction with health care facilities results from decisions in which people compare available facilities and select the one that is best in terms of distance, quality, and other characteristics. The gravity model models the allocation of those trips among health care facilities. For residents of area *i*, one can express the "utility" or value (U_{ij}) of health care facility *j* as a function of the distance or travel time to that facility (d_{ij}) and other attributes (k) of the facility that represent its attractiveness A_{kj}

$$U_{ij} = \left(\prod_k A_{kj}^{b_k}\right) / d_{ij}^{b_{k+1}}$$

The likelihood that an individual *i* will utilize facility *j*, I_{ij}, depends on the utility of that facility compared to the total utility of all facilities that could have been chosen, expressed as

$$I_{ij} = U_{ij} / \sum_{m} U_{im}$$

The parameters b_k measure the relative effects of service attributes A_{kj} and distance on utilization decisions. As with the potential model discussed earlier, the larger the parameter value, the more weight given to that particular factor in hospital choice. Parameter values can be estimated empirically via multivariate statistical methods (Congdon, 2001). Given data on flows of patients from origin areas to health care facilities, we find the parameter values that best describe these actual patient flows. Different functional forms for distance impedance can also be incorporated in estimating parameter values (DeVries, Nijkamp, & Reitveld, 2009).

Tai, Porell, and Adams (2004) analyzed hospital choices of rural Medicare beneficiaries using an origin-constrained spatial interaction model. The model included variables representing patient characteristics such as age and gender and hospital characteristics such as number of beds. Results confirmed that distance was a key determinant of hospital choice. Over half the rural Medicare patients in the sample relied on the nearest hospital facility. Patients with higher levels of income and education were more likely to bypass the closest rural hospital than were more disadvantaged patients. Results also confirmed the attraction of larger hospitals offering more complex services.

There are several other types of spatial interaction models. The ***destination-constrained model*** assumes that the total capacity of each facility (destination) is fixed, so each facility can only serve a predefined number of clients. Given this constraint, the model describes the flows of patients to facilities. These types of models have been widely used in Great Britain where health care is centrally planned and financed, and planning authorities often dictate the capacities of health care facilities (Mayhew, Gibbard, & Hall, 1986).

The value of gravity models for public health analysis lies in their ability to explain and predict health service utilization patterns. Researchers have used gravity models to estimate the changes in service areas that might occur when health facilities close, new facilities open, or other policy changes take place. Lowe and Sen (1996) examined a range of spatial interaction model applications in health care, including the impacts of hospital closure and universal health insurance on utilization and access. A key factor in their models was insurance match or the willingness of a hospital to accept the patient's insurance. Health care choices were strongly conditioned by the availability and type of insurance. In exploring the impacts of universal health insurance, the authors assumed that insurance barriers would disappear and that hospital utilization would be based solely on distance and attractiveness. The gravity model predictions revealed a marked increase in geographical access to hospitals for residents of high-poverty ZIP Codes, following the adoption of universal insurance coverage.

To predict shifts of patients following the reconfiguration of hospital emergency services in northeast London, Congdon (2001) proposed an innovative Bayesian modeling framework using an unconstrained gravity model. A potential measure of accessibility to hospitals, similar to the potential measures discussed earlier in this chapter, was used in modeling patient flows. The model also included ward-level indicators of deprivation and age structure. Predicted patient flows were compared with actual flows that were recorded after the reconfiguration of hospital services took place. Observed and predicted flows fit quite well, indicating the utility of the model for forecasting purposes.

These studies illustrate the range of gravity model applications in the context of changing health care delivery systems. As health facilities open and close, and as managed care and other new forms of health care delivery affect affordability and access to care, gravity models offer a valuable forecasting tool. Using the models in a predictive context raises several important issues, however. Because geographical patterns of health care utilization vary for diverse population groups, for different types of health services, and in different places, it is crucial that the models be tailored to the particular study context (Handy & Niemeier, 1997). In general, a gravity model should fit as closely as possible the type of service, population, and geographical area in which it is applied. The accuracy of gravity models also depends on the level of aggregation of the data on which they are calibrated. Models estimated from individual or small-area data are generally thought to be more accurate for prediction since they better describe the forces that influence individual health care choices (Handy & Niemeier, 1997). However, many gravity models used in health planning rely on patient origin data at the county or ZIP Code level or on individual data geocoded to large geographic zones. Errors arising from spatial and nonspatial data aggregation effects reduce the accuracy and utility of model forecasts.

Small-Area Variation in Health Care Utilization

A large body of research demonstrates that rates of utilization for specific types of health services or medical-surgical procedures vary substantially from place to place in the United States (Wennberg, Fisher, & Skinner, 2004). The authors of the *Dartmouth Atlas of Health Care 1998* (Wennberg, 1998, p. 2) go so far as to say that "in health care, geography is destiny." For many Americans, the quantity, quality, and type of health care received depend greatly on the capacities and practices of local health service providers. Substantial geographic inequalities in health care spending have also been documented.

The Dartmouth project on small-area variations uses GIS to create geographic areas for the comparison of utilization rates. Starting with Medicare data by ZIP Code, the ZIP Codes are grouped into hospital service areas, areas in which the Medicare population primarily uses a particular hospital. In turn, hospital service areas are aggregated into hospital referral regions, regions in which the bulk of population was referred to the given hospital for high-level surgical procedures like neurosurgery. The regions reflect actual patterns of hospital utilization (Figure 9.16) with adjustments to ensure contiguity and minimum population size. These regions form the base for statistical analysis and mapping of geographical variations in service utilization at the national scale.

Their mapping of health care utilization patterns of Medicare recipients reveals for some types of procedures, a two- or threefold variation in hospitalization rates among geographic areas, even after adjusting for age, gender, and race. Spatial variation is less for medical and surgical procedures for which there is consensus among medical specialists on appropriate treatments; procedures for which there is less consensus exhibit much more spatial variation (Figure 9.17). The authors attribute much of the latter geographical variation to hospital capacity, as the pressure to fill beds stimulates utilization. Practice variations—differences in medical decision making—are also shown to be important. Differences in population need related to age, wellness, and socioeconomic deprivation account for roughly one-fourth of the regional disparities in utilization and spending (Wennberg, Fisher, & Skinner, 2004). By documenting the significance of local variations in health care supply and practice decisions in explaining geographic differences in costs and utilization, this research highlights the need for supply-side interventions to control health care costs (Gawande, 2009).

Although the Dartmouth research has been highly influential, by focusing on Medicare recipients, it considers a population that faces few economic barriers to accessing health care. Utilization patterns are likely to differ for other population groups. In addition, although the Medicare program removes most economic barriers to health care, social, cultural and geographical barriers may limit utilization by Medicare recipients. These other types of barriers need to be carefully examined before any conclusions can be drawn about geographic variation in health care utilization among Medicare recipients.

For other population groups, economic, social, and geographical access barriers can lead to substantial differences in rates of health care utilization. Such variations have been documented for a wide range of health care procedures in a variety of contexts. Important findings have come from research on ***ambulatory care sensitive conditions (ACSC)***—medical conditions that generally can be treated successfully in an ambulatory care setting (Billings et al., 1993). Hospitalization should only be required in the case of severe illness or emergency. Although there is some disagreement on which conditions should be considered as ambulatory care sensitive, the list generally includes asthma, diabetes, and hypertension, among others. Small-area variations in hospital use rates for these conditions reflect both underlying differences in illness, poor access to primary care, or poor quality of preventive care (Probst, Laditka, & Laditka, 2009). Individuals who have no health insurance or no regular source of primary care do not get early, preventive treatment and are more likely to end up in the hospital acutely ill. Hospitalization rates for ACSC differ sharply among small areas and

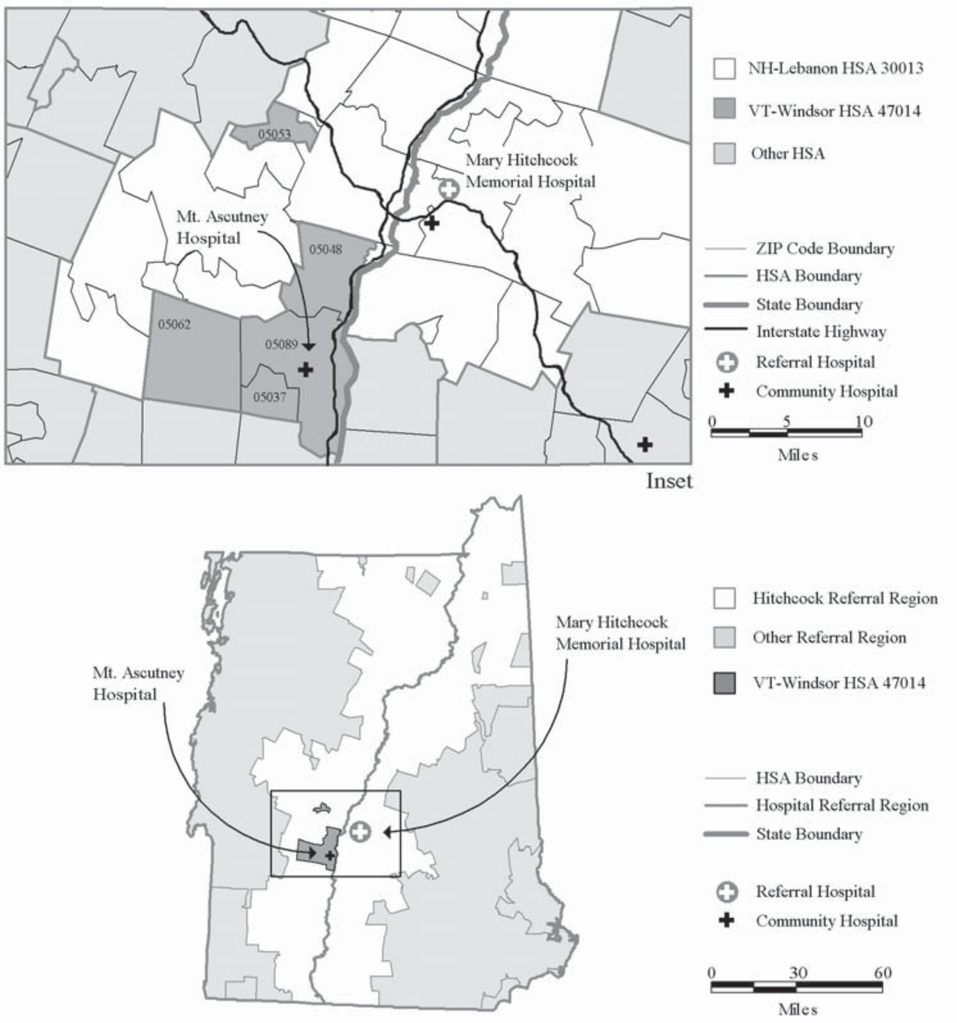


FIGURE 9.16. The inset shows a hospital service area (HSA) defined by patterns of utilization of Medicare enrollees by ZIP Code area. Medicare enrollees in five ZIP Code areas most often used the Mt. Ascutney Hospital in Windsor, Vermont. To preserve geographic contiguity in hospital service areas, ZIP Code area 05053 was reassigned to a different hospital service area. The service areas of community hospitals like the Mt. Ascutney Hospital are nested within the larger service areas of referral hospitals like the Mary Hitchock Memorial Hospital in Lebanon, New Hampshire, as shown in the main map. From Wennberg. (1998). Copyright 1998 by The Trustees of Dartmouth College. Reprinted by permission.

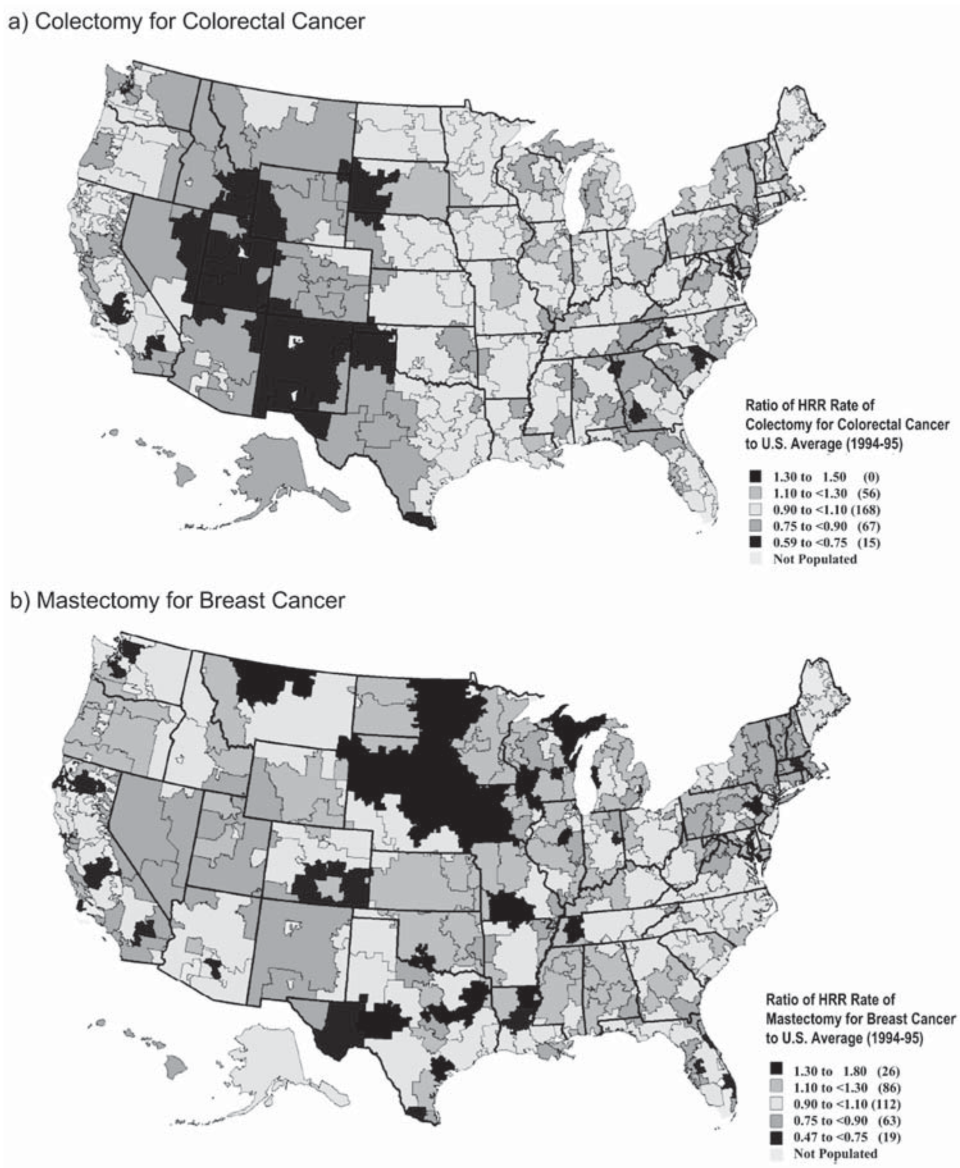


FIGURE 9.17. Utilization rates for different surgical procedures vary in relation to the U.S. average. Rates of colectomy for colon cancer showed little variation. As the map legend for Figure 9.17a shows, no regions had rates 30% or more above the national average, and only 15 regions – mostly in the West – had rates more than 25% below the national average. Rates of mastectomy for breast cancer in Figure 9.17b showed greater variation, with 26 regions in the upper central United States having rates 30% or more above the national average and 19 regions having rates more than 25% below the national average. From Wennberg. (1998). Copyright 1998 by The Trustees of Dartmouth College. Reprinted by permission.

are strongly correlated with socioeconomic status. The risk of hospitalization for ACSC is much higher in low-income areas and among people who have no health insurance (Pappas, Hadden, Kozak, & Fisher, 1997). Hospitalization for these chronic health problems can be viewed as a failure of the medical care system.

As in research on geographic variations, GIS can be used in studies of ACSC for managing the large spatial data sets that are required for examining small-area variations, for creating meaningful geographic areas to analyze, for modeling the impacts of geographical barriers on utilization, and for presenting display and visualization. The sensitivity of findings to scale and area boundaries (the modifiable area unit problem) can also be examined. The biggest challenge is interpretation—how to make sense of the geographic variations in health service use evident on the map. Geographic variation by itself is not surprising; the essential question is what processes give rise to it. The ACSC and Medicare literatures offer sharply different interpretations of similar patterns, the one emphasizing overreliance on hospitals caused by poor access to health care, and the other excess utilization caused by provider decisions. Research on individual and provider behaviors in varying geographical and health care contexts is needed to sort out these different interpretations.

Conclusion

Differential access to health care has been an important theme in public health policy in the United States and other countries for many years. The access problems of rural residents who often travel long distances to the nearest health care provider are well documented, as are the problems of low-income urban residents whose choices are limited by time–space constraints, lack of insurance, and poor transportation access. The restructuring of health care and efforts to control health care costs will continue to alter these patterns, yet the implications are poorly understood. By documenting changes in service accessibility in their geographical and social contexts, and by analyzing the differential impacts on population groups and places, GIS can play an important role in understanding evolving patterns of access and their consequences for health and well-being.

CHAPTER 10

Locating Health Services

The geography of health services delivery is an essential component of GIS applications in public health. To the extent that our information about health problems is obtained through medical care contact, our understanding of the distribution of health problems is filtered by the geographical distribution of health services and geographical factors that affect their functioning and utilization (Shannon, 1980). In addition to documenting the geographical variations in access to health services described in Chapter 9, GIS analyses have addressed issues in health services planning (McLafferty, 2003). Concern for the organization of health services is a logical outgrowth of the study of health and disease. Describing patterns of environmental contamination or uncovering the causes of disease leads us to intervention and prevention. Activities designed to prevent or address health problems include education, enforcement, and environmental modification, in addition to medical care delivery. As long as our activities occur in time and space, knowing how patterns of health, disease, and health services characterize regions will be essential to our efforts to advance human health. Like the environmental systems through which human populations are exposed to disease, health service systems have important geographies that can be effectively modeled using GIS.

As discussed in Chapter 9, the location of health services is a key factor affecting accessibility to care. The way we choose to model the distribution of health services influences the identification of underserved areas. It also influences our decisions about where additional health professionals and facilities should be located.

This chapter considers the basic components and dimensions of health service delivery systems and how they can be modeled. Given a geographical distribution of people who need access to some type of health service and a set of objectives for providing that service, patterns of health service organization can be evaluated and managed. Location–allocation models have been developed and applied to the problems of health services delivery. This chapter reviews some basic models and their use in health services research. Issues in integrating these models into a GIS are also considered. GIS can become spatial decision support systems in the health services planning process, allowing decision makers to explore complex, multiobjective problems. These techniques have been used to design and evaluate preventive health services, health services for acute and chronic health problems, and health services needed to respond to emergencies and disasters.

The development of GIS has coincided with important changes in health services delivery in the United States. Throughout the 1960s and 1970s, the federal government's role in health services delivery expanded dramatically even as many public health programs like infectious disease surveillance experienced funding cuts. In 1965, the Heart Disease, Cancer and Stroke Amendment authorized the establishment and maintenance of Regional Medical Programs, introducing the concept of regionalizing medical care in the United States. The programs were intended to promote cooperative arrangements among medical schools, research institutions, and hospitals for research and training. Fifty-six regions were established, covering the entire country, with most programs based at or near university medical schools. Federal financial support for graduate medical education increased, and the federal government became a major purchaser of health services through Medicare and Medicaid (Kovner & Knickman, 2008). The Health Care Planning and Development Act, passed in 1976, marked the culmination of almost two decades of federal support for health services by creating health system agencies across the United States. At the beginning of the 1980s, that federal support was eliminated, and the federal focus on health services shifted from subsidizing the expansion of hospital services in communities to cost containment, deregulation, and privatization.

Major insurers and providers of medical care services in the private sector in the United States have used GIS technology for institutional health services planning (Sandrick, 1998; Kennedy, 1999), but these applications have not generally been described in the research literature. The reason is partly that the locations of service centers, the structure of provider networks, and patient-origin patterns represent important business information that would be of value to competitors in a privatized system. At the same time, agencies in the public sector have recognized weaknesses in methods developed in the 1970s to identify underserved areas.

At the beginning of the 21st century, reform of the U.S. health care system again moved to the forefront of the national political agenda (Mitka, 2009), given the weaknesses in the U.S. system compared to the systems of other countries (Nuwer, 2008). Although approaches to paying for health care have dominated the debate, inequalities in the geographical supply, cost, and distribution of health providers have also been highlighted (Goodman & Fisher, 2008). Location modeling based in GIS can contribute to health system management and reform efforts.

Health Care Shortage Areas

During the period of the expanding federal role in health services financing and delivery, the federal government adopted two main systems for identifying locations where barriers to obtaining primary health care exist (General Accounting Office, 1995). The system for designating ***Health Professional Shortage Areas (HPSAs)*** was first used in 1978 to direct placements of National Health Service Corps employees to counties or facilities with a critical shortage of physicians. By the mid-1990s, close to 30 other federal programs had adopted the HPSA approach. The ***Medically Underserved Areas (MUAs)*** system was developed in 1976 to identify areas eligible for federally funded community health centers; the Community Health Center program was the system's main user. The definition was later expanded to include ***Medically Underserved Populations (MUPs).***

An evaluation of these approaches to identifying underserved areas conducted in 1995 concluded that these methods did not effectively identify areas with primary care shortages or target resources to benefit the underserved (General Accounting Office, 1995). Instead of identifying specific populations in need of care and the system resources available to meet that need, these approaches began with a place—a geographic area like a county or a specific facility like a prison—and characterized the place based on medical resource availability and population characteristics. The data used to describe the number of available physicians in a place were often neither timely nor accurate, especially when compared with information in health directories like those discussed in Chapter 3. As a result, analysts relying on these methods were not able to identify who was underserved and why.

Since the late 1990s, federal agencies have been involved in an effort to improve the designation of areas and populations that are underserved, and there is a continuing emphasis on calculating provider/population ratios for administrative units like counties or census tracts (Ricketts et al., 2007). Geographic information systems have been used to demonstrate methods for improving the designation of health professional shortage areas and medically underserved areas, especially with respect to primary care physician services (Juarez, Robinson, & Matthews-Juarez, 2002). Circular buffers of census tract centroids were used to model floating catchment areas representing the possible movement of patients to physicians across the boundaries of census tracts in a nine-county study area in northern Illinois (Luo, 2004). Physician/population ratios were calculated for areas defined using different buffer distances. The larger the radius, the fewer areas were identified as shortage areas. This occurs because the physician/population ratio is scale dependent and the greatest variability in the ratio occurs at the local scale. As the radius increases, more physicians and population are located within the catchment area, and the physician/population ratio converges to the ratio for the study area as a whole. A subsequent study applied similar techniques to identify physician shortage areas in the entire state (Wang & Luo, 2005).

In 2008, the Health Resources and Services Administration proposed a new rule for revising and consolidating criteria and procedures for designating Medically Underserved Populations and Health Professional Shortage Areas. The proposed rule describes a two-step process. First, states must identify so-called rational service areas defined in terms of U.S. Census Bureau geographical units. Once the service areas are defined, the data on providers and populations within the areas are collected and evaluated. This rule was not adopted, and a new notice of proposed rulemaking will likely be issued. In the meantime, an interactive website enables users to search for shortage areas by state and county and to identify whether or not an address entered by the user is within a shortage area (Health Resources and Services Administration, 2010). A search returns a list of geographic areas like census tracts included in the selected shortage areas.

Efforts to address concerns about physician shortages in the United States are being made, primarily at the institutional and state levels (Iglehart, 2008). Workforce issues affect the health care systems of many countries (World Health Organization, 2006). Increasingly, national physician supplies are linked by physician migration (Onyebuchi, Ogbu, & Okeke, 2008). The United States, Canada, Australia, and the United Kingdom are four major destinations for migrating physicians, and international medical graduates account for roughly a quarter of the physicians in these countries. Of this group of international medical graduates, 40 to 75% came from countries with lower incomes.

GIS provide the means to capture and verify health service capacity in locally defined service areas using data that may not be available at the national level. In some cases, health care facility service areas are defined based on utilization patterns as described in Chapter 9 (see Figure 9.10). In a study of regional variations in neonatalogists, beds, and low-birthweight newborns, neonatal intensive care regions in the United States were defined to represent markets for neonatal intensive care services (Goodman, Fisher, Little, Stukel, & Chang, 2001). Counties were assigned to regions with at least one neonatal intensive care unit based on the birth locations of very-low-birthweight infants. Regional boundaries were adjusted to minimize cross-boundary flows for care, resulting in a set of regions matching neonatal intensive care resources to the newborn populations needing care. Other research has tested methods for defining local regions for health care planning that yield nonoverlapping regions for general practitioner services based on defined population size and containment criteria (Shortt, Moore, Coombes, & Wymer, 2005). A challenge in defining regions for health care planning and analysis is that one set of regions may not be equally meaningful for all age groups and all health services (Guagliardo, Jablonski, Joseph, & Goodman, 2004). Regardless of how the region that a health services system serves and zones within it are defined, GIS functions can be used to display the components of health service systems, to investigate the distributions of specific client populations affected by different barriers to care, and to incorporate location models that assess how well the distribution of services fits the distribution of populations in need.

Components and Dimensions of Health Service Delivery Systems

Technological advances in communication have made it possible to distribute information about health, disease, and health services through the mass media and the Internet (Andreassen et al., 2007; Atkinson, Saperstein, & Pleis, 2009; Moreno, Ralston, & Grossman, 2009), and have even made it possible for patients and providers to consult in real time over long distances (Balas et al., 1997; Wilson & Branigan, 1999; Scott et al., 2007; Lam & Poropatich, 2008; McGeady, Kujala, & Ilvonen, 2008). Nevertheless, the delivery of many health services still requires some form of direct contact between the provider of the service and the person who benefits from it. As outlined in the Introduction, direct personal contact can only be achieved if people's activities can be coordinated in time and space.

A ***service delivery system*** is "a cluster of diverse agencies within an organizational network that provides services to a common client population" (Alter, 1988, p. 91). The components of a health services delivery system include the client or patient population, the provider agencies, and the relationships that connect clients to providers. An early application of GIS technology (Achabal, Moellering, Osleeb, & Swain, 1978) illustrates how interactive computer graphics can be used to display the locations of hospitals within a service area, the spatial distribution of the residential population, and the allocation of patients to service centers when patients are assigned to the hospitals so that no individual is required to travel more than a prespecified distance and no hospital is overutilized (Figure 10.1).

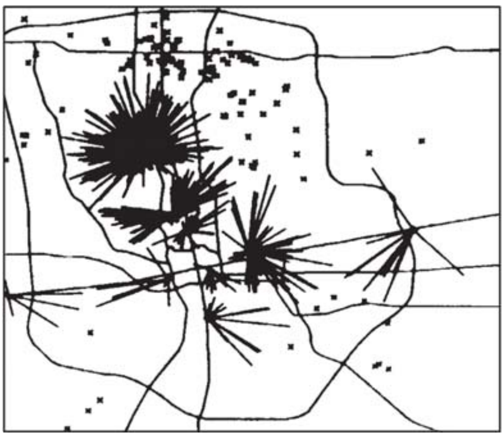


FIGURE 10.1. Allocation of residents to existing hospitals so that no individual must travel more than 6 kilometers and no hospital is overutilized. Reprinted from *Social Science and Medicine, 12*(1D), Achabal, D., Moellering, H., Osleeb, J. and Swain, R., Designing and evaluating a health care delivery system through the use of interactive computer graphics, 1-6, Copyright (1978), with permission from Elsevier.

As this example shows, the components of a health service system are usually modeled in GIS as ***objects*** (see Chapter 2). At the community scale, health service facilities are represented as point features, populations served are represented as points or aggregated as count data for areas, and the assignments of service users to service providers are represented as lines. These points, lines, and areas form a ***network*** space (see Chapter 2) for evaluating locational equity and efficiency in an existing or planned service delivery system.

Health service organizations must coordinate their activities in time and space as much as individuals do. Some organizations, such as testing and counseling centers, operate at one or more fixed locations. These service centers represent nodes in the activity spaces of service providers and service users who travel to service sites. Other individuals or organizations that provide services to people move around or circulate (the visiting nurse or physician making a house call, the emergency medical response team, the home-delivered meals service). The activity patterns of the services can be evaluated using time budget approaches like those described in the Introduction for analyzing individual travel and activity patterns. In addition to the location or set of locations where services are provided, there are other dimensions of community institutions (Alter, 1988) that have geographical implications.

Size can be interpreted as the number of service sites. In the case of the system modeled by Achabal et al. (1978), the hospital service system included nine major hospitals in the Columbus, Ohio, metropolitan area. The relationships between total size of a system (measured as the number of service sites) and ***capacity***, or volume of service, are not always straightforward. Service centers located in communities of similar size can have different capacities and can provide varying volumes of service depending on eligibility requirements and intake.

Threshold requirements, capacity constraints, and minimum standards are important characteristics of health services. A ***threshold requirement*** represents the minimum demand or volume of service needed to sustain service delivery. For example, the minimum number of deliveries in a community hospital obstetrics unit needed to make the provision of quality service viable is a threshold requirement. A ***capacity constraint*** is a maximum limit on the volume of service that can be provided. For example, the total number of hospital beds limits the number of patients who can be accommodated at any one time. An example of a ***minimum standard for service delivery*** would be that no person should live more than 10 minutes travel time from a first-response emergency service provider.

When there are threshold requirements, capacity constraints, and minimum standards for planned facilities, the optimal number, location, and capacities of service centers will be strongly influenced by the underlying geographical distribution of the population to be served. In fact, depending on the distribution of that population, it may be geographically infeasible to meet the threshold, capacity, and minimum service standards identified. This would happen if there were a small residential neighborhood that had too few people to meet the threshold requirement for a local facility and the neighborhood was located more than the desired travel distance or time from an existing provider. If provision of the service were necessary, one of the requirements would have to be broken. Either taxpayers would pay a subsidy to run a small health service center or patients would pay in excessive travel distance or time.

Centrality is another dimension of social service organizations. When the total volume of users flows through a single organization, that organization has a high degree of centrality. There is a strong relationship between differentiation of a service organization's functions and its degree of centrality.

The geographical relationships among service centers with varying degrees of centrality is the focus of central place theory (Christaller, 1933; King, 1984) and subsequent research on human settlement systems and public and private service systems (Foot, 1981; Ghosh & Rushton, 1987). When a population is uniformly distributed, those service centers with smaller threshold requirements—for example, physicians' offices—will be more common in the landscape and spaced relatively close together. Those service centers with larger threshold requirements—for example, tertiary care centers—will be less common in the landscape and spaced relatively far apart. The service areas of small activity sites are commonly nested in the service areas of larger activity sites.

The relationships between the number of service sites and the number and level of services provided at each site have been explored using GIS. A study of the provision of mental health advocacy services in London, England, mapped the locations of service sites and their catchment areas as well as the provision of advocacy services in acute care and community settings (Foley & Platzer, 2007). Although each borough was served by at least one local service and by specialist service covering all of London, no single borough had the full range of provision and no single organization offered a full range of services or delivered its services in the full range of settings. The study also identified significant changes in the organization of services over time. Close to a quarter of organizations providing advocacy in 1998 were operating differently or were no longer delivering services at all in 2002.

The changing pattern of service provision at locations was analyzed in a study of the spread of health services and fertility change in rural Nepal from 1945 to 1995 (Brauner-Otto, Axinn, & Ghimire, 2007). A GIS application modeled accessibility to four types of maternal and child health services—child immunization, family planning, prenatal care, and pregnancy/delivery assistance—provided by a range of organizations at different locations at different points in time. Different women had different levels of access to different combinations of services over their lifetimes. The analysis uncovered significant independent relationships between the availability of maternal and child health services and family planning services and fertility limitation. The pathways individuals follow to access services are also important.

Integration refers to the relationships or forward and backward linkages among units within a system. In the case of health services systems in the United States, alternative pathways through the service hierarchy have generally been very common. Residents of a particular neighborhood are not usually "assigned" to a particular service center, although managed care systems may attempt to direct patients through the health service hierarchy of providers.

GIS have been used to investigate integration in health service systems. Researchers in Japan studied patterns of pediatric inpatient referral from secondary care hospitals to a tertiary care university hospital (Toyabe & Kouhei, 2006). The analysis showed that the university hospital functioned as both a secondary and tertiary level facility. Patients who lived near the hospital were often directly admitted to the hospital, and patients who lived farther away from the hospital were more likely to use the hospital as a tertiary care provider. Spatial differences in referral to care after discharge were also observed among the patient population.

Service linkages are especially important in planning for special populations recently released from inpatient treatment programs or prisons. A study using GIS modeling and logistic regression analysis found that discharge of patients with mental health and substance use disorders from an acute inpatient unit to the preadmission address, along with other individual and neighborhood factors, reduced the likelihood of attending a program of postdischarge outpatient treatment (Stahler et al., 2007). Analysts in New Jersey used GIS to assess the number, demographic characteristics, and needs of a parolee population in Newark, New Jersey, in light of the availability, location, and characteristics of health and human services agencies providing services that would assist prisoner reintegration into society (Mellow, Schlager, & Caplan, 2008). Spatial distribution of services and the degree of spatial overlap in service provision to the client population were documented.

Client Population Distribution

An important geographical pattern to investigate in planning and evaluating health services delivery systems is the distribution of the population who will be receiving care. The residential distribution of the population is usually considered the most relevant in health services planning, especially for home-delivered services, but also for services requiring the help seeker to travel to a fixed service delivery site. As noted in earlier chapters, the residential distribution of population is rarely uniform. GIS are effective tools for developing useful representations of population distribution.

This is particularly important for services designed to meet health problems affecting particular age or age/sex cohorts because these groups are probably not distributed equally across the distribution of the total residential population. Mammography, for example, is recommended for women aged 40 years and older on a regular basis. A map of the distribution of women aged 40 and older shows that the distribution of this age/sex group differs from the distribution of the total population (Figure 10.2).

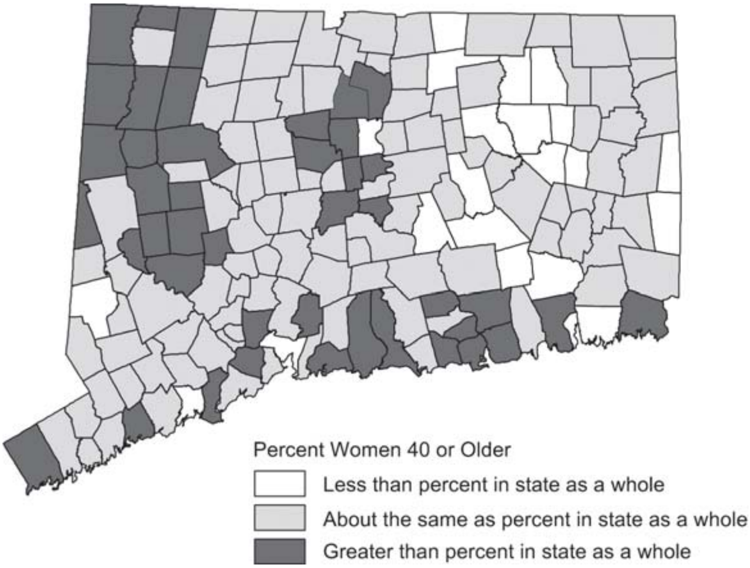


FIGURE 10.2. Mapping the age–sex specific need for mammography services. The map shows that older women are concentrated in towns in the northwest, along the coast, and near Hartford. These towns have a higher percent of women 40 or older in their populations than in the population of the state as a whole. The towns with the highest concentrations have 21% of the state's total population, based on data from the 2000 census, but they have 24% of the state's population of women 40 or older.

Travel distance is an important consideration in locating health service centers. Time spent in travel—either by providers or by help seekers—increases the cost of providing care and decreases accessibility. For some types of care, like emergency medical service or home-delivered meals, there may be critical service response times after which the service is of little or no value. The importance of avoiding unnecessary travel means that opening service or dispatch sites at central locations within the distribution of the client population is a key objective of health services planning.

The Meaning of "Centrality" in Health Service Facility Location

The different operational meanings of a "central" location within a distribution of points can be explored through a simple example. In this hypothetical setting, the planning task is to open a single service center to meet the needs of nine people requiring care. The residential distribution of the population is shown in Figure 10.3. Again, to simplify, the population is distributed in a one-dimensional space, along a hypothetical coastline or the foothills of a mountain range. In the

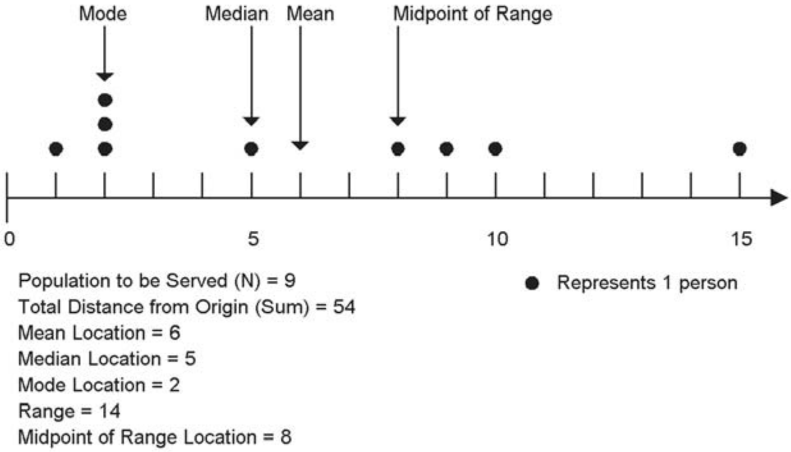


FIGURE 10.3. An example showing the location of a single central facility to serve nine clients distributed along a single dimension. The location of the facility changes depending on how centrality is defined.

example, location is measured in absolute terms from an arbitrary origin located outside the range of the data. Distance is calculated as difference between the starting and ending points along the number line.

Four measures of central tendency are available to define the "center" of this distribution of population: mean, median, mode, and midpoint of the range. To calculate the ***mean*** as the center of the distribution of residential locations, we would sum the distances of each residence to the origin and divide by the total number of residential locations. In this case, the mean is 6 and the "central facility" would be located at position 6, a place where no one actually resides. The mean —and therefore its associated location — has the special property of minimizing variance in travel distance. Position 6 is the location ensuring that the variation in distances people must travel to receive service is minimized; that is, the sum of the squared distances from each residence to position 6 is a minimum. Locating the facility at any other position would result in a greater sum of squared distances from the facility location.

To calculate the ***median***, we can arrange the distances from the origin in order from lowest to highest and identify the distance value in the middle of the distribution. In this example, the median position is position 5. The median has the special property of minimizing total distance. That is, the sum of the absolute differences between position 5 and the other locations is a minimum. Locating the facility at any other position would result in a greater total distance traveled to the facility location.

To calculate the ***mode***, we would identify the location that occurs most frequently in the residential distribution. This is position 2. The mode has the special property of maximizing access to the facility by locating it most conveniently for the greatest number of people. Position 2 is the location where the population to be served is concentrated.

Finally, to calculate the ***midpoint of the range***, we would first calculate the range of the distribution. The ***range*** is a measure of dispersion and is the difference between the highest and lowest values in the ordered distribution. In this case, the range is 14. The midpoint of the range is calculated by dividing the range in half and adding that value to the lowest value in the distribution. When these calculations are performed for the hypothetical example, the midpoint of the range is 8. Locating the facility at position 8 minimizes the maximum distance that any single person would have to travel to obtain care—in this case, 7 units of distance. Locating the facility at any other position would increase the maximum travel distance for the most remotely located individual.

These measures can also be computed in the bivariate space of the map (Figure 10.4) (Ebdon, 1985; Wong & Lee, 2005). As these hypothetical examples

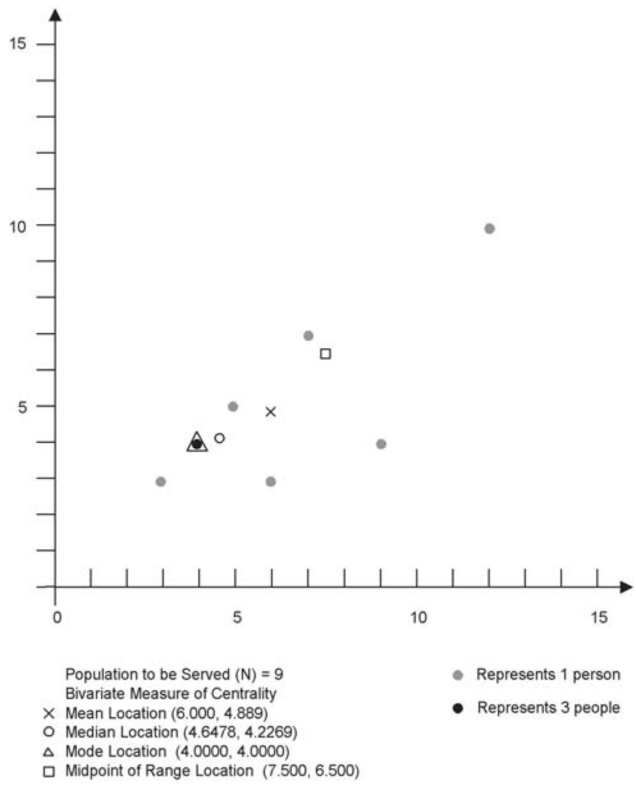


FIGURE 10.4. An example showing the location of a single central facility to serve nine clients distributed in a two-dimensional space. The location of the facility changes depending on how centrality is defined.

illustrate, there is more than one way to define a "central" location for a single facility depending on the particular travel distance function that the selected measure of centrality maximizes or minimizes. Facility locations based on measures of central tendency like the median or mode emphasize ***locational efficiency*** in the delivery of services because these measures minimize total travel effort or maximize accessibility (Morrill & Symons, 1977). Facility locations based on the mean or midpoint of the range emphasize ***locational equity*** in the delivery of services because these measures minimize variation in travel effort or reduce travel distances for those farthest from population centers.

Normative Models of Facility Location and Service Delivery

Normative Models and Mathematical Programming Methods

Normative models of facility location or service delivery do not seek to describe existing facility locations or flows. Instead, they are designed to identify the facility locations or flows that maximize or minimize a mathematical function that expresses the objective of the decision maker. ***Allocation models*** assume that facility locations are fixed—as they are in the short term—and identify the assignment of patients to facilities that maximizes or minimizes the objective function—for example, the assignment that minimizes total distance traveled to service sites. ***Location models*** seek the set of locations from among a set of candidate sites that maximize or minimize the objective function. ***Location–allocation models*** identify the optimal locations and assignments.

Location–allocation problems are solved through the application of mathematical programming techniques (Greenberg, 1978). ***Mathematical programming*** is a set of numerical methods for solving optimization problems. These methods are not based in multivariate inferential statistics. Most public health professionals have probably not received intensive training in these methods. Mathematical programming techniques, however, are applied in industrial and business planning to optimize various aspects of production, including facility location. They have also been used in choropleth map classification (Cromley, 1996) and health data anonymization (Wieland et al., 2008). Although other numerical methods, like calculus, can be used to solve an optimization problem (e.g., finding the minimum of an average cost function), mathematical programming methods are used when optimization problems involve quantities that cannot be negative. All location–allocation problems have these nonnegativity constraints. We cannot travel a negative number of miles to an outpatient facility or assign a negative number of patients to a hospital. Location–allocation models have been used within medical geography since the 1960s when the algorithms for solving them could be run on mainframe computers (Godlund, 1961; Gould & Leinbach, 1966; Rushton, 1975; Bennett, 1981; Mohan, 1983).

The Transportation Problem

A mathematical programming model is specified by an objective function and a set of constraints. One of the most commonly modeled problems is ***the transportation problem*** (Scott, 1970). It is an assignment or allocation problem because the geographical distributions of supply and demand are known and fixed. The objective of the transportation problem is to assign demand associated with a set of demand points to facilities that can supply the needed service (supply sites) so that the total cost of the assignment (total distance or travel time) is minimized (Table 10.1). This assignment is subject to the constraints that all demand must be served and the capacity of a supply site cannot be exceeded.

Mohan (1983) used the transportation problem to assess strategies for hospital location in the Durham Health District in England (Figure 10.5). The demand sites were represented by grid cells 1 kilometer square superimposed over the Durham Health District area; the volume of demand was total population per grid cell. In the initial analysis, the two existing supply sites (hospitals) were used, and the minimum aggregate travel distance and the average travel distance for each hospital were calculated based on an optimal assignment of demand sites to service sites. Once the minimum aggregate travel achievable with optimal use of the existing hospital system was established as a benchmark, the impact of adding or modifying the existing hospital system could be evaluated. Two alternatives were examined. In the first, a third hospital facility

TABLE 10.1. Mathematical Programming Formulation of the Transportation Problem

Objective function:	Minimize $Z = \Sigma_{i \in I} \Sigma_{j \in J} d_{ij} x_{ij}$
Subject to the constraints:	
All demand at a demand site must be served.	$\Sigma_{j \in J} x_{ij} \geq r_i$ for all i
The capacity at a supply site cannot be exceeded.	$\Sigma_{i \in I} x_{ij} \leq q_j$ for all j
The number of people assigned from a particular demand site to a particular facility site cannot be negative.	$x_{ij} \geq 0$ for all (i, j)

Where:

Z is the objective function.

I is the set of demand areas, usually nodes on a network, and the subscript i is an index denoting a particular demand area.

J is the set of candidate facility sites, usually nodes on a network, and the subscript j is an index denoting a particular facility site.

d_{ij} is the distance or time (travel cost) separating place i from candidate facility site j.

x_{ij} is the number of people from demand site i assigned to receive service at facility site j.

r_i is the total number of people to be served at demand site i.

q_j is the total capacity of facility site j to provide service.

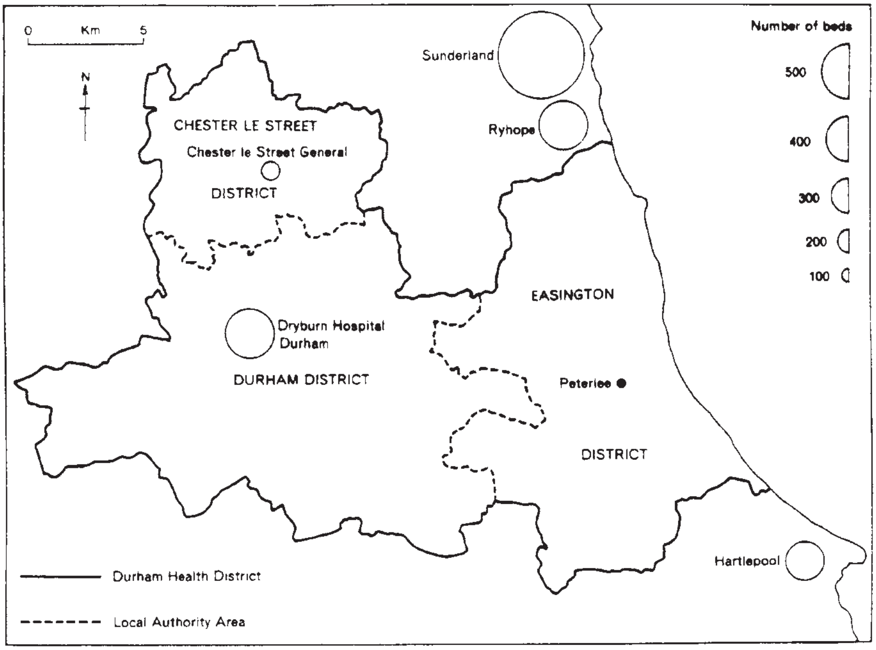


FIGURE 10.5. Hospital location in the Durham Health District showing the location of Peterlee New Town in relation to existing hospitals. Reprinted from *Social Science and Medicine, 17*(8), Mohan, J., Location-allocation models, social science and health services planning: An example from North East England, 493–499, Copyright (1983), with permission from Elsevier.

was "added." In the second, one of the two existing hospitals was retained, but the second existing hospital was "closed" and replaced with a facility at another location, Peterlee. Both of the alternative configurations resulted in substantial reductions in total travel time to obtain hospital service (Table 10.2). Locating a hospital at Peterlee instead of Chester-le-Street decreased aggregate travel distance by a third, from 1,570,021 to 1,054,316 kilometers.

As this research illustrates, total travel distance generally decreases as the number of service sites increases. But locating facilities in sparsely settled areas remote from large population centers may lead to facilities that are underutilized. The ***bounded transportation problem*** is a variant of the transportation problem that places a lower bound or service population threshold on each supply site as well as an upper bound on each site's capacity (Table 10.3). When a minimum level of service must be provided to ensure quality of care or economic viability, adding more service centers may actually increase the total travel cost of assigning patients to providers. This will happen if patients must be diverted to more distant service centers to ensure that threshold requirements are met there (Green, Cromley, & Semple, 1980).

TABLE 10.2. Aggregate and Average Travel Statistics for Various Combinations of Hospital Locations

Hospital locations	Population served	Aggregate travel distance (km)	Average travel distance (km)
Two existing sites			
Dryburn	162,606	1,431,227	8.80
Chester-le-Street	49,442	138,794	2.80
Total	212,048	1,570,021	7.40
Two Existing Sites and One New Site			
Dryburn	9,194	457,905	4.76
Chester-le-Street	49,439	138,745	2.80
Optimal site for third facility (Peterlee)	66,415	197,656	2.97
Total	212,048	794,306	3.74
One Existing Site and One New Site			
Dryburn	144,485	845,021	8.85
Optimal site for second facility (Peterlee)	67,363	209,295	3.09
Total	212,048	1,053,316	4.97

Note. Reprinted from *Social Science and Medicine*, *17*(8), Mohan, J., Location-allocation models, social science and health services planning: An example from North East England, 493-499, Copyright (1983), with permission from Elsevier.

Facility Location

MINIMIZING TRAVEL EFFORT

One of the most commonly modeled location problems is the ***p-median problem*** (ReVelle & Swain, 1970; Church & Sorensen, 1996). The objective of the p-median problem is to locate a given number of facilities among a set of candidate facility sites so that the total travel distance or time to serve the population assigned to the facilities is minimized, using the median as the measure of central tendency. Unlike the transportation problem, in which the number and locations of supply sites are known in advance, the p-median problem specifies only the number of facilities, p, to be located from a larger set of possible facility sites.

The solution is subject to a set of constraints. Every place where users of the service originate (every demand site) must be assigned to one and only one facility, ensuring that all service needs will be met. Each potential facility site must either receive or not receive a facility in the solution. The number of facilities located must equal the given number p exactly. These concepts can be written using mathematical programming notation (Table 10.4).

The following input is required to solve a p-median problem: the number of demand sites and the volume of demand at each site; the number of possible supply sites; the per-unit distance, time, or cost of travel from every demand site to every potential supply site; and p, the number of facilities to be opened.

TABLE 10.3. Mathematical Programming Formulation of the Bounded Transportation Problem

Objective function:	Minimize $Z = \Sigma_{i \in I} \Sigma_{j \in J} d_{ij} x_{ij}$
Subject to the constraints:	
All demand at a demand site must be served.	$\Sigma_{j \in J} x_{ij} = r_i$ for all i
The capacity at a facility site cannot be exceeded.	$\Sigma_{i \in I} x_{ij} \leq q_j$ for all j
The minimum level of service provided at a facility site must exceed a threshold level.	$\Sigma_{i \in I} x_{ij} \geq t_j$ for all j
The number of people assigned from a particular demand site to a particular facility site cannot be negative.	$x_{ij} \geq 0$ for all (i, j)

Where:

Z is the objective function.

I is the set of demand areas, usually nodes on a network, and the subscript i is an index denoting a particular demand area.

J is the set of candidate facility sites, usually nodes on a network, and the subscript j is an index denoting a particular facility site.

d_{ij} is the distance or time (travel cost) separating place i from candidate facility site j.

x_{ij} is the number of people from demand site i assigned to receive service at facility site j.

r_i is the total number of people to be served at demand site i.

q_j is the total capacity of facility site j to provide service.

t_j is the minimum amount of service (the threshold) that must be provided at facility site j.

The formulation of the problem reveals an important difference between the transportation problem and the p-median problem. All the variables in the transportation problem are continuous, but some of the variables in the p-median problem are discrete. The number of facilities to be located is an integer. It would not be possible to open half a facility. Similarly, a demand site is either assigned to a facility or it is not, so the decision variable to assign demand to a site is a zero–one integer variable. Integer programming problems are solved with different algorithms than those used to solve linear programming problems like the transportation problem.

In the research on evaluating hospital locations in Columbus mentioned earlier in this chapter, the optimal location for a new hospital among five candidate locations was identified by finding the location that minimized total travel distance for all user areas (Achabal et al., 1978). The candidate locations were concentrated in the north end of town after a model allocating population to the existing hospitals revealed that patients in that part of the city could not be allocated to existing hospitals.

TABLE 10.4. Mathematical Programming Formulation of the *p*-Median Problem

Objective function:	Minimize $Z = \Sigma_{i \in I} \Sigma_{j \in J} a_i d_{ij} x_{ij}$
Subject to the constraints:	
An individual demand site must be assigned to a facility.	$\Sigma_{j \in J} x_{ij} = 1$ for all i
Demand must be assigned to an open facility.	$x_{ij} \leq \Sigma x_{jj}$ for all (i, j)
Exactly p facilities must be located (the number of communities assigned to themselves equals the number of facilities to be located).	$\Sigma_{j \in J} x_{jj} = p$
All demand from an individual demand site is assigned to only one facility.	$X_{ij} = (0,1)$ for all (i, j)

Where:

Z is the objective function.

I is the set of demand areas, usually nodes of a network, and the subscript i is an index denoting a particular demand area.

J is the set of candidate facility sites, usually nodes of a network, and the subscript j is an index denoting a particular facility site.

a_i is the number of people from demand site i.

d_{ij} is the distance or time (travel cost) separating place i from candidate facility site j.

x_{ij} is 1 if demand at place i is assigned to a facility opened at site j or 0 if demand at place i is not assigned to that site.

p is the number of facilities to be located.

For some facility location problems, however, the p-median problem may not be appropriate. The optimal solution minimizes total travel effort but it does not necessarily limit the travel effort of an individual service user. For many health services, critical travel or response time standards have been established. How can facilities be located to ensure that all users or as many users as possible will be served within the critical travel or response time?

One approach involves adding a constraint to the original p-median formulation to require that each demand site be served by a supply site within the critical distance or time. This formulation is known as the ***p-median with maximum distance constraints*** (Khumawala, 1973; Hillsman & Rushton, 1975). In the research on hospital locations in Columbus just mentioned, the researchers applied a maximum service distance of 6 kilometers in an urban area (Achabal et al., 1978). In the study of hospital location in rural Ohio, a maximum service distance of 40 miles was used (Green et al., 1980). Once these distance constraints were added, the question arose as to whether p facilities—the number specified at the outset of analysis—would be sufficient to ensure that all users could be covered within the maximum service distance. If not, the problem would have an infeasible solution. The need to determine the minimum number of facilities that would be required to cover a set of demand sites gave rise to the location set covering problem (Church & ReVelle, 1976).

MAXIMIZING COVERAGE AND EMERGENCY SERVICE LOCATION

The ***location set covering problem*** (LSCP) identifies the minimal number and the locations of facilities required to "cover," or provide service to all users, within a prespecified critical travel distance or time (Toregas, Swain, ReVelle, & Bergman, 1971). The objective function of the LSCP is to minimize the total number of facilities to be "opened" from a set of potential facility locations (Table 10.5). This solution is subject to the constraint that every demand site must be within the critical distance or time of at least one open facility.

GIS tools have been used to model coverage areas of facilities. These analyses are useful in identifying areas that are not currently covered by an existing facility within the specified critical distance or time. Coverage areas for methadone clinics in Hong Kong were identified using a 1.5-kilometer straight-line buffer (Pang & Lee, 2008). It is also possible to identify coverage zones based on street network distance, as described later in this chapter and in Figure 10.8. GIS analyses also show the trade-off between the critical service time or distance and the number of facilities required to cover a population. When the service standard was relaxed from 5 miles to 10 miles in a study of dental facilities in Ohio, the number of areas not served by dentists fell from 307 to 45 (Horner

TABLE 10.5. Mathematical Programming Formulation of the Location Set Covering Problem

Objective function:	Minimize $Z = \Sigma_{j \in J} x_j$
Subject to the constraints:	
An individual demand site must be within the critical service distance or time of at least one open facility site.	$\Sigma_{j \in N_i} x_j \geq 1$ for all i
A candidate facility site must be either opened or closed.	$x_j = (0, 1)$ for all j

Where:

Z is the objective function.

I is the set of demand areas, usually nodes of a network, and the subscript i is an index denoting a particular demand area.

J is the set of candidate facility sites, usually nodes of a network, and the subscript j is an index denoting a particular facility site.

x_j is 1 if a facility is opened at candidate site j or 0 if a facility is not opened at candidate site j.

N_i is the set of facilities where the distance between demand site i and candidate facility site j is less than the critical distance or time, or $d_{ij} \leq s$.

d_{ij} is the distance between a demand site i and a candidate facility site j.

s is the critical service response distance or time.

& Mascarenhas, 2007). It was estimated that covering the 45 sites would require as few as 24 additional service sites.

The mathematical formulation of the LSCP reveals that it, like the p-median problem, is an integer programming problem because the decision variable x_j is a zero–one integer variable: a candidate facility site will either be opened or it will not be opened in the solution. The following input is required to solve a LSCP: the number of demand sites, the number and location of possible supply sites, the critical service distance or time, and the distance or time from each demand site to each possible supply site. The last-named makes it possible to identify the set of all possible supply sites that can serve an individual demand site within the critical service constraint.

An obvious limitation of the LSCP is that the number of facilities required to cover 100% of the population may be beyond the budget available for providing the service. To address this problem, analysts developed the ***maximal covering problem***, incorporating elements of both the p-median problem and the LSCP (Church & ReVelle, 1974). The objective function of the maximal covering problem is to locate p facilities within a set of possible supply sites so that the number of users receiving service at a facility located within a critical service distance or time is as large as possible, or maximized, which is equivalent to minimizing the number of users beyond the critical distance (Table 10.6). The data required to solve a maximal covering problem are the same as for the LSCP, with the addition of the volume of demand at each demand site and the number of facilities to be opened.

Maximal covering problems are particularly appropriate for planning and evaluating the location of emergency service facilities. Emergency medical service (EMS) delivery is only effective if the response can be made within a critical time period. The maximal covering formulation was used as part of a study conducted in Austin, Texas, to determine what EMS services should be provided, by whom, using what numbers and types of equipment, and sited at which locations (Eaton, Daskin, Simmons, Bulloch, & Jansma, 1985). A software package to perform analyses of emergency call histories from various zones in the city, computer mapping programs, and a program to solve the maximal covering problem were used in the study.

Data from the call history analysis became inputs to both the computer mapping programs and the location model. Eight surrogates for EMS demand were modeled: total calls, critical calls, noncritical calls, total population, black population, Hispanic population, Anglo population, and elderly population. These were modeled with a range of vehicle fleet sizes and a variety of critical response times. Clear trade-offs in service became apparent. Locating 12 vehicles to maximize coverage of black residents allowed for 97% of the black population to be served within 5 minutes. The 12 vehicle sites that best covered the Anglo population would reach only 60% of the black population within 5 minutes. After decision makers considered a variety of options, the final plan agreed upon deployed 12 vehicles in a two-tiered advanced and basic life-support system.

TABLE 10.6. Mathematical Programming Formulation of the Maximal Covering Problem

Objective function:	Minimize $Z = \Sigma_{i \in I} a_i y_i$
Subject to the constraints:	
An individual demand site must be within the critical service distance or time of at least 1 open facility site or it is not covered.	$\Sigma_{j \in N_i} x_j + y_i \geq 1$ for all i
Exactly p facilities must be located	$\Sigma_{j \in J} x_j = p$
A candidate facility site must be either opened or closed.	$x_j = (0, 1)$ for all j
A individual demand site is either covered within the critical service distance of a facility or it is not.	$y_i = (0, 1)$ for all i

Where:

Z is the objective function.

I is the set of demand areas, usually nodes of a network, and the subscript i is an index denoting a particular demand area.

J is the set of candidate facility sites, usually nodes of a network, and the subscript j is an index denoting a particular facility site.

a_i is the number of people at demand site i.

N_i is the set of facilities where the distance between demand site i and candidate facility site j is less than the critical distance or time, or $d_{ij} \leq s$

s is the critical service response distance or time.

x_j is 1 if the facility is opened at site j or 0 if the facility at site j is not opened.

y_i is 1 if the demand site i is not covered by an open facility within s and 0 if the demand site i is covered by an open facility within s.

p is the number of facilities to be located.

Research on emergency services response recognizes that response time has several components in addition to travel time from the response vehicle departure point to the scene. Responders need time to prepare to leave after a call is received. In rural areas where personnel are volunteers, additional time may be needed. In urban areas, response times need to take into account the time it takes to travel through multistory buildings to reach an individual in need or ***vertical response time***. A prospective observational study of response times in the New York City 911 emergency medical services system found that time intervals from on-scene arrival to patient ranged from 0.5 minute for outdoor scenes to 2.8 minutes for residential buildings (Silverman et al., 2007). Overall, 28% of the actual response time was accounted for by time to reach the patient after arrival at the scene.

MAXIMIZING MEDICAL OUTCOMES

When the underlying distribution of demand for health services is not uniform, siting facilities to ensure coverage within the desired travel distance or time may lead to low utilization of facilities in less densely populated regions. Aside from the economic implications of this situation, there are also implications for health outcomes. The "***patient volume effect*** ... refers to the relationship between the number of patients treated in a facility and the rates of mortality and morbidity among those patients" (McLafferty & Broe, 1990, p. 298). Enhancing coverage by ensuring geographical accessibility and raising the level of care by centralizing services are two desirable objectives for the spatial organization of a health service system to improve patient outcomes. In some regions, however, satisfaction of both objectives may not be feasible.

In a study of coronary care services in upstate New York, the trade-offs between geographical accessibility and centralization of services were explored through the application of a location–allocation model designed to locate coronary care units to maximize patient survival (McLafferty & Broe, 1990). The objective function maximizes the difference between two terms. The first term shows the number of coronary care patients surviving after travel to the hospital, which is a function of distance to the hospital. The second term indicates the number of those patients who die in the coronary care unit (CCU), which is a function of the volume of care provided. The difference between the two terms is the total number of patients who survive, an important measure of health outcomes.

The results of the analysis suggested that the number of CCUs in the study region could be reduced and that a system with fewer but better located CCUs could provide better outcomes than the current system that has a greater number of dispersed units. Closing CCUs in small rural hospitals might, however, undermine the viability of those hospitals and result in an adverse effect on health if hospital closure leads to a loss of other services. This research underscores the point that health care facilities are part of a hierarchical system.

MODELING OPTIMAL HIERARCHICAL FACILITY SYSTEMS

Normative modeling techniques that explicitly address the need to determine both the locations of facilities and the levels of care they provide have also been developed. One model addressed the development of a system of facilities in which each facility offered one of three levels of care to serve 150 demand sites in the Suhum District of Ghana (Yasenovskiy & Hodgson, 2007). The assignment of clients to facility sites was enhanced by incorporating spatial interaction modeling formulations so that the attractiveness of the high- and mid-level facilities influencing client choice could be taken into account as well as the disutility of travel. In this approach, clients need not be strictly assigned to the closest facility. A study of patient choice in eastern England found that only 56% of the population had registered with the general practice closest to their home (Haynes, Lovett, & Sunnenberg, 2003). Spatial interaction models, as discussed in Chapter 9, are also used to measure accessibility to health services.

The objective of the hierarchical location–allocation spatial interaction model was to maximize overall client benefit, which depends on the spatial benefit function (Table 10.7). This function depends on the level of the facility (higher-level facilities are more attractive), the distance the client must travel to the facility, and the attractiveness of all potential destinations in the neighborhood of the facility, operationalized as the population size of the neighborhood. The constraints ensure that all demand at all levels is served and that demand at a particular level will be served only by facilities providing an equal or higher level of care. No location can have more than one open facility. The number of facilities that can be opened at each level given the budget is also modeled as a constraint.

The model can be adapted to explore other aspects of facility location. It can be modified to analyze situations where spacing facilities far apart would be desirable. Also, characteristics of the facility neighborhood other than population size can be modeled. There is evidence that the characteristics of the facility neighborhood, as opposed to the residential neighborhood of the client, may be important factors affecting health care accessibility and health outcomes, as discussed in Chapter 9.

As noted, health services delivery most commonly involves travel on the part of either the service provider or the patient. Most location models do not address the specific routes that patients might take in traveling for medical care, although GIS functions have made it easier to model travel based on specific routes. In applying location–allocation models in Ghana, Oppong (1996) demonstrated the impact of road closures during the rainy season on the optimal pattern of facility location.

Finding Optimal Routes for Service Delivery

SHORTEST PATH ANALYSIS

For services like EMS which require service providers to travel to the person in need, the optimal route for the service provider to take from the dispatch site to the location where care will be delivered may also be an important issue. In general, this has been treated more as an operational issue than as a planning issue.

In location analysis, ***shortest path algorithms*** are used to find the shortest distance (or least cost) path from one point in a transportation network to another point. "The computation of shortest paths is an important task in many network and transportation related analyses" (Zhan & Noon, 1998, p. 65). Since the introduction of the problem in the late 1950s (Dijkstra, 1959), the development, testing, and application of shortest path algorithms has been a research focus in geography, transportation, operations research, and management science (Gallo & Pallottino, 1988). Like most location–allocation models, shortest path algorithms assume a network that consists of a set of nodes or points connected by paths or arcs. Each arc begins at one node and ends at another. Each arc also has associated with it a numerical value representing the distance or cost incurred when the arc is traversed. These kinds of networks are easily modeled in a vec-

TABLE 10.7. Mathematical Programming Formulation of a Hierarchical Location–Allocation Problem

Objective function:	Maximize $Z = \Sigma_{i \in I}\Sigma_{j \in J}\Sigma_{c \in C}\Sigma_{k \in K(k \geq c)} W_i U_c B_{ij}^{ck} X_{ij}^{ck}$ where $B_{ij}^{ck} = S_j^k / \exp(\beta^c d_{ij})$
Subject to the constraints:	
All demand at all levels is served.	$\Sigma_{k \in K(k \geq c)}\, \Sigma_{j \in J}\, X_{ij}^{ck} = 1$ for all i, for all c
Demand at a particular level of service is served by a facility at the same or higher level.	$Y_j^k \geq X_{ij}^{ck}$ for all i, for all j, for all k, for all $c \leq k$
A candidate site may have only one facility of any level.	$\Sigma_{k \in K} Y_j^k \leq 1$ for all j
Exactly p_k must be located for each level k.	$\Sigma_{j \in J}\, Y_j^k = p_k$ for all k
X_{ij}^{ck} is 1 if demand at place i for service level c is assigned to a level k facility opened at site j or 0 if demand at place i for service level c is not assigned to that site.	$X_{ij}^{ck} = (0, 1)$ for all i, for all j, for all c, for all k
A level k facility at candidate site j must be either opened or closed.	$Y_j^k = (0, 1)$ for all j, for all k

Where:

Z is the objective function.

I is the set of demand areas, usually nodes of a network, and the subscript i is an index denoting a particular demand area.

J is the set of candidate facility sites, usually nodes of a network, and the subscript j is an index denoting a particular facility site.

C is the set of service levels, and the subscript c is an index denoting a particular service level.

K is the set of facility levels, and the subscript k is an index denoting a particular facility level; in the hierarchy of facility levels, higher level facilities provide all of the services of lower level facilities; $k \geq c$ indicates that a particular demanded service level c is only provided within the range offered at the particular facility level k.

W_i is the number of people at demand site i.

U_c is the portion of total demand for all service levels accounted for by service level c.

B_{ij}^{ck} is the total benefit to patrons from the assignment of patron demand for services at particular levels to facilities providing the appropriate levels of service; S_j^k is the attractiveness of a level k facility at site j; β^c is the distance impedance parameter for service level c; d_{ij} is the distance between demand site i and facility site j.

X_{ij}^{ck} is 1 if the demand for service c at site i is served by level k facility at site j and 0 if it is not.

Y_j^k is 1 if a level k facility is located at site j and 0 if it is not.

p_k is the number of facilities of level k to be located.

tor GIS. As a consequence, shortest path algorithms, more than other types of normative models, have been incorporated into GIS software packages.

ROUTING PROBLEMS

For some types of service delivery, the design of effective systems is more complex than finding the shortest path between two points through a network. Home-delivered care, like the services provided by visiting nurses, requires the service provider to make a series of stops along a route. In rural areas, where travel times between stops are likely to be longer, less service can be delivered by a single provider because more time must be devoted to travel. In this case, the locations of the stops that have to be made are known, and the distances between each pair of stops can be readily determined. Finding the order in which the stops are made that minimizes the total distance traveled is an optimization problem known as the "Traveling Salesman Problem" (Lawler, Lenstra, Rinnooy Kan, & Shmoys, 1987).

This problem is one in an extended set of vehicle routing and scheduling problems that have been formulated for problems with multiple routes and dispatch sites. If a large number of homebound people require care, the home health agency likely has more than one nurse to schedule. This means that the agency will need to identify multiple routes, one for each provider, and assign each person needing care to a particular place on a particular route. Although the locations of the people who need care are fixed, at least in the short term, the agency needs to evaluate the best location for one or more dispatch sites. Variations in the length of time required to make each stop or visit can also be incorporated into these models.

Incorporating Normative Models of Facility Location and Service Delivery into GIS

Normative models of facility location and service delivery have been developed to address a wide range of health services delivery problems over the last several decades (Walsh, Page, & Gesler, 1997). Some software systems include algorithms to solve mathematical programming problems. Several themes related to the integration of normative modeling methods and GIS have been identified (Church & Sorensen, 1996):

- Representing demand for services and the implications of demand aggregation.
- Identifying feasible sites for facility location.
- Modeling coverage areas based on the road network.
- Modeling service delivery routes.
- Finding solution methodologies that can be implemented in the GIS.

Approaches addressing these themes are evident in a variety of GIS applications.

Representing Demand for Services and Demand Aggregation

An emerging role for GIS in health services planning lies in defining and characterizing localities that represent demand areas for health planning purposes. In Delaware, a GIS application mapped counts and rates for census tracts for 10 factors related to community health needs: teen mothers, prenatal care, poverty, employment, public assistance, transportation, home ownership, education, language, and children in single-parent house holds (Berry & Jarrell, 1999). A composite score was calculated, and then the scores were mapped to identify communities with the greatest need for a new service initiative integrating state departments, school districts, and nonprofit organizations in service delivery partnerships. Similar research is being conducted in the United Kingdom to define catchment areas for various health service providers; to generate demographic, social, and residential profiles for patients who use particular providers; and to examine patient travel patterns for other activities like work and school (Hirschfield, Brown, & Bundred, 1995; Bullen, Moon, & Jones, 1996; Niggebrugge, Haynes, Jones, Lovett, & Harvey, 2005).

In the development of service area or neighborhood profiles and the identification of demand sites for services, the people requiring services are usually grouped together by residence. Distances to care are usually not calculated from individual residential locations to service centers. Instead, demand or need is aggregated to a set of area centroids or other central points. For example, the number of children requiring immunization might be aggregated to census tract areas, or the number of motor vehicle collisions requiring an emergency response might be aggregated to the nearest intersection. Demand aggregation reduces the complexity of location and routing problems, but it has some important implications for location modeling.

When demand is aggregated (Current & Schilling, 1990), the true distance to accomplish health service delivery to an individual is replaced by the distance from the point of aggregation (Figure 10.6). In some cases, the true distance exceeds the modeled distance; in other cases, the true distance will fall short of the modeled distance. For models like the p-median problem, this will result in under- or overestimation of the true travel distance or cost, and the modeled optimal facility location pattern may not be optimal in fact. For covering problems, demand aggregation may result in an under- or overestimation of coverage. If the location of a person needing care is translated to an aggregate demand site that is closer to a proposed facility than the person's residence, the person's residence may lie outside the critical service distance or time even when the aggregate demand site can be served within the critical service distance or time.

Given the ability of GIS to manage large volumes of spatial data, one solution to this problem might be more disaggregate representations of demand. This

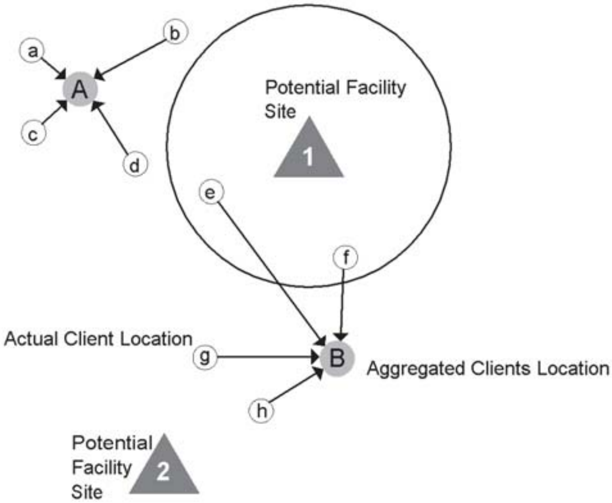


FIGURE 10.6. A schematic example of demand aggregation shows clients a, b, c, and d aggregated from their actual location to location A and clients e, f, g, and h aggregated from their actual locations to location B. The true travel distance for client h to Potential Facility Site 1 would likely be underestimated by this aggregation because location B is closer to Facility Site 1 than client h is. The true travel distance for client e would likely be overestimated as a result of this aggregation because location B is farther from Facility Site 1 than client e is. The true coverage would also be misrepresented by this aggregation on the basis of the critical coverage distance radius shown for Potential Facility Site 1. Clients e and f are actually within the critical distance from Facility Site 1, but this would not be apparent if they were aggregated to location B.

would increase the number of demand sites as well as the computational effort required to solve most location–allocation problems. Analysts have also investigated the use of dasymetric mapping techniques, discussed in Chapter 6, to improve the representation of population distribution and modeling of demand locations (Langford, Higgs, Radcliffe, & White, 2008). Alternatively, GIS can be used to assist health service analysts make intelligent choices in aggregating demand. Analysts who have studied the problems of demand aggregation suggest several strategies for reducing its effects (Daskin, Haghani, Khanal, & Malandraki, 1989; Current & Schilling, 1990). First, demand should only be aggregated to places where some demand is actually present (Figure 10.7). Second, demand should only be aggregated to a location if the demand would be covered by a facility located at the aggregate site. Finally, analysts might wish to aggregate only those individuals covered by the same set of potential service sites. The literature addressing demand aggregation issues for a range of location problems is growing, but there continues to be a lack of consensus on the most appropriate methods for measuring error resulting from demand aggrega-

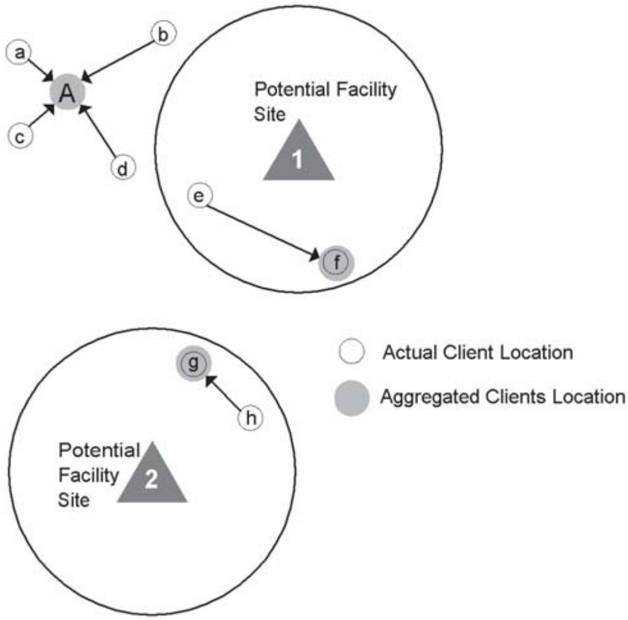


FIGURE 10.7. Error resulting from demand aggregation can be reduced by aggregating demand to locations where some clients are actually located and by aggregating demand only if the distance between the actual client location and the aggregated location is less than the critical distance specified in a covering problem.

tion and the most useful methods for dealing with these errors (Francis, Lowe, Rayco, & Tamir, 2009).

Identifying Feasible Sites for Potential Facility Location

In addition to describing and analyzing the demand for health services, GIS analysis also has a role to play in identifying candidate facility locations. Most network models of facility location work with a set of potential facility sites from which the facilities to be opened are selected. In the traditional formulation of location–allocation models, candidate facility sites are usually nodes in the network space; they are included because they are also demand sites or because of their relative location to demand sites. Site characteristics affecting the feasibility of actually constructing a facility at the candidate supply site are generally ignored.

GIS, through its ability to integrate data layers spatially, provides an opportunity to take both location and site characteristics into account in identifying candidate facility locations. A team of planners from Maryland used a GIS to identify and rank sites for new primary medical care facilities (Marks, Thrall, & Arno, 1992). The GIS application included data layers describing parcel size, distance to facilities and demand centers, percent of local population older than 65 years, existing land use on the site, site availability, percent slope on the site, and availability of infrastructure like water and sewer systems.

Modeling Coverage Areas Based on Road Networks

Many early location–allocation models implemented outside a GIS environment required distances between demand and supply sites as direct data input. These distances might frequently be measured as straight-line Euclidean distances between sites. GIS provides an opportunity to improve measurement of coverage areas based on travel over actual street networks. Network analysis functions in the GIS make it possible to identify a node in the street network as a starting point and to identify the portions of street network segments within a specified travel distance from the starting point (Figure 10.8).

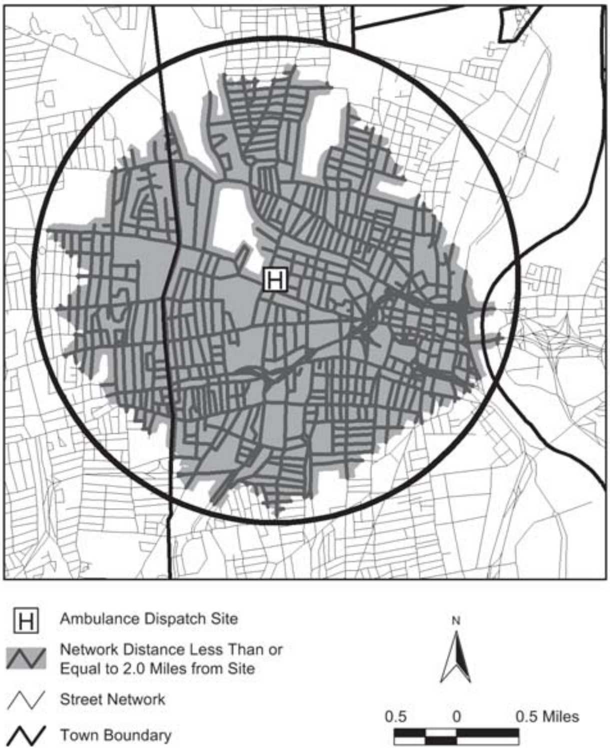


FIGURE 10.8. GIS network functions identify street network segments within a specified travel distance or travel time from a starting point. For comparison, a 2.0-mile circular buffer around the ambulance dispatch site is also shown. The area covered based on network distance is smaller than the area covered within the circular buffer based on Euclidean distance, as discussed in Chapter 9.

A standard of 30-minutes travel time to a primary care physician or a standard of 8 minutes for EMS response, for example, could be effectively modeled in a GIS using these techniques, leading to more accurate descriptions of underserved areas. Mapping coverage areas would indicate regions and populations in the larger study region that probably cannot be served at the minimum standard with the existing arrangement of resources. Utilization data showing longer actual travel or response times than expected based on the modeled coverage area would indicate neighborhoods where barriers to service delivery exist (Peters & Hall, 1999).

Accurate modeling of road travel time response areas using GIS can be used as a basis for site selection. In a study designed to identify which of two tertiary care hospitals in British Columbia would be the best choice for expanding a helicopter emergency services program (Schuurman, Bell, L'Heureux, & Hameed, 2009), network analysis was used to model one-hour service areas for each facility based on road travel. Once these areas were identified, population and historical usage within the modeled catchment areas suggested that the Royal Inland Hospital would be the best site for an expanded helicopter EMS.

A different approach to modeling travel time for site selection was developed to locate new primary health care facilities in KwaZulu-Natal, South Africa (Tanser, 2006). This model used a 30 meters × 30 meters grid of the study area and created cost surfaces based on the least travel time from every cell to the most accessible target clinic site based on walking and public transport. These surfaces incorporate spatial impedance, the difficulty people have traveling from a given origin to a given destination, into the model. By mapping total person impedance by km^2, the largest contiguous area of high spatial impedance was identified. This area would be an attractive site for locating a new clinic.

Modeling Service Delivery Routes

GIS applications have also been useful for evaluating optimal routes for services delivered to those in need. The network analysis functions of GIS can be used to display stops and identify optimal routes. Analysts can specify whether or not the route must begin or end at a particular site and model the impact of barriers (Figure 10.9).

GIS routing functions were used to evaluate the optimality of delivery routes in a meals-on-wheels program in southeastern Connecticut (Wong & Meyer, 1993). Five vans were used to deliver meals, with each van covering a particular route after it departed from the main kitchen. A single-depot vehicle routing problem with time windows to account for meal delivery time was solved using the GIS. The application required a street network database. Residential locations of delivery stops were geocoded using the GIS and assigned to nodes on the street network. The routing procedure was used to allocate clients to five routes to minimize the time required to serve all clients, an important consideration in home-delivered meal service. Four of the five routes described in the results of the GIS analysis differed from the five actual routes.

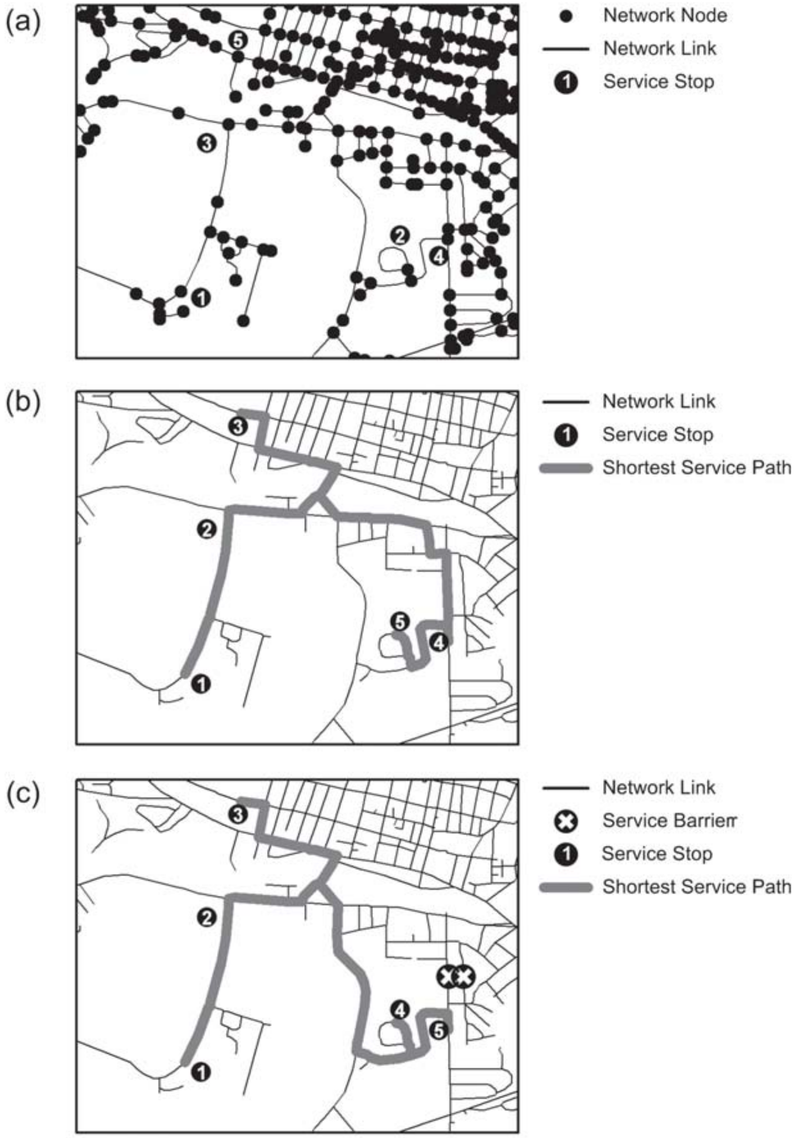


FIGURE 10.9. Using GIS to model service delivery routes. Figure 10.9a shows a set of five service delivery stops including the dispatch location at Stop 1, mapped on a network of links and nodes. Stops 2, 3, 4, and 5 are numbered in the order in which they were entered into a database of clients needing the service. Figure 10.9b shows the optimal route to serve the stops if the service vehicle must depart from Stop 1. Note that the stop numbers have changed to correspond to the optimal order of service stops on the shortest path. Figure 10.9c shows how the shortest path would change if an event like a flood made two road segments impassable. The optimal route shifts and the order of stops also changes for the last two stops.

Solution Methodologies

Exact solutions to optimization problems can be found through linear programming or integer programming algorithms. These algorithms guarantee an optimal solution if one can be found. These methods, however, are computationally intensive, especially for real-world problems like a p-median problem involving many demand and supply sites. Integer and mixed-integer programming solutions are more computationally complex than linear programming solutions, even for problems of the same size. Hardware and software developments have made it possible to more easily solve large problems, like the hierarchical location–allocation model with 150 nodes and three levels described in a previous section of this chapter (Yasenovskiy & Hodgson, 2007).

The alternative to mathematical programming solutions is heuristics. A ***heuristic*** is an algorithm that finds an approximate solution to an optimization problem in a reasonable amount of computation time. The solution is not guaranteed to be optimal, but it may indeed be optimal. Benchmarking studies have evaluated the most commonly used heuristics for solving complex problems. Research investigating the most appropriate heuristics for integration into GIS is ongoing.

Although most GIS packages do not yet include functions to determine locations for new facilities or to select facilities for closure, data on distances between demand and supply sites can be exported from a GIS and imported into spreadsheets where simple calculations can be performed to compare the average travel distances of clients to services sites so that the facility patterns that minimize travel distances can be identified. Rushton (1999) illustrates this approach for closing some existing facilities and reopening them at other locations, for locating facilities in an unserved area, and for evaluating changing patterns of utilization associated with changes in the location and number of clinics.

The mathematical structure of shortest path problems makes solution by exact methods more practical. As a result, shortest path algorithms are commonly found in many GIS software packages. Many of these algorithms, however, were developed using hypothetical street networks. The development of GIS provides an opportunity to test them using real road networks. Real street networks differ from many of the hypothetical networks used in testing shortest path algorithms because road density is variable in real-world networks. Density is higher in urban centers that are in turn surrounded by suburban areas with distinct subnetworks, further surrounded by sparse rural roads. A test of 15 shortest path algorithms on two real road networks modeled in a GIS indicated that different algorithms would be preferred for different kinds of problems (Zhan & Noon, 1998).

Clearly, the effective integration of normative models of facility location and routing into a GIS would give the analyst the opportunity to select from a variety of problems and solutional techniques. Much of the literature emphasizes the value of altering estimates of demand, the set of candidate supply sites, critical service distances, and solutional techniques to evaluate alternatives (Birkin, Clarke, Clarke, & Wilson, 1996). Increasingly, GIS for health services delivery are being viewed as spatial decision support systems.

Spatial Decision Support Systems

The concept of the ***spatial decision support system*** (SDSS) grew out of the ***decision support system*** (DSS) first championed by Geoffrion (1983). SDSS has the following in common with DSS: (1) an explicit approach to solving ill-defined problems, (2) easy-to-use interfaces, (3) flexibility in combining analytical models with data, (4) ability to evaluate alternatives, (5) models reflecting a variety of decision-making styles, and (6) support for interactive and recursive problem solving (Malczewski, 1999). Densham (1991) identified some distinguishing features of the SDSS, including (1) support for spatial data input, (2) representation of spatial relations and structures, (3) availability of spatial analytic techniques, and (4) support for output in a variety of forms, including maps. The components of an SDSS include a geographic database, a set of models, a database management system (DBMS), and a user interface (Figure 10.10). The DBMS manages spatial and attribute data and provides the GIS with capabilities for data input, storage, retrieval, and manipulation. The user interface can provide access to data and tools through the web (Schuurman, Leight, & Berube, 2008). An SDSS called DOCLOC was designed to assist health practitioners make decisions about practice locations in Idaho (Jankowski & Ewart, 1996).

A DSS approach was used to identify routes for high-level radioactive waste shipments. The system was designed to address the multiple facets of route choice

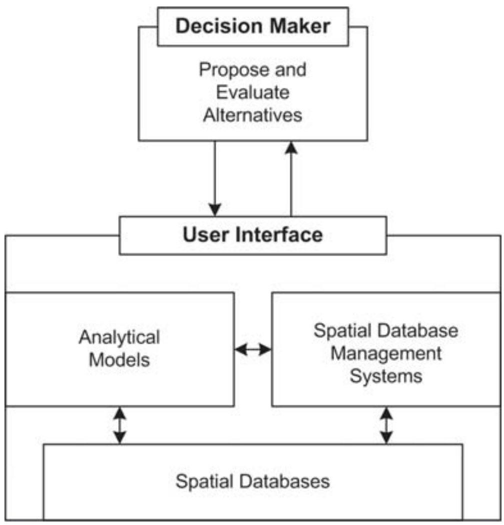


FIGURE 10.10. The components of a spatial decision support system.

and emergency response team siting comprising the waste shipment problem (List & Turnquist, 1998). Each of these facets had multiple objectives associated with it. For route choice, the objectives included finding the routes that would minimize population at risk, probability of an accident, and delays. For emergency response team siting, the objectives included minimizing distance to road links with the highest shipment volumes and minimizing the maximum distance a team would have to travel to reach any possible accident location. The siting of the response teams depends, in part, on the choice of routes. The system was tested using the anticipated shipments of waste from around the country to a Waste Isolation Pilot Project site in southeastern New Mexico. Health services delivery needs in response to disasters are increasingly modeled using GIS.

Health Services Delivery in Response to Disasters

The public health community recognizes the important role of public health in managing emergencies and disasters (Landesman, 2005). Emergency management encompasses a wide range of activities. GIS are recognized by response agencies at all levels as playing a critical role in data management. In the context of disaster and emergency services, an ***emergency*** is an event or a series of events that endanger or adversely affect people, the environment, or property. A ***disaster*** is an emergency of such scope that the management capability of local resources is exceeded. Disasters usually result in great damage, destruction, or loss. As discussed in Chapter 6, emergencies may be natural in origin, human in origin (planned or unplanned), or mixed. Emergencies of human origin cover a wide range of events—everything from chemical spills to utility failures to airplane crashes, from riots to warfare to terrorism.

Emergency management activities take place in five phases. ***Planning*** includes analyzing and documenting the possibility of an emergency and its consequences. Modeling areas vulnerable to flooding and assessing the potential impacts on people, their residences and other properties, human service facilities, and the environment is a planning activity. ***Mitigation*** involves taking action to eliminate or reduce the probability of a disaster, including long-term activities designed to reduce the damaging effects of unavoidable disasters. ***Preparedness*** includes developing plans to save lives and minimize damage in the event that mitigation measures do not or cannot prevent disasters. This includes compiling resource inventories, conducting drills, installing warning systems, developing and testing evacuation plans, and training responders. Activities may also include measures to enhance response, for example, stockpiling supplies.

Even when effective planning, mitigation, and preparedness efforts have been made, emergencies and disasters will occur. The response and recovery phases complete the integrated cycle of disaster and emergency management. ***Response*** activities following an emergency or a disaster provide immediate assistance to victims, stabilize the situation, reduce the probability of additional damage (for example, patrols to prevent individuals from entering unsafe structures or areas), and speed recovery (for example, damage assessment). ***Recovery*** activities include activities necessary to return all systems to normal or to a state that improves upon predisaster conditions. Short-term recovery includes activities such as cleanup, whereas long-term recovery includes services such as redevelopment loans, community planning, and legal assistance.

Activities in all phases of the disaster and emergency management process require data. Geospatial data held at the local level are especially important during the response and recovery phases (Napier, 2003). Because disasters, by definition, overwhelm the capabilities of local responders, people from outside the local community are involved in response and recovery. They may not be as knowledgeable as local residents about the community affected, so geospatial data and flexible systems for making data available in the field are needed.

GIS applications have been developed to serve information needs in every phase of the disaster and emergency services management process. GIS applications in planning have identified both vulnerable environments and vulnerable communities of people. At the global scale, research has used GIS to investigate urban settlement in low elevation coastal zones. These settlements are growing and are especially vulnerable to a range of environmental risks (McGranahan, Balk, & Anderson, 2007). Research is increasingly recognizing the importance of integrating information on the built environment with information on environmental risks. International teams have partnered in using GIS technology to investigate earthquake hazards and the structural vulnerability of building stock in historical areas in San Giuliano di Puglia, Italy, and in Valparaiso, Chile (Indirli, 2009). In the context of risk from wildfire, the wildland–urban interface has been the focus of vulnerability assessments. The spatial organization of dwellings—isolated, scattered, or clustered—has a major impact on fire occurrence (Lampin-Maillet, Jappiot, Long, Morge, & Ferrier, 2009).

Equally important is research documenting the vulnerability of particular populations, showing the relationship between increased hazard risk and socioeconomic characteristics of communities (Morrow, 1999). In many communities, limited economic and material resources limit people's ability to mitigate, prepare for, respond to, and recover from disasters. A substantial body of research has documented significant inequalities among social groups with respect to exposure to natural disasters (Fielding, 2007; Neumayer & Plumper, 2007). ***Community vulnerability maps*** show where at-risk groups, including the elderly, female-headed households, homeless, renters, and others, are concentrated. A Social Vulnerability Index based on county-level socioeconomic and demographic data for the United States found the most vulnerable counties clustered in metropolitan counties in the East, south Texas, and the Mississippi Delta region (Cutter, Boruff, & Shirley, 2003). Temporal and spatial changes in patterns of vulnerability in the United States over several decades have also been observed (Cutter & Finch, 2008).

GIS-based community vulnerability analyses have also been conducted at the local scale. In a study of areas within Hillsborough County, Florida, where Tampa and St. Petersburg are located, social vulnerability was assessed in the context of geophysical risk of hurricanes and flooding (Chakraborty, Tobin, & Montz, 2005). The number of occupied housing units without telephones or vehicles and the size of the institutionalized population and the population over 5 years of age with disabilities were mapped at the block group level to identify areas where it might be difficult to evacuate people. Research conducted in Georgetown County, South Carolina, assessed the vulnerability of places to 12 environmental threats; described the population and social characteristics of places with different levels of vulnerability to hazards; and mapped place vulnerability against emergency response lifelines and infrastructure (Figure 10.11) (Cutter, Mitchell, & Scott, 2000). The degree of vulnerability to hazards depends on the severity of biophysical risk, the size and characteristics of the population, and the level of resources available for response in particular places.

A study of the population of New Orleans in the aftermath of Hurricane Katrina extended the concept of vulnerability mapping to include all phases of emergency management (Curtis, Mills, & Leitner, 2007). Low-income and minority residents were not only more likely to live in areas likely to be affected by flooding, but they were also less likely to have been evacuated, so they experienced the stress of the storm directly. They also lived in neighborhoods that had the worst prospects for recovery after the disaster.

GIS was used in an investigation of the effectiveness of a grassed waterway developed to mitigate a muddy flood hazard affecting villages in the Belgian loam belt (Evrard, Persoons, Vandaele, & van Wesemael, 2007). Muddy floods occur when runoff from agricultural catchments after intense rainfall accumulates in dry valleys downstream, causing damage to housing and infrastructure. After an extreme flooding event in 2002, a grassed waterway and retention dam were built as mitigation measures. A hydrological model used with a GIS simulates discharge at various points in the catchment and makes it possible to model the effectiveness of the mitigation effort in preventing another extreme event of the same magnitude as transportation and land use patterns in the area continue to change.

Public health professionals have responded to recent natural and technological disasters by providing tools for public health preparedness. The Agency for Healthcare Research and Quality in the United States has developed the Emergency Preparedness Resource Inventory (EPRI) as a tool for local, regional, and state planners (Agency for Healthcare Research and Quality, 2005). EPRI is a downloadable tool enabling local or regional planners to assemble an inventory of critical resources that would be useful in responding to bioterrorism or other public health emergencies.

The ability of maps to identify a community's vulnerability to disaster and available community resources was evaluated in a study conducted in New York (Zarcadoolas, Boyer, Krishnaswami, & Rothenberg (2007). Researchers conducted interviews with 178 English- and Spanish-speaking residents of east and central Harlem and presented participants with a map used in the City's Office of Emergency Management Storm Surge Report. A majority of adults who had not completed high school were not able to read the map to identify whether they

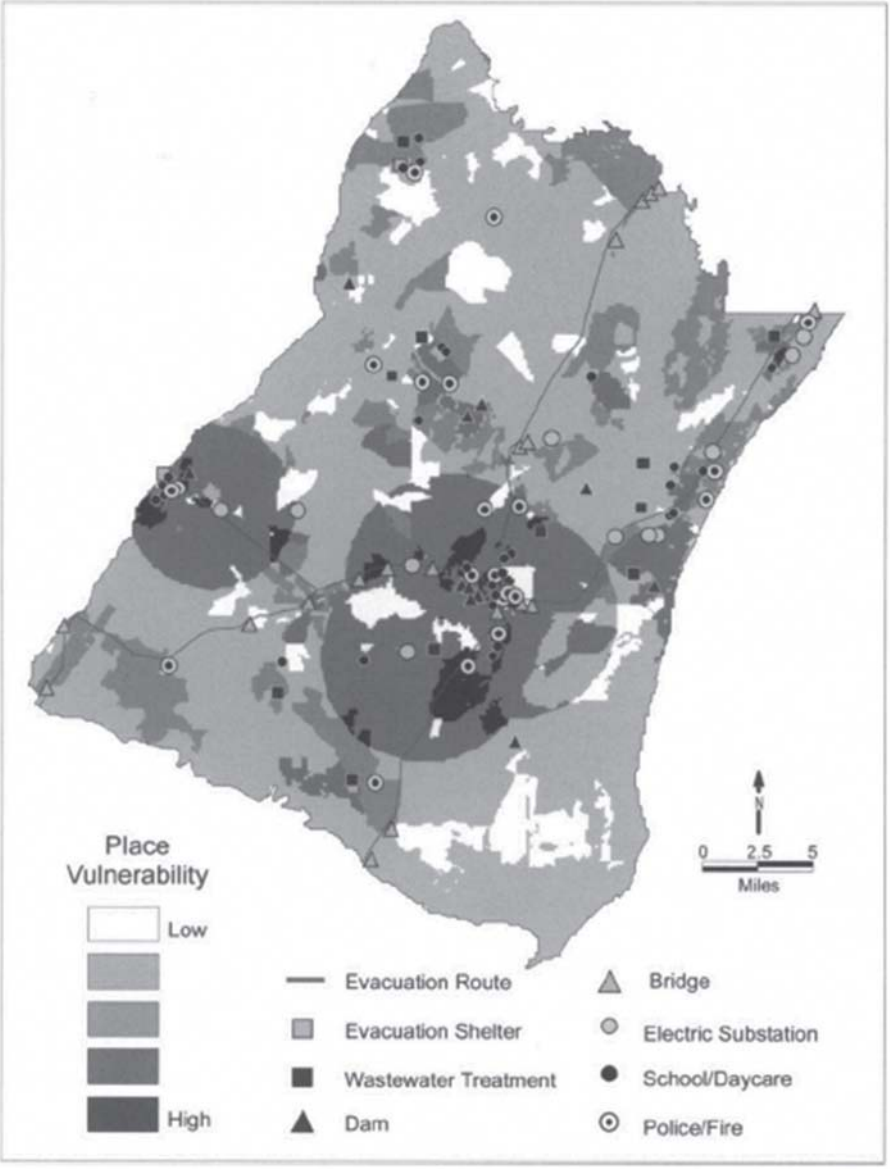


FIGURE 10.11. Mapping place vulnerability based on the spatial distribution of overall hazard scores shows that many important resources for emergency response are located in highly vulnerable areas. From Cutter, Mitchell, and Scott (2000). Copyright 2000 by the Association of American Geographers. Reprinted by permission of the Taylor and Francis Group.

lived in a hurricane evacuation zone or to identify the evacuation center that was nearest to their home. This research raises important questions about how best to communicate with the public for emergency preparedness.

Evacuation planning is an emerging focus of research on disasters. Studies have investigated ways to identify neighborhoods facing transportation difficulties during evacuation (Cova & Church, 1997) and transportation networks that might lead to significant problems in evacuation (Church & Cova, 2000). For example, a traffic simulation study for a canyon community near Salt Lake City, Utah, which was prone to wildfires investigated subneighborhood variation in household evacuation travel times. GIS was used to map the effects on evacuation times of adding a second access road to the community (Cova & Johnson, 2002). Traffic signal timing on urban streets has been investigated as a factor affecting evacuation (Chen, Chen, & Miller-Hooks, 2007).

Even with the best of planning and preparation, disasters will occur. Public health response following a disaster such as a tropical storm is one important part of the total response to disasters. GIS have been used to support rapid epidemiological assessments following weather-related disasters. In Texas, following landfall of a tropical storm near Galveston in 2001, the most severe flood-related damage observed to date was recorded in the Houston metropolitan area (Waring et al., 2005). The GIS application was used to facilitate a modified cluster sample of households in the areas most affected by flooding. A total of 420 households participated in the survey, which was accomplished within one week following the storm. A sizeable number of households reported serious damage to their homes, and more than a third of these were outside the 500-year floodplain zone. Significant increases in illness were noted in residents of flooded homes compared to those in homes that were not flooded.

Health information collected in the aftermath of disasters is important not just for response to immediate needs; these data also help us to understand the health consequences of disasters. A study of fatal and inpatient injuries due to the 1994 Northridge Earthquake in California used GIS to map all injury locations (Peek-Asa, Ramirez, Shoaf, Seligson, & Kraus, 2000). Injuries were studied in relation to the distance from the earthquake's epicenter, the intensity of the earthquake, peak ground acceleration, and damaged residential buildings. Seismic hazard and building damage did not completely predict injury incidence or severity. Injuries caused by falling parts of buildings were, on average, closer to the epicenter, but injuries caused by falls and cutting or piercing were spread over a larger geographic area. This research, like the research on flooded residences, suggests that GIS can help provide a better picture of the spatial extent of health problems following a disaster.

Maps and data have been used effectively in the recovery phase by the Greater New Orleans Community Data Center (Nonprofit Knowledge Works, 2011). The Center was in existence and providing information to the community before Hurricane Katrina came ashore in August 2005. The site's server and data and several key staff were located outside of the community in different states, and so the site remained in operation during the disaster and early recovery period. Because so many other organizations and institutions were affected by the hurricane, the Community Data Center site was one of the few sources of information documenting the state of the city before the disaster occurred.

The Center became a key source of information on the recovery of New Orleans. Innovative approaches to mapping included working with the postal service to map postal deliveries as a way of helping service providers to visualize which areas of the city were experiencing the return of population. Thanks to data compiled by the Center, the city of New Orleans was able to challenge the Census Bureau's initial estimate of the city's population. The Census Bureau's decision to revise its estimate upward by almost 50,000 people resulted in an increase of federal funds to the city amounting to $45 million.

Conclusion

GIS are reviving interest in the application of normative modeling techniques in services delivery planning, operation, and evaluation. These techniques have sometimes been criticized because they do not attempt to describe observed patterns of behavior, but there are also drawbacks to relying solely on utilization data or manifest patterns of travel in planning health services. First, utilization patterns for existing services may be of little use in planning for new services because the new sites will alter the geographical set of opportunities for receiving care. If the service becomes more accessible through the addition or relocation of resources, then the true cost of the service will decrease and individuals will be able to utilize more of the service without increasing total out-of-pocket expenditures. Second, for some health services, like EMS, geographic areas must be covered whether or not there has been high utilization. Finally, given that utilization data are not always available, normative methods make it possible to study the pattern of health services that would result if particular objectives were desired. Even if conflicts in the multiple objectives that decision makers want to achieve make it difficult to implement optimal solutions in practice, the results of location–allocation models and the output from SDSS can establish quantitative benchmarks for what it is possible to achieve in designing a service delivery system.

In developing meaningful analyses of health services delivery systems, it is important to recognize that the location, hours of operation, and scale of service are not simply matters of convenience or economics. They truly affect the quality of care provided and the medical outcomes that result. In addition to the studies described in this chapter on the impact of distance on whether care can actually be provided, the research reviewed in Chapter 9 on the links between distance and utilization indicates that the geography of health services influences health status. Numerous studies have shown that the volume of care provided is related to the quality of the care and the pattern of health outcomes.

More research is needed on the relationships between the timing of care and health outcomes. The medical research literature is showing an increasing interest in the relationships between biological clocks and human health (Hrushesky, 1985). Strokes and heart attacks are more likely to occur during the morning hours. What are the implications for EMS provision? The effectiveness of treatments like chemotherapy may be a function of when they are administered. Are health services designed so that therapy can be provided at the optimal time and place?

The consequences of natural and technological disasters in vulnerable communities and for health service delivery systems themselves are of increasing concern in all countries. The aging of the world population and the redistribution of population into urban environments are reshaping patterns of vulnerability. GIS can support health services planning efforts to make sure that quality health services are available when and where we need them.

CHAPTER 11

Health Disparities

Variations in the health of populations are apparent at the global, regional, and local scales, and these patterns have been observed for centuries (Vallin & Meslé, 2004; Gibbons, 2005). As the preceding chapters show, spatial variations in the distributions of disease agents and environmental toxicants explain some of the differences in health from place to place. In the decades after World War II, however, ***health disparities***—differences in health status among social groups—came to be seen as a public health issue that was as important as the overall level of health in a population. In countries where the life span was increasing and the overall incidence of specific diseases was declining, improvements in health did not always occur equally among all groups (Figure 11.1). Recognition of these differences in health status by class, race or ethnicity, gender, and geography has fostered the study of health disparities, especially in countries with high per capita national incomes, although inequalities in health also exist in less affluent countries (Braveman & Tarimo, 2002). In the last several decades, observed declines in health status even in wealthier countries and the emergence of new health problems like HIV/AIDS have also affected different groups disproportionately. The size of these inequalities has led some to rank health disparities as the most important public health problem (Graham, Boyle, Curtis, & Moore, 2004).

This chapter considers the role of geography in the study of health disparities in terms of theory and methodology. In the first section, we consider the various levels at which health status and factors influencing health are measured, and we discuss the concepts of contextual and compositional factors. The next section reviews how GIS have been used to visualize differences in income and environmental conditions across a range of spatial contexts. Because poverty has been identified as an important individual and contextual variable in health disparities research (Krieger, Chen, Waterman, Rehkopf, & Subramanian, 2003), the geography of income in relation to health is discussed. Environmental conditions also contribute to health disparities. In this section, we pay particular attention to how GIS have been used to measure characteristics of the built

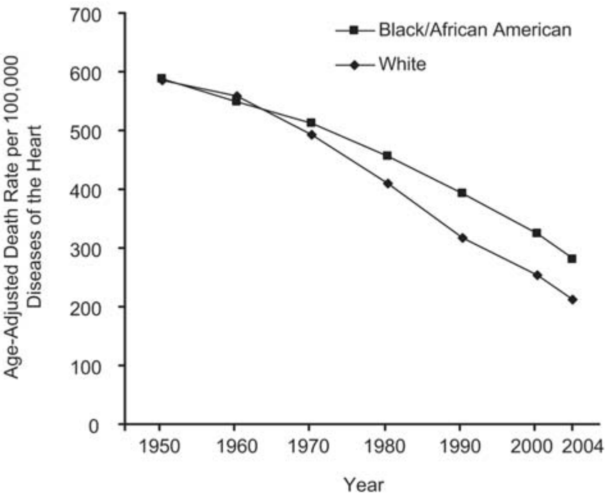


FIGURE 11.1. Age-adjusted death rates for diseases of the heart declined for blacks/African Americans and whites in the United States from 1950 through 2004. The rate declined more rapidly for whites than for blacks, resulting in a disparity. Individuals in both groups may be of Hispanic origin. Rates for Asians/Pacific Islanders and Native Americans are not graphed because they are known to be underestimated and were not reported for all years. Data from National Center for Health Statistics (2006b).

environment that may influence health and review some of the evidence on the relationships between these characteristics and health.

The third section reviews the role GIS can play in defining neighborhoods. How the areas or neighborhoods used to represent the contexts for individual health are defined is an important conceptual and methodological issue. In some cases, data are reported for political and administrative units like census tracts, and GIS functions are used to place individuals in these neighborhoods. In other cases, GIS functions are used to create neighborhoods based on individual residential, school, or workplace locations. It is also possible to use GIS to map the area that a person considers to be within his or her neighborhood.

GIS plays a role in implementing analytic methods used to investigate health disparities, including multilevel modeling. In many studies, area-level variables are used to investigate health disparities. Spatial statistical models help analysts to account for spatial dependencies across neighborhoods and to explore spatial variability in the relationships between neighborhood conditions and health outcomes. Finally, the role of location processes such as migration in creating and maintaining the socioeconomic and neighborhood inequalities that give rise to health disparities is discussed. These processes explain how regions come to differ in terms of the resources and risks that affect their populations and in the composition of their populations. In efforts to understand the relationships between people, where they live, and their health, geographical analysis is helping to provide answers to the question "Does where you live matter to your health?"

Context and Composition

Health problems can be studied at multiple levels. Lower level units, sometimes called ***micro units***, are nested within units at a higher level, ***macro units***. In an analysis of children in schools, for example, individual participants in the study would be the lower level units, and the studies themselves would be the higher level units. If the analysis involved measurements made for each participant over three points in time, the measurements would be lower level units nested within individuals as the next higher level, and individuals would be nested in schools at the highest level. In many studies of health disparities, macro units correspond to geographical areas or neighborhoods.

The idea that characteristics of individuals in a population cannot alone explain the distribution of health problems within the population is a key premise of social epidemiology (Krieger, 2002). ***Ecological studies*** of health compare large groups of people rather than individuals. Although the social epidemiologic perspective has been controversial within the discipline of epidemiology itself (Krieger, 1994, 2001; Krieger & Zierler, 1996, 1997; Kaufman, 2001; Macdonald 2001; McPherson, 2001; Savitz, 1997; Siegrist, 2001; Zielhus & Kiemeney, 2001), the connections between poverty, social and occupational class, housing and living conditions, and health have been widely studied in the social sciences and addressed in public health policy. Taking the milieu of the person into account in studies of health involves using analytic methods that extend models based on individual characteristics to include variables measuring family, social group, neighborhood, and regional contexts and the composition of the populations in these settings.

Contextual effects are differences in an outcome observed at a lower level that can be attributed to the effects of variables observed at a higher level after controlling for individual-level confounders (Diez Roux, 2003a). If mean income observed at the neighborhood level has an effect on an individual level outcome such as diabetes after controlling for individual income, this is considered a contextual effect. ***Compositional effects*** are differences in an outcome—for example, the diabetes rate—that can be attributed to the characteristics of the individuals comprising the different family, neighborhood, or regional groups or contexts rather than to the nature of the setting (Duncan, Jones, & Moon, 1998).

For geographers, these higher levels have spatial dimensions, and the methods of geography provide guidance in defining the spatial aspects of a particular context. How contexts are defined is important because, in addition to heterogeneity among individuals, contextual analysis emphasizes heterogeneity among contexts (Duncan, Jones, & Moon, 1998). ***Heterogeneity*** means that individuals or contexts are diverse in kind or nature. Exploring heterogeneity among individuals, contexts, and individuals within contexts is also important because there may be different degrees of individual heterogeneity within contexts (Bullen, Jones, & Duncan, 1997). Two communities may have the same mean income, but in one of the communities everyone has roughly the same as the mean income and in the other community there is high variance in individual incomes.

Conceptualizing what is an attribute of the individual and what is an attribute of the context is not always easy in contextual analysis. Individual-level variables like age or income are measures that characterize individuals. Group-level variables are measures that describe groups. This distinction is blurred when a group-level variable is used as a proxy for an individual-level measure because data at the individual level are not reported or are not reliable (Bassett & Krieger, 1986). Research has shown that area-level measures may not be valid as substitutes for individual-level data (Hanley & Morgan, 2008). In multilevel analysis, group-level variables are intended to measure group-level constructs.

Different types of group-level or contextual variables have been described (Diez Roux, 2003b). ***Derived group-level variables*** mathematically summarize characteristics of individuals in the group. Mean income is an example. It is calculated from the incomes of individuals in the group. ***Integral group-level variables*** also measure characteristics of the group, but they are not derived from the characteristics of individuals in the group, and they have no individual-level analogues. Variables measuring laws governing tobacco use in the contextual setting or population density to classify the context as urban or rural are examples. Characteristics of social interactions among members of the group are also considered integral group-level variables and sometimes called ***structural group-level variables*** (Diez Roux, 2004). The term "environmental variables" has been used by some to refer to area-level characterizations of physical and chemical conditions that are used exclusively as proxies for individual-level exposure to environmental conditions (Diez Roux, 2003b). These variables have traditionally been derived by sampling conditions at particular places and modeling surfaces of environmental quality, as described in Chapter 6. ***Environmental variables*** are not mathematically derived from measures of individual people but from measures of a sample of individual places. Increasingly, however, technology is being used to assess environmental quality in the places where an individual person actually circulates. Thus, it is possible that an individual's exposure to air pollution could be different from the characterization of the level of air pollution in the general area where the person lives.

Explaining the role of contextual and compositional effects in the health of populations has been a challenge in research on health disparities. A measure like the age structure of the population in a place might be conceived of as a compositional variable or a contextual variable (Cagney, 2006). Rather than focusing on rigid distinctions between contextual and compositional effects, research is increasingly concerned with the relationships between people and where they live (Macintyre, Ellaway, & Cummins, 2002; Cummins, Curtis, Diez-Roux, & Macintyre, 2007). The idea that where we live affects our well-being has face validity. Our sense of place consists, in part, of our assessments of the connections between place and well-being, about whether particular dwellings and neighborhoods are supportive and safe. Individuals and groups are not indifferent to neighborhood changes—either the movements of residents or the land use changes that occur when businesses and institutions relocate. How place contributes to health and well-being, how to uncover these effects, and how to assess whether these effects vary from place to place over time are questions that have engaged health analysts in many countries. Whether there is something about the setting itself that affects the health of individuals and social groups or whether there is something in the characteristics of the people who comprise the community of interest that explains their health or both, place differences are a central element in health disparities research.

Visualizing and Measuring Area Characteristics

GIS provide a useful tool for visualizing area differences in income and measuring and visualizing area characteristics, particularly characteristics of the built environment. An important step in studying health disparities is explicitly investigating spatial patterns of community characteristics by mapping them and, as discussed later in the chapter, by assessing the degree of spatial dependency using statistical methods. Variations in income and other area characteristics are apparent at a variety of spatial scales. It is not just the degree of variability that is a concern; the spatial arrangement of areas with high and low incomes or high and low environmental quality is also important.

The Geography of Socioeconomic Inequality

Income is one of the most widely used variables in health disparities research in the United States, as an individual-level variable and as a contextual variable explaining mortality and differences in self-reported health (Pickett & Pearl, 2001). Income contributes to well-being because it provides the means to command resources, including housing, food, education, and medical care. Sometimes, the level of resources at the individual's disposal is measured in terms of poverty (Krieger et al., 2003), wealth (Duncan, Daly, McDonough, & Williams, 2002), or occupation and education rather than income. The literature generally supports the conclusion that individual or household income is inversely related to poor health in a nonlinear way and that the income level of the local area also affects the health of individuals (Jones, Duncan, & Twigg, 2004).

Variations in income are apparent at the global, regional, and local scales (Figure 11.2). Furthermore, there are spatial patterns in income at every scale. These patterns can be visualized using GIS and analyzed using spatial statistical methods.

Income inequality within places has also attracted attention as a factor in population health (Kawachi & Kennedy, 1999; Subramanian & Kawachi, 2003; Jones, Duncan, & Twigg, 2004; Ross, Wolfson, Berthelot, & Dunn, 2004). Inequality in income and inequality in wealth have increased substantially in the

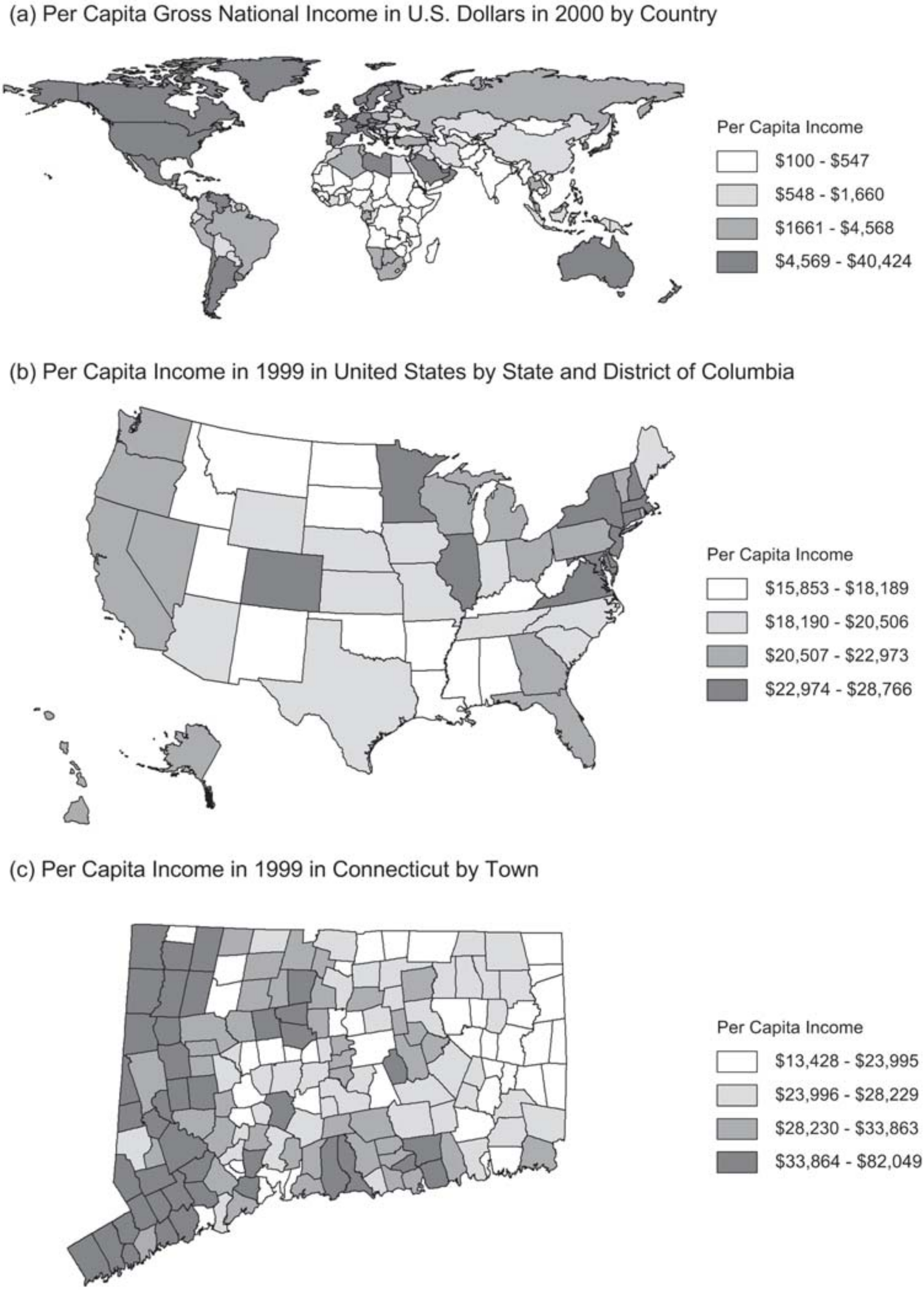


FIGURE 11.2. Variations in income at the global, regional, and local scales.

United States over the last 30 years (Subramanian & Kawachi, 2004). Research to date has highlighted a number of issues in exploring the possible relationship between income inequality and health.

Relative income inequality may be as important as overall income inequality. A study of the incidence of Type 2 diabetes was conducted in Tayside, Scotland, using data reported for Output Areas, the smallest reporting areas for census data (Cox, Boyle, Davey, Feng, & Morris, 2007). Area deprivation was positively related to diabetes incidence, and the inequality in deprivation between areas and their neighbors was negatively related to diabetes incidence. Although the disease was more common in deprived areas, incidence was lower in deprived areas that were surrounded by relatively less deprived areas, and less deprived areas surrounded by relatively more deprived places had higher diabetes incidence than would be expected based on the level of deprivation in the area itself.

There is also evidence that income inequality may affect the health of different population subgroups differently. For example, there may be a ***cross-level interaction***—a modification of the effects of lower-level variables by characteristics of the higher-level units to which they belong or vice versa—such that income inequality has particularly harmful effects among those of low income. Geographic scale also matters in research on income inequality and health. For smaller geographic areas that may be more homogeneous in terms of income, social segregation might be more important than income inequality (Ross, et al., 2004). More longitudinal research is needed to investigate the possible temporal lag effects of increasing income inequality. In addition, occupation, race and ethnicity, and gender are all associated with income and may confound the relationship between income and health.

Occupation, along with income and education, are components of socioecononomic status. Although these measures are related, they are not interchangeable (Braveman et al., 2005). In the United States, ethnic and racial groups have different levels of income for a given level of education, different levels of wealth for a given level of income, and different levels of neighborhood socioeconomic status for given levels of individual socioeconomic status. Income and education are the most commonly used measures of socioeconomic status in the United States, in part because of the way occupations are classified in U.S. data systems. Studies on the relationships between occupation and health disparities conducted in western European countries have found strong relationships between occupation and health (Kunst, Groenhof, Mackenbach, & EU Working Group on Socioeconomic Inequalities in Health, 1998), even in studies that have controlled for income and education. Employment in hazardous occupations exposes individuals to health risks. To the extent that hazardous industries are concentrated in particular places, there are likely to be geographic variations in health tied to occupation and to individual behaviors like smoking or drinking alcohol that are often more prevalent among individuals in particular occupations (Stirling, 1978).

Area-based measures of material deprivation have also been used in health disparities research (Townsend, Phillimore, & Beattie, 1988; Carstairs, 1995; Cox et al., 2007). A ***deprivation index*** measures the absence of resources needed for survival. Indices like those developed in the United Kingdom by Townsend and Carstairs relied on census data, some using surveys of health experts to weight the census variables included in the index (Jarman, 1983). GIS and multicriteria analysis have also been used to identify patterns of social deprivation in British Columbia (Bell, Schuurman, & Hayes, 2007). As with occupation, income, and education, the components of a deprivation index may not be interchangeable and the resources of an individual or household living within an area may not match the resources measured at the area level (Macintyre, Hiscock, Ellaway, & Kearns, 2004).

In the United States, race and ethnicity have received particular attention in the investigation of health disparities. An analysis of the effects of income inequality on mortality in cities and states in the United States found no relationships in 1980 or in 1990 between income inequality and mortality across cities or states once the fraction of the population that was black was controlled (Deaton & Lubotsky, 2003). The analysis found that white mortality rates were higher in places where a higher fraction of the population was black, and this relationship also held within large regions of the country. A replication of the study essentially confirmed the findings and uncovered evidence of spatial variability in the relationship between inequality and health (Ash & Robinson, 2009). Both black mortality and white mortality were negatively affected by the fraction of population in metropolitan areas that was black, reflecting the presence of detrimental social and environmental determinants of health. The racial composition of the area was seen, in the American context, as evidence of inequality in political and economic power (Williams & Collins, 2001).

Gender is also related to health, although attention to the role of gender in relation to income, education, and occupation in health disparities is more recent than research focusing on gender and health alone (Macintyre & Hunt, 1997). Data from the First National Health and Nutrition Examination Survey (1971–1993) including baseline data for men and women on educational attainment and household income were analyzed with data on incident coronary heart disease drawn from hospital records and death certificates over 22 years of follow-up (Thurston, Kubzansky, Kawachi, & Berkman, 2005). Although the association between income and education and health was traditionally believed to be weaker for women than for men, the study found that having less than a high school education was associated with a stronger risk of coronary heart disease for women than for men.

Income, wealth, and other measures of socioeconomic status such as occupation and education that differ by race, ethnicity, and gender are important factors in individual well-being, but characteristics of the environment are also important. People in the same income group may live in settings that are very different in terms of the built environment. Characteristics of areas other than the socioeconomic status of their residents have also been studied in relation to health disparities.

The Geography of the Built Environment

The neighborhoods where we live are created by people. The ***built environment*** is everything that is made or maintained by humans to fulfill human purposes (Bartuska, 2007). Much of the built environment mediates or changes environmental conditions for human comfort or well-being, with results that have positive and negative effects on ecosystems and the human–environment relationships that they comprise. When population density is used as a measure, variations in human impact on the environment are, like variations in income, apparent at the global, regional, and local scales (Figure 11.3).

The development of GIS technology and spatially referenced databases has had a significant impact on studies of the relationships between the built environment and health because they have provided support for creating objective measures of the built environment and modeling neighborhood conditions across a wide range of settings. Nevertheless, research on the quality of the built environment and health emerged in the 19th century in response to the dramatic changes in the built environment associated with the Industrial Revolution (Chadwick, 1965; Davis, 1973). In 1889, the British Medical Association mapped the principal concentrations of rickets in England and Scotland, showing their correlation with industrial areas where smoke and overcrowding limited exposure to sunlight (Owen, 1889; Loomis, 1970). Rickets was then seen as a disease tied to the environment rather than individual behavior, given that diets in the urban industrial areas were generally better than those in poorer rural communities. Interestingly, there is renewed interest in the role of vitamin D in health (Holick, 2009). Recognition of the links between the built environment and health has continued in the decades since.

In 1948, the American Public Health Association published a set of standards highlighting "the basic health criteria which should guide the planning of residential neighborhood environment" (Committee on the Hygiene of Housing, American Public Health Association, 1948, p. v). The standards cover many aspects of the built environment: essential physical characteristics of housing sites; availability of water supply and sewage disposal, solid waste removal, power, fuel, and communications, fire and police protection; freedom from accident hazards, noise and vibration, odors, smoke, and dust, disease hazards, and moral hazards; access to community facilities by pedestrian, automobile, and public transit, and bicycle ways; and the availability of essential city and district facilities including education, retail, employment, outdoor recreation, and health services. Detailed quantitative standards are provided for many features. For example, tables show the total neighborhood park size in acres recommended for neighborhoods based on their type of housing and population size and the recommended distances for access to neighborhood facilities.

In the 1960s, as part of the Detroit Geographical Expedition and Institute, geographers and community residents of neighborhoods in Detroit documented characteristics of the built environment (Bunge, 1971; Horvath, 2006).

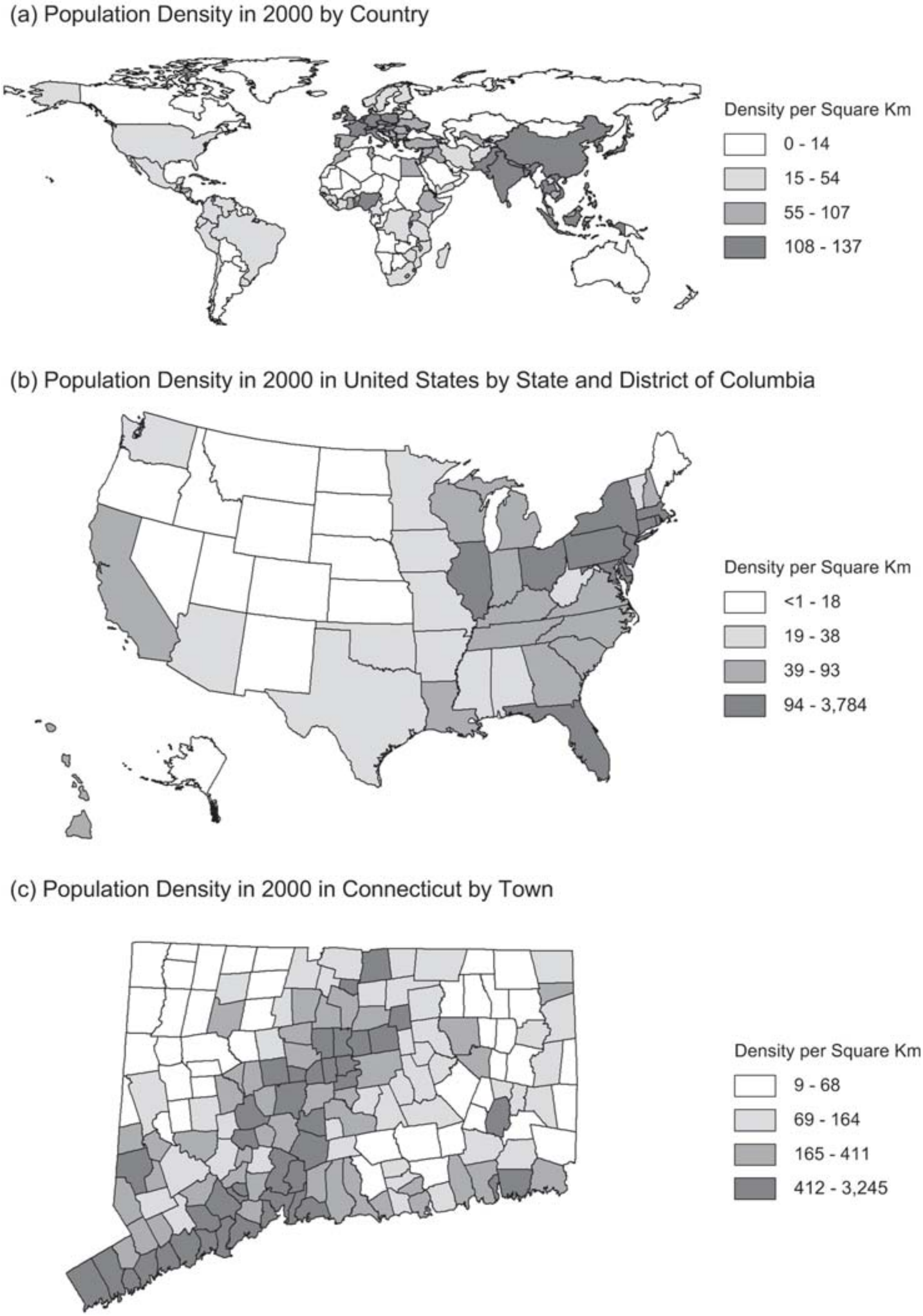


FIGURE 11.3. Variations in population density at the global, regional, and local scales.

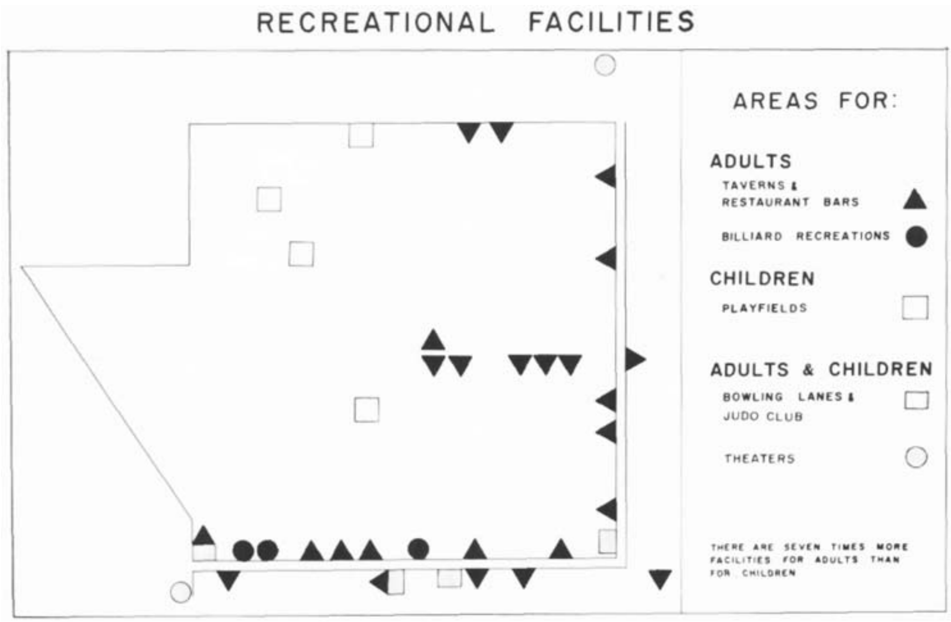


FIGURE 11.4. Recreational facilities for children and adults in the Fitzgerald neighborhood of Detroit, Michigan, 1966. The neighborhood had more facilities for adults than for children. From Bunge (1971, 2011). Copyright (2011) by University of Georgia Press. Reprinted by permission.

Their research yielded information on the location of neighborhood recreational facilities (Figure 11.4). They also found that the prices of milk, peanut butter, and other foods were 25 to 79% higher in the inner city compared to the suburbs in Detroit chain stores. These earlier efforts, however, did not benefit from the spatial data handling, measurement, and analysis capabilities of GIS, which have made it possible to measure and analyze characteristics of the built environment in health studies involving large numbers of people and many regions.

MEASURES OF THE BUILT ENVIRONMENT

Like income, which can be measured in a number of ways, the built environment is a multidimensional concept. It encompasses both the local physical infrastructure such as roads, sidewalks, residences, and facilities and the local social infrastructure such as networks of community support, social order, and social capital, both as objectively measured and as subjectively perceived (Macintyre, Ellaway, & Cummins, 2002). ***Objective measures*** of the built environment are based on measurements of geographic features—for example, the density of the street network in an area or the number of parks within a certain distance from a residence. ***Subjective measures*** of the built environment are based on residents' perceptions of environmental conditions—for example, how many parks a person believes to be within a specified distance from the residence.

Geographic information systems have proven highly useful in developing objective measures of the built environment for health research. Many of these measures such as population density, street network connectivity, land use mix, and facility density were first developed in urban planning and transportation, and the formulas for calculating the measures were published before the widespread adoption of GIS. Features of the built environment have been used in studies of specific health outcomes like overweight and obesity or substance use and in relation to general health.

THE BUILT ENVIRONMENT AND OBESITY

The rapid increase in overweight and obesity in the populations of high-income countries has been one of the most notable public health problems of the last two decades (Caballero, 2007; Friel, Chopra, & Satcher, 2007). Overweight and obesity in individuals are tied to ***energy balance***, the difference between the calories consumed in food and drink and the energy expended through physical activity. Food environments and physical activity environments have both been studied in relation to overweight and obesity.

The ***community nutrition environment*** is described by the number, type, and location of food sources, including groceries, convenience stores, restaurants, farmer's markets, and other outlets (Glanz, 2009). Communities with limited access to grocery stores and other retail outlets offering nutritious foods at reasonable prices have been labeled ***food deserts*** (Reisig & Hobbiss, 2000; Walker, Keane, & Burke, 2010). Data on the location of food outlets can be obtained by neighborhood survey and from license records, directories, and commercial databases like Dun & Bradstreet or InfoUSA (Wang, Gonzalez, Ritchie, & Winkleby, 2006). Commercial databases of facilities with geographical coordinates are sometimes bundled with GIS software. Otherwise, spatial databases of facility locations are developed from surveys, license records, or directors by address-match geocoding, as discussed in Chapter 3.

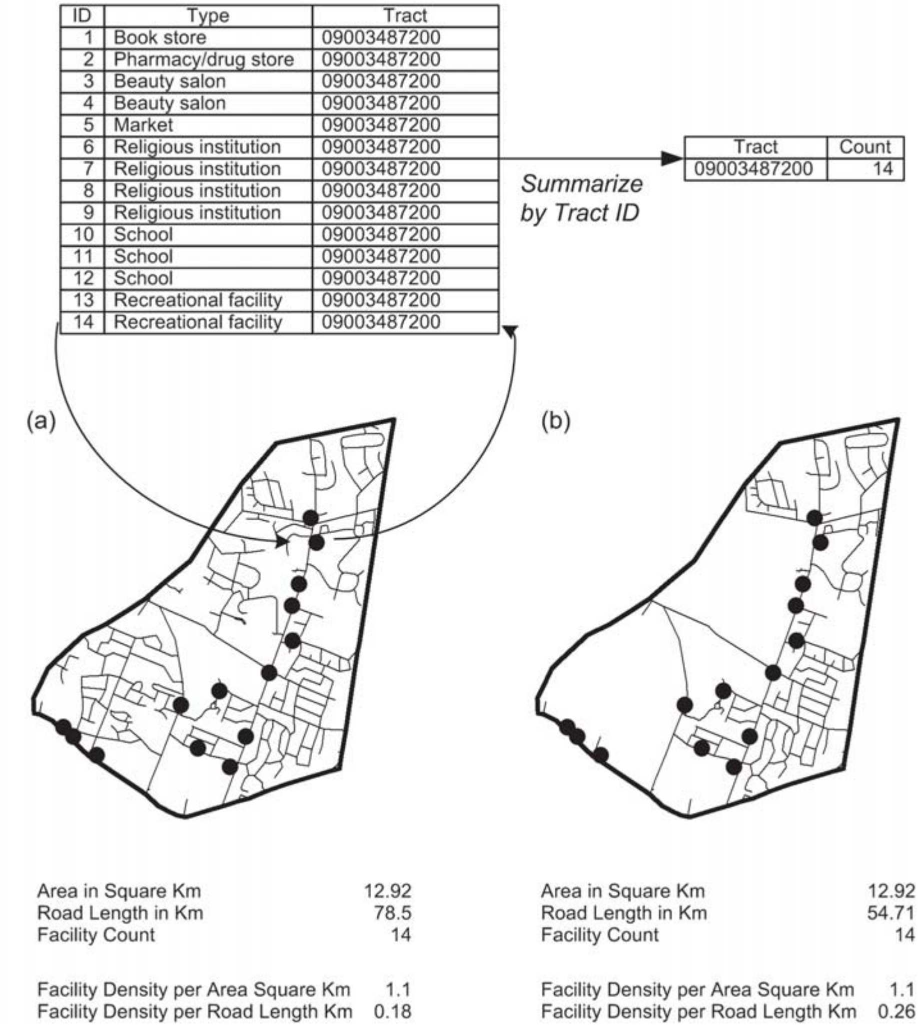
Increasingly, efforts are being made to validate data on facilities used to characterize community nutrition environments and environments for physical activity (Boone, Gordon-Larsen, Stewart, & Popkin, 2008a; Bader, Ailshire, Morenoff, & House, 2010). The methods used to assess completeness and positional accuracy of the data on facilities and facility locations have been criticized (Zandbergen, 2008; Boone, Gordon-Larsen, Stewart, & Popkin, 2008b). However, an equally serious weakness of these studies is that not enough attention is paid to the spatial distribution of the errors. If errors are concentrated in particular places, then the measure of neighborhood environmental quality will be more valid for some individuals than for others.

With these caveats, these databases have been useful for describing the locations of food outlets. GIS have been used to map the locations of food outlets in relation to overweight and obesity and to measure accessibility to and density of food outlets in different neighborhoods. Accessibility can be measured using techniques described in Chapter 9. Density of food outlets can be derived by performing a point-in-polygon analysis to determine the area where a facility is located, summarizing the data by area, and then joining a table of facility counts by area to a GIS database of area features. GIS software functions can return the geographic area of each polygon as an attribute, and density can be calculated by dividing the count of facilities in a polygon by the length of the street network in the zone or by the area of the polygon (Figure 11.5). Density can also be calculated in relation to local population size.

There is less information in facility databases, however, about the characteristics of facilities, for example, the kinds of food actually available in the outlets. This information usually requires community surveys of the types of food for sale, because the purchasing and sales records of the stores are proprietary data. The ***consumer nutrition environment*** includes what people encounter in and near the places where they buy food (Glanz, 2009). The connections between the locations of food outlets, what people actually eat, and their weight are complex (Lytle, 2009).

As an example, research in a number of countries has demonstrated that disadvantaged areas have better geographic accessibility to fast-food outlets and are likely to lack supermarkets (Block, Scribner, & DeSalvo, 2004; Cummins, McKay, & Macintyre, 2005; Burns & Inglis, 2007). Yet, a national study of the association between neighborhood access to fast-food outlets, diet, and weight found that residents who were furthest from multinational fast-food outlets were more likely to have the recommended intake of vegetables but were also more likely to be overweight, controlling for individual and area characteristics (Pearce, Hiscock, Blakely, & Witten, 2009). With the growing number of studies conducted in a wide range of community settings, variations in these effects among regions within countries have been noted. A study conducted in a range of community settings in Scotland found that the most deprived areas had the highest access measured in terms of travel time to grocery stores and stores selling fresh produce (Smith et al., 2010). Access to fresh produce was lowest among affluent island communities. A study conducted in census tracts located in Mississippi, North Carolina, Maryland, and Minnesota found that people living in areas with particular combinations of food stores had higher obesity and overweight (Morland, Diez Roux, & Wing, 2006). In particular, people in areas with grocery and/or convenience stores but no supermarkets were at risk. People who lived in areas with only supermarkets were less likely to be overweight or obese.

Measures of the built environment have also been used in studies of physical activity (Sallis, 2009). Data on the locations of parks and recreation facilities are compiled in some of the same ways as data on food sources, from neighborhood survey, from public records, from directories, and from commercial databases. As with food sources, the databases on parks and recreation facilities gener-



ID	Type	Tract
1	Book store	09003487200
2	Pharmacy/drug store	09003487200
3	Beauty salon	09003487200
4	Beauty salon	09003487200
5	Market	09003487200
6	Religious institution	09003487200
7	Religious institution	09003487200
8	Religious institution	09003487200
9	Religious institution	09003487200
10	School	09003487200
11	School	09003487200
12	School	09003487200
13	Recreational facility	09003487200
14	Recreational facility	09003487200

Tract	Count
09003487200	14

FIGURE 11.5. GIS can be used to map facilities in the built environment. A point-in-polygon analysis can be used to identify the area, like a census tract, within which each facility is located. GIS functions can be used to summarize the table of individual facility points and their census tract locations to obtain a count of facilities in each area. The density of facilities per square kilometer of area is calculated by dividing the facility count by the area of the tract in square kilometers. Because facilities are usually located on streets, facility density by area road length might also be calculated by dividing the facility count by the total road length in the area in kilometers. Note that facility density per square kilometer is the same in Figures 11.a and 11.b above, but the facility density based on the road network is about 30% higher in the area shown in Figure 11.b above, because there are fewer roads. Two census tracts with the same number of facilities and same area can have different facility densities depending on the road network.

ally provide little information on the attributes of these places. The same kind of distinction between community nutrition environment and consumer nutrition environment applies to the community built environment and community physical activity environments. Because physical activity can take place almost anywhere, research on the built environment and physical activity encompasses a wide range of measures beyond the presence of parks and recreation facilities.

Three main approaches to collecting data on the built environment as it relates to physical activity have been identified: reports from individuals surveyed by telephone or self-administered questionnaires on their perceptions of features in the environment; systematic field audits of neighborhoods and recreational facilities; and geographic databases compiled using GIS (Brownson, Hoehner, Day, Forsyth, & Sallis, 2009). Many of the audit tools have been validated (Sallis, 2009). The validity and reliability of geographic databases used to derive measures of the built environment are more difficult to assess, although studies are being conducted in the field to validate data on facilities, as noted above.

Nevertheless, geographic databases have been particularly useful in studies that cover a large number of dispersed neighborhoods. Five main classes of measures derived from spatial databases have been used in physical activity research: population density, land use and land use mix, access to recreational facilities, street network characteristics, and attributes of the street network environment that may enhance or discourage physical activity such as presence of sidewalks, levels of traffic, and crime. Sometimes, composite variables or indices are developed from these measures.

As discussed in Chapters 2 and 3, the foundation databases available for deriving measures of the built environment vary widely from place to place, and the costs of collecting and processing the required data can be high. Data on the availability of sidewalks, for example, are often difficult to obtain. It may not be possible to replicate studies conducted in one setting in other places or to find comparable data for studies conducted across a wide range of communities. Also, analytic procedures can be implemented in different ways depending on the software and on the analyst affecting the reliability and validity of the measures. To begin addressing these issues, attention has been paid to developing standards for documenting GIS procedures and data in research on food and physical activity environments (Forsyth, 2007).

A study of recreational cycling in Melbourne, Australia, examined associations between cycling, individual characteristics, area socioeconomic characteristics, and objectively measured built environment characteristics assessed by audits (Kamphuis et al., 2008). The built environment characteristics included design features (cycling paths, street width), safety (lighting, traffic control), facilities (bicycle parking, retail stores, recreational facilities), and aesthetics (views, maintenance). Residents of deprived areas were less likely to cycle for recreation, and aesthetics were worse in deprived areas. Two safety characteristics, the level of surveillance and the absence of driveway intersections, were significantly associated with recreational cycling after adjusting for safety features and compositional factors.

THE BUILT ENVIRONMENT AND SUBSTANCE USE

Built environment measures have also been used in research on substance use, especially tobacco and alcohol. In these studies, the focus is primarily on density of outlets and accessibility to outlets. The density of tobacco retail outlets around schools has been linked to adolescent smoking initiation and purchasing habits (Leatherdale & Strath, 2007; McCarthy et al., 2009). For adults, the density of tobacco outlets around the home, as well as the proximity of tobacco outlets to the home, has been associated with the number of cigarettes consumed per day among adult smokers (Chuang, Cubbin, Ahn, & Winkleby, 2005).

The picture with alcohol use is more complex, in part because of the array of outlets for alcohol sales and connections between drinking setting and drinking behavior. Higher densities of off-campus alcohol outlets allowing on-premise drinking were strongly related to drinking outcomes in college students, controlling for individual predictors of college drinking (Scribner, Mason, et al., 2008). The structure of drinking opportunities may also affect other health outcomes. Declines in gonorrhea rates from 1988 to 1996 were steeper in Los Angeles neighborhoods where more alcohol outlets closed following a period of social unrest in 1992 (Cohen et al., 2006). Even after controlling for the effects of property damage, the number of alcohol outlets per roadway mile was positively associated with gonorrhea rates. It is possible, however, that community efforts to prevent closed alcohol outlets from reopening were associated with other changes in the neighborhoods that influenced sexual risk taking in the local population.

Distinct patterns of drinking outlet utilization were found among age, gender, and ethnic subgroups in a study of 25,000 drinkers in communities in California and South Carolina. These patterns were differentially linked to acute drinking problems (Treno, Alaniz, & Gruenewald, 2000). Reasons for drinking and drinking setting together influenced consumption in a study of more than 8,000 students at 18 universities (Kairouz, Glicksman, Demers, & Adlaf, 2002). Students drank for different reasons in different contexts. A comparative multilevel analysis of contextual drinking in American and Canadian adults concluded that interactions between locations and demographic variables of individuals may differ in different societies (Kairouz & Greenfield, 2007).

THE BUILT ENVIRONMENT AND GENERAL HEALTH

Efforts have also been made to develop more comprehensive sets of built environment measures that might be linked to health. One innovative yet vital approach begins with a set of human needs, including air, water, food, shelter, security, education, information, social relationships, political capital, and play (Cummins, Macintyre, Davidson, & Ellaway, 2005). Approaches to finding and collecting data to operationalize community characteristics that meet these needs are described.

A study using specific measures hypothesized to be important for a healthy life found that a number of features were associated with self-rated health (Cummins, Stafford, Macintyre, Marmot, & Ellaway, 2005). The study used data from the Health Survey for England and the Scottish Health survey for almost 14,000 men and women 16 years of age or older. Six neighborhood attributes were significantly associated with fair to very bad self-ratings of health. Poor physical quality of the residential environment, lower political engagement, higher unemployment, and lower access to private transportation were among the attributes associated with poorer health after adjusting for individual age, sex, social class, and economic activity. The associations were larger for unemployed residents than for employed residents, suggesting that the effect of some built environment characteristics, like neighborhood income, may differ across groups.

INCOME, RACE AND ETHNICITY, GENDER, AND THE BUILT ENVIRONMENT

Relationships between income and built environment characteristics were investigated in a cross-sectional study of 32 neighborhoods in Seattle, Washington, and Baltimore, Maryland (Sallis et al., 2009). Daily minutes of moderate to vigorous physical activity measured by accelerometer were greater in areas with high walkability, and the relationship did not differ by income. Walkability was measured through an index based on built environment features including intersection density, net residential density, retail floor area ratio, and land use mix.

Without measuring any area characteristics directly, the effects of race and ethnicity as opposed to area characteristics were investigated in the United States by using dummy variables for each place (Do et al., 2008). Data on self-rated health from the National Health Interview Survey were used. Controlling for residential context reduced disparities in self-rated health among blacks and whites that could not be accounted for by individual-level controls by 15 to 76%. The contribution of residential context to explaining disparities between blacks and whites declined with age and was smaller among females than males.

A study conducted using health survey data for England and Scotland found that contextual effects including measures of the built environment were related to self-reported health for men and women (Stafford, Cummins, Macintyre, Ellaway, & Marmot, 2005). The study included measures of the built environment, including access to food stores, number of public recreation sites, health services providers, and vacant and derelict land. The size of the effects was larger in every case for women than for men.

Summary

There is clearly evidence that the socioeconomic and environmental characteristics of areas affect the health of the people who live in them, although the results of individual studies conducted in different settings using different measures and methods vary. It is important to include a range of community characteristics in studies of health disparities. The poorest residents do not always have the worst access to facilities (Macintyre, Macdonald, & Ellaway, 2008; Smith et al., 2010), in part because higher incomes enable people to live in low-density residential areas distant from facilities. Not all deprived areas have high levels of pollution and not all polluted areas have poor populations, but some places have a "double burden" of economic deprivation and poor environmental quality (Crouse, Ross, & Goldberg, 2009). The fact that some health patterns are not fully explained by the composition of areas or available measures of area conditions (Riva, Curtis, Gauvin, & Fagg, 2009) is leading researchers to develop better ways of conceptualizing neighborhoods and to investigate spatially varying processes.

Defining Neighborhood Contexts

One of the most challenging aspects of exploring the effects of neighborhoods on health is defining the neighborhood (Spielman & Yoo, 2009). In general, there are three approaches to defining neighborhoods in research concerned with contextual effects. First, reporting units like census tracts or postal code areas can be used. Second, GIS functions can be used to create neighborhoods that reflect some kind of spatial process. Finally, individuals' reports of conditions in their neighborhoods can be used, although these reports are usually collected without obtaining an explicit description of the spatial extent of the perceived neighborhood.

For contextual variables such as income and measures of socioeconomic status which are usually aggregated from individual census returns, the fact that data are reported for political/administrative units like states, counties, or census tracts often leads analysts to define the "neighborhood" contexts in terms of these units (Krieger et al., 2003). The advantages and disadvantages of some units over others—for example, census tracts versus ZIP Codes—have been considered (Krieger et al., 2002). Whatever reporting unit is used, the neighborhood of the individual is determined by performing a point-in-polygon analysis in GIS to identify the areal unit within which the individual resides or engages in some other activity.

Data reported for political/administrative zones are subject to the modifiable areal unit problem (MAUP), also discussed in Chapter 4. Flowerdew, Manley, and Sabel (2008) investigated the issue of neighborhood definition by creating several sets of "pseudo-wards," essentially redrawing existing ward boundaries in communities in England. The different sets of alternative zones had approximately the same number of zones, but the zones were drawn with different boundaries designed to reflect different properties, including equality of size, regularity of shape, and internal homogeneity. Each of these properties can be measured for a set of zones using the statistical formulas given. The impact of using different zonation schemes, including the actual ward zones, on the observed relationships between neighborhood contextual characteristics and health outcomes was modeled by analyzing the relationships for each of the zonal schemes. The actual ward boundaries, like most administrative boundaries, were found to have been drawn in a way that arbitrarily partitions the underlying distribution of interest, and the conclusion from the initial analysis was that there was no apparent neighborhood effect on health. When other zonal systems were used, significant neighborhood effects on health were uncovered. This analysis provided clear evidence that where the boundaries of political/administrative units are drawn matters.

The socioeconomic perspective is only one way of looking at neighborhoods (Lebel, Pampalon, & Villeneuve, 2007). For contextual variables such as many of the built environment variables that are not reported in the census, GIS have been used to develop neighborhoods whose boundaries do not necessarily correspond to reporting units but instead reflect some underlying process. Some studies have used circular buffers around people's homes to describe the neighborhood of interest, but the network analysis capabilities of GIS discussed in Chapters 9 and 10 have made it possible to model so-called ***network buffers*** based on distance measured along the street network (Figure 11.6). These buffers may more accurately reflect how people actually travel in the environment, although care must be taken in choosing the buffer distance. Population density,

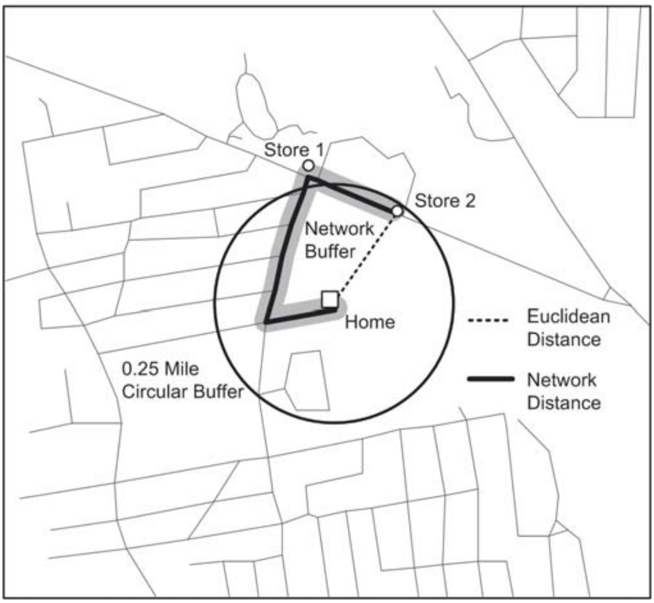


FIGURE 11.6. The consequences of using a circular buffer versus a network buffer to define the neighborhood around a residence. The Euclidean distance between Home and Store 2 is just under a quarter of a mile (1,296 feet) so the store would be considered within the neighborhood of the home as defined by the circular buffer area. The network distance from Home to Store 2 is 3,609 feet. Store 1 is actually closer to Home than Store 2 by almost a quarter of a mile, but it would be outside the neighborhood defined by the circular buffer.

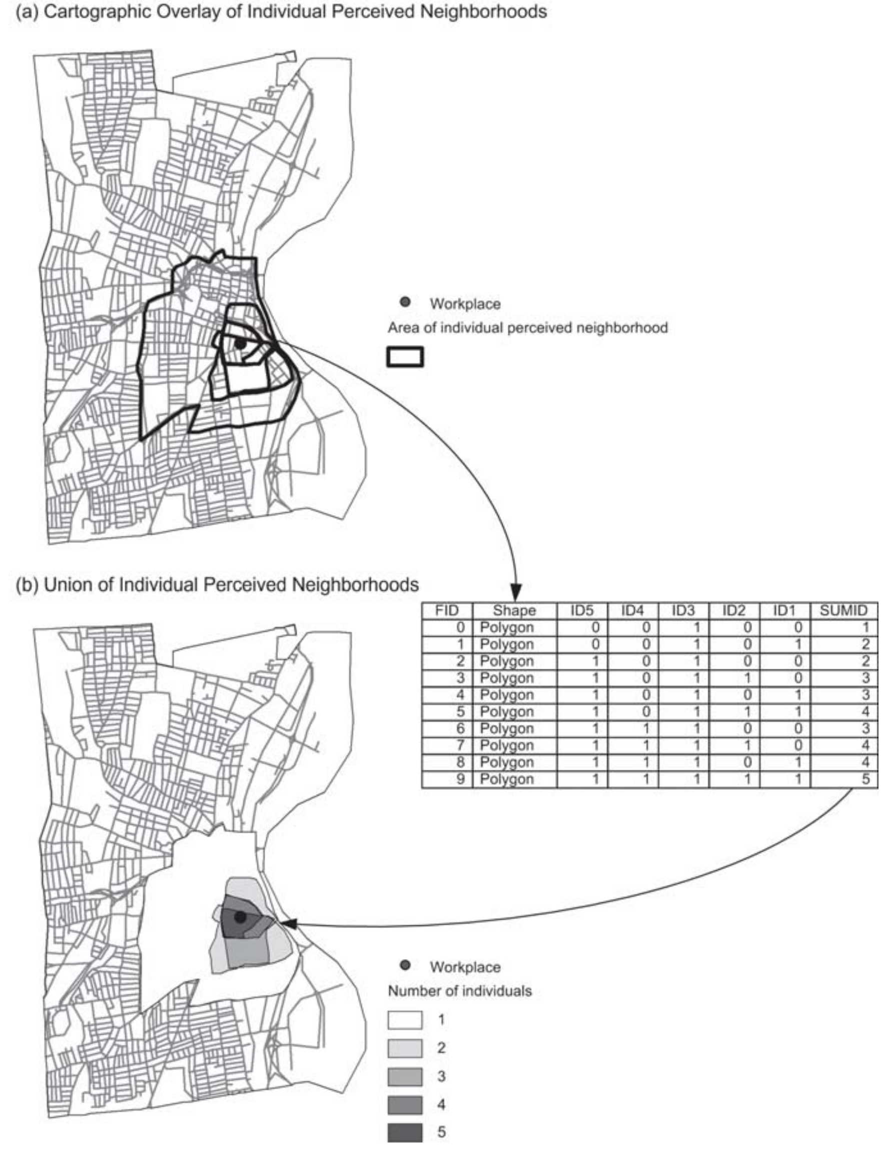
number and density of retail outlets, land use, and other measures of the built environment can then be derived for these areas.

In addition to modeling the area within a certain street network distance from a person's home, school, or workplace, geospatial technologies are being used to identify the actual neighborhoods where people engage in activities of interest. GPS receivers have been used with accelerometers to monitor individuals' levels of physical activity and where the activity actually occurs (Troped et al., 2010). An ***accelerometer*** is an electromagnetic instrument that measures acceleration. Measurements from these devices, which can be worn by study participants, are analyzed to assess when participants are engaged in low, moderate, or vigorous physical activity. These studies are useful in identifying where physical activity actually occurs and suggesting what the appropriate buffer distances might be to capture the characteristics of neighborhoods as contexts for physical activity.

In addition to their social, spatial, and statistical definitions, neighborhoods have a perceptual dimension. Our cognitive maps include our perceptions of districts or neighborhoods (Lynch, 1960). When people are asked in surveys to assess conditions in their neighborhoods, they are rarely asked to describe what they perceive as the boundaries of the neighborhood. Nevertheless, GIS can be used to document these neighborhoods, too. The boundaries of an individual's perceived neighborhood can be represented on a map either by translating the person's description to the map or having the person draw the limits of the neighborhood on the map (Figure 11.7). These maps can be scanned, and the neighborhood boundaries can be screen digitized or acquired by vectorization of the raster image. To identify the area that most people who live in the same place would consider as their neighborhood, a union operation can be performed on the individual data layers. Similar techniques can be used to investigate the spatial coincidence of various historical zones used to define neighborhoods over time (Lebel, Pampalon, & Villeneuve, 2007). It may not always be practical in large-scale studies to investigate participants' perceived neighborhoods in this way, and there is likely to be disagreement about neighborhood boundaries (Flowerdew, Feng, & Manley, 2007). Neighborhood defined according to people's activity spaces or their perceptions will likely vary among individuals living in the same place in terms of age, gender, social roles, and mobility (Cummins et al., 2007). Nevertheless, these techniques can be useful in incorporating local knowledge into studies modeling neighborhood effects on health.

Modeling Neighborhood Effects on Health

Multilevel modeling, spatial modeling and analysis in GIS, and other techniques are commonly used to assess neighborhood effects on health (Luke, 2005). These methods have been employed in cross-sectional research on the relationships between health and area characteristics. Studies using multilevel modeling techniques have raised important conceptual and methodological issues about



FID	Shape	ID5	ID4	ID3	ID2	ID1	SUMID
0	Polygon	0	0	1	0	0	1
1	Polygon	0	0	1	0	1	2
2	Polygon	1	0	1	0	0	2
3	Polygon	1	0	1	1	0	3
4	Polygon	1	0	1	0	1	3
5	Polygon	1	0	1	1	1	4
6	Polygon	1	1	1	0	0	3
7	Polygon	1	1	1	1	0	4
8	Polygon	1	1	1	0	1	4
9	Polygon	1	1	1	1	1	5

FIGURE 11.7. Using GIS to model perceived neighborhoods of individuals at a common workplace. The individual perceived neighborhood areas were screen digitized from scanned maps as shown in Figure 11.7a. The individual polygons were combined using GIS polygon overlay procedures to find the union of the first pair of neighborhoods, the union of the result with the next neighborhood, and so on, carrying forward the ID numbers of each individual as shown in the table and in Figure 11.7b. The number of individuals who counted the area of a union polygon as part of their perceived neighborhood of the workplace can be summed and mapped to identify the areas of agreement for all individuals.

how neighborhoods are defined and about spatial dependencies across neighborhoods. Research has shown that the relationships between neighborhood characteristics and health may themselves be spatially variable. Finally, there is a growing body of work that considers how change in the locations of individuals through migration and changes in neighborhood areas over time affect health disparities.

Multilevel Modeling and Health Disparities

Multilevel modeling is one of the most widely used methods for studying health disparities (Diez Roux, 2000). For health research, multilevel models view individual health outcomes as influenced by factors at different levels. For example, individual health disease risk is influenced by individual characteristics such as age and gender and characteristics of the census tract of residence. A two-level multilevel model is generally formulated as follows:

$$Y_{ij} = \gamma_{00} + \gamma_{01}C_j + \gamma_{10}I_{ij} + \gamma_{11}C_{ij}I_{ij} + U_{0j} + U_{1j}I_{ij} + e_{ij}$$

where Y_{ij} is the health outcome for the *i*th individual in the *j*th context or group, γ_{00} is the common intercept across contexts, γ_{01} is the fixed effect of group-level variable C_j, γ_{10} is the fixed effect of individual-level variable I_{ij}, and γ_{11} is the effect of their interaction $C_{ij}I_{ij}$ on the individual outcome. The remaining terms specify a complex error structure where U_{0j} is a random intercept, U_{1j} is a random slope, and e_{ij} are individual errors. The random intercept and random slope components of the error for observations within groups are correlated, and error variance is not constant.

This formulation is for a simple model, but more complex models with different assumptions about the distribution of the dependent variable, fixed and random effects, nonlinear relationships, and overlapping or cross-classified contexts are possible. Software functions for solving multilevel models have not been incorporated into most GIS packages. Instead, GIS are used for deriving neighborhood environment characteristics and for determining which neighborhood contexts a person falls within.

Multilevel modeling was used to investigate whether or not characteristics of the neighborhood internal housing and the external built environments were related to depression in an analysis of 1,355 residents of New York City (Galea, Ahern, Rudenstine, Wallace, & Vlahov, 2005). The 59 community districts in the city were used as the neighborhoods in the study. The New York City Housing and Vacancy Survey provided data on internal characteristics of housing units in neighborhoods and on external characteristics of properties in each neighborhood, along with some additional information on the external environment provided by the Fire Department and the Department of Sanitation. Data on neighborhood socioeconomic characteristics were also included. Adjusting for an individual's age, race/ethnicity, sex, and income and for the neighborhood's level of income, the research showed that people who lived in neighborhoods with poorer quality internal and external built environments were more likely to report having symptoms of depression in the last 6 months and lifetime depression. A review of cross-sectional and longitudinal studies of depression and neighborhoods found that many, but not all, confirmed associations between at least one built environment characteristic and depression, despite the variability in neighborhood definitions and measures, study populations, and study design (Mair, Diez Roux, & Galea, 2008).

Modeling Spatial Effects

A key weakness of multilevel modeling as an approach to studying health disparities is that spatial effects are not always explicitly addressed. In studies where neighborhoods are contiguous, spatial autocorrelation may be present, as discussed in Chapter 5. The maps in Figures 11.2 and 11.3 suggest that there often are spatial patterns in income and built environment characteristics at all scales. For example, poor neighborhoods may be clustered together within a metropolitan region. Spatial regression models have been used in research on health disparities to account for spatial autocorrelation.

A typical model in an ordinary least squares regression specification takes the general form

$$y_i = \alpha + \sum_k \beta_k x_{ik} + e_i$$

where y_i is the observed outcome for i, α is the intercept (a constant), x_{ik} is the value of the independent variable k for observation i, β_k is the estimated parameter for variable k, and e_i is error (the difference between the observed and expected values for y_i). In ordinary least squares regression, e is random error. If spatial autocorrelation is present, the error terms are correlated with each other and nonrandom, therefore violating one of the basic assumptions of the regression model and resulting in incorrect parameter estimates.

Spatial regression models incorporate direct modeling of spatial autocorrelation (Anselin, 2003b). The model includes a spatial weights matrix (see Chapter 5) describing neighborhood relationships among the set of observations I in relation to the set of observations as neighbors J. A spatial errors model takes the general form

$$y_i = \alpha + \sum_k \beta_k x_{ik} + \lambda \sum_j w_{ij} e_j + \mu_i$$

Here, the nonrandom spatial error terms are incorporated into the third term, which captures the spatial structure of the spatially dependent error, where λ is the spatial autoregressive coefficient or error parameter and $\Sigma_j w_{ij} e_j$ is the sum of the spatial weights multiplied by the spatially dependent error with respect to observation *i*. The last term, μ_i, is the random error. The spatial error model formulation addresses the situation of spatial autocorrelation in the error term alone. It is also possible to design spatial lag models that include a term for spatial autocorrelation in the dependent variable as mentioned in Chakraborty (2009).

A spatial error model estimated using GeoDa software (Anselin, 2003a) was used to predict neighborhood scores on a hazard density index tied to the presence of *maquiladoras* (final assembly manufacturing plants) in Ciudad Juárez, Mexico (Grineski & Collins, 2008). Independent variables included neighborhood-level social class measured using a combination of variables, the percent of children under 14 years of age in the neighborhood, and the neighborhood level of formal housing development measured using a combination of variables measuring the quality of residential construction and availability of water and sewer. Because the neighborhood units were heterogeneous in size and shape, distance rather than contiguity (see Chapter 5) was used to define neighbors in developing the spatial weights. In the inverse distance weighting procedure, neighbors of a given neighborhood unit were defined as units whose centroids were within 3,000 meters of the centroid of the given neighborhood unit. The percent of children under 14 was a significant positive predictor of hazard, but social class was not. When housing quality was included, however, all three variables were significant, indicating that neighborhoods of lower social class, better quality housing, and higher percentages of children had significantly higher scores on the hazard density index.

Spatial error models are global in scope because the regression parameters are interpreted to be constant across the study space, with variation resulting from the spatial heterogeneity of the variables alone. Local spatial statistics have been developed for situations where there is spatial variability in the parameters as well as in the explanatory variables (Fotheringham, Brunsdon, & Charlton, 2002). Investigations of spatially varying processes are also relevant to research on health disparities.

Spatially Varying Processes

Global statistics summarize data for entire regions yielding a single statistic. As such, they may mislead analysts about the nature of relationships in particular places. Local statistics summarize data for individual places within entire regions. They yield multiple statistics, one for each place, and these statistics can be mapped using GIS. Global and local statistical methods for studying clusters of health events are discussed in Chapter 5.

Local statistics are useful in exploratory data analysis, confirmatory analysis, and the development of global models. The measured particulates at individual air monitoring stations are local statistics, whereas the mean level of particulates across a state is a global statistic. When data are compiled and reported at a disaggregate level, local statistics can be calculated. When data are reported only for a single entity, such as for states or the nation as a whole, local statistics cannot be calculated and GIS can have only a limited role.

Geographically weighted odds ratios and geographically weighted regression are local statistics that have been used to study health problems. In a study of individual, behavioral, and environmental factors contributing to motor vehicle collisions in Connecticut, geographically weighted odds ratios were calculated for areas where high numbers of collisions occurred (Cromley, 2007). In the state as a whole, fixed object crashes (striking an object like a telephone pole or a tree) accounted for 19% of crashes. In some of the high-frequency collision zones, the local proportion of fixed object crashes was significantly higher. In the state as a whole, 65% of crashes occurred on dry roads. At some of the high-frequency sites, weather was a greater factor, with only 10% of crashes occurring when roads were dry. Local odds ratios for the role of weather differed significantly from the global odds of weather contributing to a collision. This suggests that, in addition to weather conditions themselves being spatially variable, the connection between weather and crashes is different in different places.

Geographically weighted regression (GWR) extends the basic regression model to provide locally varying parameter estimates (Fotheringham, Charlton, & Brunsdon, 1998). The model takes the form

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + e_i$$

where y_i is the predicted value of the dependent variable for location i, x_{ik} is the value of the independent variable k at point i, β_0 and β_k are continuous functions of (u_i,v_i), the spatial coordinates of the ith point, and e_i is the error at point i.

GWR was used in a study of mortality in the Atlanta metropolitan area (Holt & Lo, 2008). Socioeconomic status, race, and urbanization were significantly associated with standardized mortality rates across 431 census tracts in the area. Local parameter estimates were mapped. Low socioeconomic status, for example, was found to be particularly significant in estimating mortality in the northern suburbs of Atlanta, but had much less significant effect in the southern region of the study area (Figure 11.8).

GWR has also been used to explore environmental inequities in the density of air toxic releases in New Jersey (Mennis & Jordan, 2005). In some areas, the association between minority populations and concentration of air toxic releases was mediated by high poverty rates. In other areas, the presence of industrial and commercial land uses was a factor.

Like spatial error and spatial lag models, GWR can be performed using freestanding software (Fotheringham, Charlton, & Brunsdon, 2009). GWR functions have also been fully coupled with some GIS software packages. These analyses show the important role of GIS in investigating the links between location and well-being.

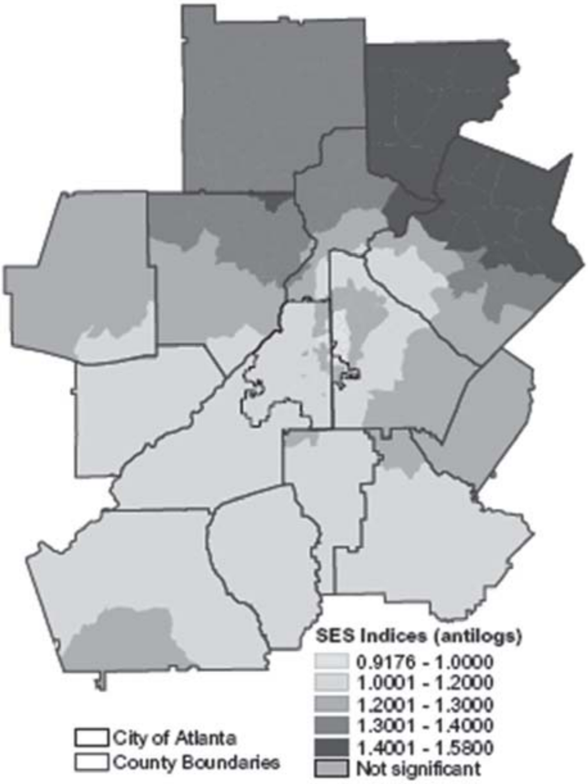


FIGURE 11.8. Geographically Weighted Regression (GWR) parameter estimates for the socioeconomic status variable in a model predicting mortality reveal that the relationship between socioeconomic status and mortality varied in the Atlanta metropolitan area. Low socioeconomic status was significant in predicting mortality in the northern suburbs but had a lower and less significant impact in the southern parts of the area. Reprinted from Holt and Lo (2008). Copyright 2008 by Elsevier. Reprinted by permission.

Location Processes and the Link between Location and Well-Being

Location processes are the mechanism by which the income and neighborhood environmental inequalities underlying health disparities are created and maintained (Cox, 1979). In a very real sense, we create each other's environments through our locational choices and our behavior in the environment. Our activities generate ***externalities***, unpriced costs and benefits for our neighbors. The location of a plant for incinerating municipal solid waste generates pollution that also generates external income and environmental quality effects if it lowers the value of neighboring properties and exposes their residents to a health hazard. These are examples of negative asymmetric externalities. The location of a park, on the other hand, enhances the value of neighboring properties and provides easy access to recreational facilities, an example of positive asymmetric externalities, if the area also has low levels of crime. These external effects are asymmetric because the costs and benefits are not shared. Some positive and negative externalities are symmetric. When individuals in a community are immunized, there is an individual benefit, and the increase in immunity in the population as a whole makes it less likely that an outbreak of infectious disease will occur—a positive symmetric externality. Congested highways generate negative symmetric external effects in that all of the vehicles are contributing to the problem they are experiencing.

All activities generate positive and negative effects. Even facilities or activities that are seen as goods are not necessarily good neighbors. Most of us want to know that an emergency medical vehicle can reach us within several minutes, but we would not want to live next door to a fire department or an ambulance station. Often negative external effects are highly localized in space while the benefits are more diffuse. "Neighborhoods" can be thought of as the resources and risks associated with particular locations (Bernard et al., 2007).

Locational conflict occurs when differences of opinion arise about how location and flows of goods and people should be regulated. In federal states like the United States, these conflicts can lead to considerable geographical variation in the juridical context, the existing laws and their administration, at state and local levels. Geographical differences in fatal occupational injury rates are significant in the United States, ranging from 1.7 per 100,000 in Connecticut to 24.3 per 100,000 in Alaska (Loomis et al., 2009). After controlling for differences in industry and workforce composition, higher rates of fatal occupational injury in 1980 and 1995 were associated with state policies favoring business over labor.

When external effects decrease income or environmental quality sufficiently, people have two choices. They can move, or they can take action to regulate the location choices and behaviors of others. Because moving is expensive and there is no guarantee that neighborhood quality will be preserved in the new location, people and institutions frequently engage in the political process to regulate the location and movement of others, including government institutions that themselves locate facilities. These regulations, at all levels of government and affecting all scales, are directed at controlling the locations and activities of other actors to improve the real or perceived local welfare advantage. Research has examined the effect of migration and neighborhood change on health.

Migration

One response to research showing contextual effects on health, especially built environment effects, is that studies do not control for self-selection (Riva, Gauvin,

& Barnett, 2007). If physically active people choose to move to areas with more parks, for example, then it is the composition of the area and not the qualities of the physical environment that lead to higher levels of physical activity. One approach to addressing this issue is to collect data on the self-reported reasons people give for having moved to their current location, although adjustments for self-selection may show only minor changes in the relationships between place and health (Sallis et al., 2009).

Self-selection would most likely be a factor in situations where people have high mobility and everyone who wants to move to a particular housing unit that best matches their desired unit and neighborhood characteristics can do so. Research on residential mobility shows this is rarely the case, especially for people with limited incomes who experience housing discrimination. Even in a fast-growing region where government programs designed to broaden housing options for low-income households are in place, half of the low-income units ended up being highly clustered in areas of high poverty, concentrated minority populations, poor educational programs, and high crime (Van Zandt & Mhatre, 2009).

Longitudinal analysis investigating the relationships between health and place is needed. The relationships between health and place over time are difficult to uncover because some people move and "migration and health are jointly dependent events" (Boyle & Duke-Williams, 2004, p. 131). Some people may move because of their health, but the move may also positively or negatively affect health status over time.

Analysis of data for England and Wales revealed notable patterns in the health of migrants (Norman, Boyle, & Rees, 2004). Over a 20-year period, the largest absolute migration stream involved relatively healthy individuals moving away from more deprived areas to less deprived areas, raising rates of morbidity and mortality in the origin areas and lowering rates in the destination zones. A significant group of people in poor health relocated from less deprived to more deprived areas, but there was also a counterflow of unhealthy people who moved into less deprived areas. Migration, rather than changes in the characteristics of places that people who did not move lived in, accounted for most of the observed increase in health inequalities across areas over the time of the study.

Neighborhood Change

Neighborhood change, especially change in characteristics of the built environment, is difficult to measure over time. Many GIS databases describing built environment characteristics, especially at the local level, are developed by government agencies to describe current conditions. Reconstructing where sidewalks were or water and sewer lines were 10 years ago may be difficult if spatial databases describing earlier conditions have not been saved. Neighborhood history calendars have been proposed as a method using GPS and GIS for collecting and integrating event histories of neighborhood changes over time (Axinn, Barber, & Ghimire, 1997).

The level of deprivation in an area may change over time. A study of people in England and Wales who had not moved between 1971 and 1991 and who lived in households that were not deprived over the time period was conducted to investigate whether the change in deprivation in the area where they were living influenced health (Boyle, Norman, & Rees, 2004). People living in areas that become less deprived during the study period had lower standardized illness ratios, and changing conditions in the most deprived areas had a clear effect on morbidity.

Not only the characteristics of places but also the boundaries of areas used to assess effects on health may change. Here, too, political forces come into play. A study using GIS to analyze changes in municipal boundaries in small towns in the South due to annexation found that blacks living in areas adjacent to municipalities that might be candidates for annexation were systematically excluded from incorporation (Lichter, Parisi, Grice, & Taquino, 2007).

Location processes work to produce and reproduce living environments over time, often in ways that reinforce health disparities. GIS was used to digitize Booth's 1896 map of social class and aggregate the data to London's ward structure in the 1990s (Dorling, Mitchell, Shaw, Orford, & Smith, 2000). The index of poverty derived from Booth's data contributed more to explaining some health problems observed among neighborhood residents in the 20th century than the more recent data. Research using GIS to investigate health disparities is only beginning to address key issues of longitudinal effects and the pathways by which environmental conditions affect health.

Conclusion

Health disparities are an important public health issue. The differences among groups of people in different places are not trivial. The gap between the highest and lowest life expectancies observed among people grouped by race and county in the United States was more than 35 years in 2001 (Murray et al., 2006). In many settings, disparities in income, environmental quality, and health status are widening. Nevertheless, it is possible to address health disparities, and the techniques for investigating the connections between health and place have also been used to evaluate the success of interventions to promote health (Meersman, Breen, Pickle, Meissner, & Simon, 2009).

Measures of individual and neighborhood-level income and environmental quality are key variables associated with health disparities in many places. Research has revealed the importance of including both measures of economic power and environmental quality. Not all deprived areas have poor environmental quality and not all places with poor environmental quality have poor populations. Furthermore, the relationships between income, environmental quality, and health appear to be themselves spatially variable.

The links between location and health raise important public policy questions. Who has power to influence the location processes of land use change, trade, and migration that affect the neighborhoods where we live? Several studies discussed in this chapter considered the role that information and political participation play in health disparities. GIS are widely used by government agencies, private enterprises, and research scientists, but public access to spatial data and the tools needed to analyze and map them is often limited. The final chapter in our book discusses public participation GIS and its role in public health.

CHAPTER 12

Public Participation GIS and Community Health

The themes, concepts, and applications discussed in this book demonstrate how to develop a geographic foundation for the study of health problems using GIS. The examples in the preceding chapters illustrate the different kinds of spatial data and spatial methods available for health-related GIS analyses. Many of the examples are drawn from the research of health geographers, epidemiologists, and medical and public health professionals working in universities or federal and state agencies. These settings represent particular institutional contexts for GIS implementation that in turn influence the outcomes of GIS analyses. Some critiques of GIS emphasize the potentially harmful social consequences of the diffusion of GIS technology, including reinforcing the power of state agencies, facilitating surveillance, and promoting an at-best naive, technocratic view of social problems (Pickles, 1995b; Sheppard, 2005; Schuurman & Pratt, 2002). At the same time, the development of GIS and the hardware, software, databases, and networking systems they rely on, coupled with the expansion and development of the World Wide Web, have given the general public greater access to health and environmental information and the ability to visualize and analyze that information in new and innovative ways.

This chapter explores the role of GIS in community-based efforts to improve health and well-being. To provide a framework for understanding the role of GIS in community organizations, we begin by examining the links between GIS and society and the ways that GIS can be "structured" by diverse organizations and stakeholders. We then consider the concepts that underpin the development of community-based, "participatory" or "public participation" GIS (PPGIS) and how these GIS differ from the systems implemented in other institutional contexts. We examine PPGIS as a means of increasing community participation in public health planning, the kinds of data and technologies used in PPGIS, and the advantages and challenges of various strategies for implementing PPGIS including university–community group partnerships.

GIS and Society

The previous chapters in this book highlight the dynamic nature of GIS technologies and the wide range of application areas. This suggests that GIS is not a fixed or unchanging technology but rather one that both emerges from and is embedded in wider social contexts. Investigating the interplay between GIS and society—how society influences GIS and how GIS technologies influence societies—has emerged as an important theme in GIS research since the mid-1990s (Pickles, 1995). Referred to as "GIS and society" (Sheppard, Couclelis, Graham, Harrington, & Onsrud, 1999), "critical GIS" (Sheppard, 2005) or "heterodox GIS" (St. Martin & Wing, 2007), this research takes a critical look at the use and design of GIS while pointing to new areas of innovation and development. Themes raised in these debates have important implications for the use of GIS by community groups and others interested in improving health and access to health care.

One of the key insights from the GIS and society literature is that GIS is a socially constructed technology, a social practice (Sieber, 2000). It represents a combination of hardware, software, institutions, and people who interact in specific geographical and historical settings and whose activities influence how GIS is created and used and for what purposes. Attributes of GIS technology play a role in how GIS applications are structured. GIS is clearly a digital technology that has emerged and expanded in the digital era. As discussed in Chapter 1, changes in GIS technologies have been closely linked to advances in data storage and computing. As a digital technology, GIS also privileges particular ways of understanding the world. GIS is rooted in Cartesian conceptions of space in which places, and the relationships among them, are referenced according to a two- or three-dimensional coordinate system (Sheppard et al., 1999). It is difficult to include and model nonspatial (non-Cartesian) relationships—for example, those based on social or political networks—in GIS, without reference to locations on the earth. In general, quantitative approaches are privileged in GIS, although qualitative data and methodologies are increasingly being incorporated (Matthews, Detwiler & Burton, 2005; Cope & Elwood, 2009).

Within the broad technological constraints of GIS, users construct the technology in different ways that reflect their own goals, opportunities, and constraints in particular contexts. In planning a GIS, an organization makes a series of decisions about which GIS platform to use, the types of data and spatial analytic tools to be included in the GIS, and so on (Table 12.1). One of the most important decisions concerns the spatial data to be included or excluded in a GIS. Which specific data layers are incorporated, at which geographical scales? As discussed in earlier chapters, scale limits the spatial variation that can be observed and mapped in a data set, thus limiting the kinds of conclusions that can be drawn. Data access is also important. Organizations decide who will have access to GIS data and create procedures for granting access.

The social construction of GIS also involves decisions about the kinds of analytic tools to be included in the GIS and descriptions and guidelines for their

TABLE 12.1. The Social Construction of GIS

Main element	Element	Considerations
GIS Design	Data	Types of data included and excluded Spatial and temporal scales of data Geographic extent of study area Metadata Data access restrictions and procedures
	Spatial analysis	Types of tools included and excluded Ease and clarity of use Tool descriptions and instructions for use Access to tools
	User interface	Ease and clarity of use Graphical features
	Platform	Web- versus computer-based
Actors		Designers of the GIS Users of the GIS People with access to GIS data, tools, and results Participants in GIS deliberations
Use		Organizations sponsoring and directing the GIS Goals of GIS use Types of activities supported by the GIS Impacts of GIS on communities

use. Although most commercial GIS come with a predefined analysis toolkit, organizations may want to develop customized tools for particular applications. Which tools are included or excluded in a GIS? There are also key decisions about whether or not to distribute GIS data and tools over the Internet, the design of the user interface, and the administration and updating of the system. Thus, the characteristics and operations of a health GIS will differ in different contexts.

Questions about the use of GIS—who uses the technology, how, and for what purposes—are also central to understanding the relationship between GIS and society. Like all technologies, GIS is supported and shaped by powerful economic and political interests (Pickles, 1995b). Examining the rapid expansion of GIS in the late 1980s and 1990s, critics have described the role of GIS in military operations and in resource extraction and geodemographic mapping activities by large corporations. The needs and goals of these powerful institutions are reflected, in part, in the development and rapid expansion of the GIS market. In the health arena, little is known about GIS use by large, health care firms. However, in the United States we know that health insurance companies, pharmaceutical companies, and health service providers control large, proprietary geospatial databases, and it is likely that they are using GIS to support their commercial activities.

Even within a particular type of GIS application, who uses the technology and how can vary (Table 12.1). Most public health applications of GIS involve a diverse array of stakeholders, including public health officials, GIS experts, community members, and health service providers. What role, if any, do these groups have in designing the GIS, proposing GIS queries, and viewing and discussing the outcomes of GIS? The technical complexity of GIS means that experts are often involved, but their roles can range from facilitating community or agency goals and queries to, on the other extreme, completely directing the GIS investigation with limited input from other stakeholders.

The role(s) of community groups and individual residents in public health GIS is a critically important dimension of the social construction of GIS. ***Top-down GIS*** are systems directed and implemented by planners and decision makers with limited community input. In contrast, ***bottom-up GIS*** have substantial community involvement in all aspects of GIS design and implementation, enabling residents to characterize their local neighborhoods and specify GIS queries (Talen, 2000). Historically, community involvement in health GIS has been relatively limited. Although residents and community groups often draw attention to perceived health issues in their neighborhoods, the subsequent GIS analysis of those issues is often conducted with limited community input. In many cases this is intentional: community input is viewed as potentially compromising the scientific validity of the GIS analysis (Neutra, Swan & Mack, 1992). Yet, in a few notable cases, community groups have implemented their own GIS and conducted their own GIS investigations, typically with the assistance of scientific experts (Brody et al., 2004). Thus, there are diverse models for GIS use and implementation ranging from top-down, expert-driven systems to bottom-up, community-based systems.

To illustrate various dimensions of the social construction of health GIS, consider the case of the GIS for Breast Cancer Studies on Long Island (LI-GIS) (National Cancer Institute, 2010b). As discussed in Chapter 4, the LI-GIS was established by the National Cancer Institute in response to a federal mandate (PL 103-43) calling for the creation of a "geographic system" to evaluate "environmental and other potential risk factors contributing to the incidence of breast cancer in Nassau and Suffolk counties in New York." Although delayed for several years, the system is now easily accessible on the Internet (National Cancer Institute, 2010b).

The LI-GIS has several notable features. First, it incorporates a wide range of social and environmental data layers including data on medical facilities, air and water quality, hazardous sites, and population characteristics. It also includes a diverse set of GIS-based tools, including tools for detecting spatial clusters. Some tools include an easy-to-follow description for the nonexpert user, while others assume scientific expertise. The system also contains links to numerous scientific studies on breast cancer and the environment, thus serving as an educational and a translational tool.

Although the LI-GIS contains diverse health and environmental data sets, accessing the data is complex, requiring submission of a research proposal that describes research goals and hypotheses to be tested. Obtaining geospatial data on breast cancer incidence involves a separate request to the New York State

Health Department. It is unclear if cancer data are available below the ZIP Code scale. In sum, this top-down system is set up to facilitate hypothesis-driven academic research as opposed to exploratory, community-based mapping and analysis. As might be expected in a system directed by a prominent federal health agency, scientific and biomedical approaches to understanding breast cancer are privileged. Thus, the LI-GIS's "social construction" contrasts sharply with that of breast cancer GIS developed earlier by community advocacy groups to explore environmental associations (Carlin, 2002; McLafferty, 2002).

Public Participation GIS

The expanding role of GIS in community-based health initiatives in places like Long Island is connected to the broader movement to develop public participation GIS. ***Public participation GIS (PPGIS)*** are systems that facilitate and enhance the participation of individuals and community groups around issues of local concern (Sheppard et al., 1999). PPGIS aim to bring GIS tools to community groups and residents and incorporate their concerns and knowledge in decision making. Such systems enable communities to explore and visualize local concerns and resources, to contribute their own local knowledge of important issues, and to participate in identifying effective solutions. A "tool for the empowerment of social movement groups" (Sieber, 2000, p. 775), PPGIS is a means for giving communities control over issues that affect them. Community groups, academic researchers, and public health agencies are increasingly embracing PPGIS in their efforts to improve health and access to health care in local communities. The expansion of PPGIS in public health reflects not only increasing interest in GIS use among community groups, but also increasing awareness of the need for community input in understanding health issues and planning public health interventions (Partridge & Fouad, 2010). The past decade has witnessed a resurgence of interest in community-based health research and planning, and the growth of PPGIS mirrors this trend.

One cornerstone of PPGIS is "to accommodate an equitable representation of diverse views" (Sheppard et al., 1999, p. 811). Historically, most GIS were developed for use by government agencies and private firms with relatively narrowly defined programmatic or commercial interests. Designed to facilitate spatial database management and to support decision making, the systems included powerful tools for managing the large spatial databases these organizations often rely on and for solving well-defined geographical problems, as described in earlier chapters of this book. In contrast, PPGIS are designed to accommodate a much more diverse set of participants and viewpoints. This implies enhancements to GIS design and functionality, as well as an understanding of the sociopolitical contexts of community participation.

Despite general agreement about the value of PPGIS, its features are difficult to pin down. Sieber (2006, p. 292) describes PPGIS as a "co-produced concept composed of multiple disciplinary approaches and actors, rapidly changing technologies, and numerous as well as occasionally transgressive goals.” This diversity, however, encompasses certain guiding principles, as shown in Table 12.2 (Aberley & Sieber, 2002). These principles relate to the overall goals of PPGIS and the contexts in which it is used; access and participation; PPGIS design including platform, data and tools; and PPGIS deliberations and outcomes.

To provide an overarching framework, we can view PPGIS as a set of interconnected processes aimed at enhancing community participation and empowerment (Figure 12.1). These processes are clustered in three phases (Jankowski & Nyerges, 2003). The ***convening phase*** comprises processes that bring together stakeholders in a PPGIS or inhibit their participation. Social, cultural, and institutional influences on participation are important in this phase, as are the planning, designing, and scheduling of the PPGIS. The ***deliberation phase*** encompasses activities, discussions, and debates among participants and facilitators during PPGIS sessions. Processes that influence interactions between people and groups, and the role of GIS in shaping these processes, are key components of the deliberation phase. Finally, the ***outcomes phase*** focuses on the results of the PPGIS and on how outcomes translate into community improvements and

TABLE 12.2. Selected Principles of Public Participation GIS

Is an interdisciplinary research, community development, and environmental stewardship tool.
Is practiced in diverse contexts relating to place, organizations, government, or sector.
Endeavors to involve people and groups traditionally marginalized from decision-making process.
Can be used to help solve problems in particular sectors of society or provide broader regional assessments of place-based or bioregional identity.
Is best applied via partnerships between individuals, community organizations, academic institutions, social or religious organizations, governments, and the private sector.
Is linked to social theories and methods.
Is linked to applied qualitative research tools, including participatory action research.
Is applied using a wide variety of data formats, qualitative and quantitative.
Enables public access to cultural, economic, and biophysical data generated by public, voluntary, and private institutions.
Supports a range of interactive approaches from face-to-face contacts and web-based applications.
Promotes development of software that is accessible to broad acquisition and ease of use.
Supports lifelong learning of its practitioners in a manner that helps bridge divides between cultures, academic disciplines, gender, and class.
Is about sharing the challenged and opportunities of place and situation in a transparent and celebratory manner.

Note. Adapted from Aberley and Sieber (2002). Reprinted by permission.

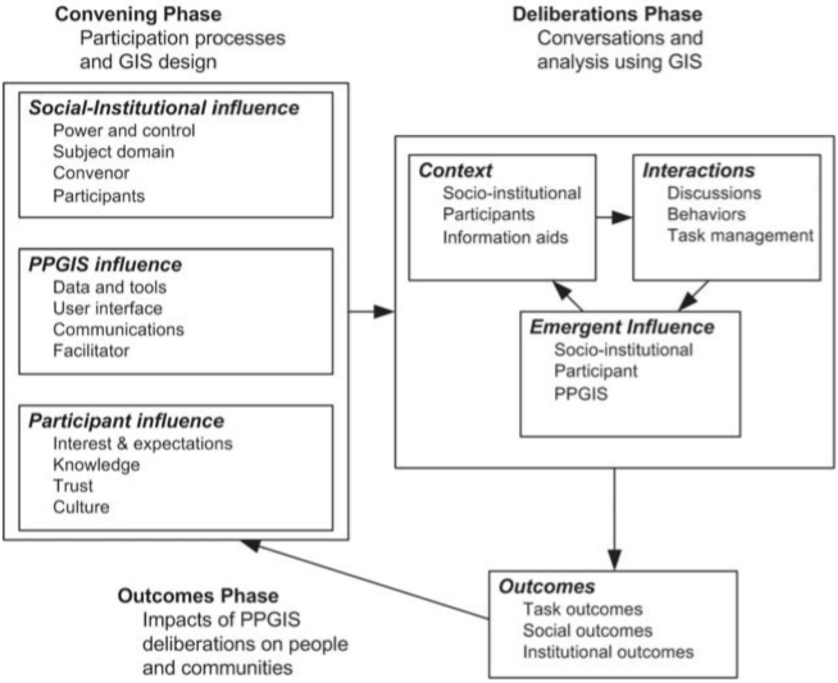


FIGURE 12.1. A PPGIS comprises three interconnected processes: the convening phase, the deliberations phase, and the outcomes phase. Adapted from Jankowski & Nyerges (2003). Copyright 2003 by P. Jankowski.

empowerment. There is also an important feedback from the outcomes phase to the convening phase: outcomes in turn affect community resources and organizations as well as the everyday lives of participants, thus influencing the social, political, and geographical contexts for implementing new PPGIS. According to Jankowski and Nyerges (2003), these interlinked processes depict technology, people, and institutions as mutually dependent, continually shaping and reshaping each other—a perspective known as ***adaptive structuration***.

PPGIS can be structured and implemented in a number of different ways. In one model, the PPGIS is created by public or nonprofit organizations or academic researchers who implement the PPGIS at meetings with community participants. Participants respond to GIS maps, but they have little or no say in how the GIS is constructed. A second type of PPGIS involves partnerships between community organizations and GIS experts in academia or government. Although community participants make key decisions, implementing those decisions falls on GIS facilitators, or "chauffeurs" who provide technical support (Nyerges, Jankowski, Tuthill, & Ramsey, 2006). These are termed ***chauffeur-driven PPGIS***. GIS partners implement and operate the PPGIS, but with substantial input and direction from community participants. The third model involves creating a stand-alone

GIS for a community organization. The organization has full control over the GIS and can use and modify the system to suit its needs. Typically, education and training are provided to community group participants so that after an initial learning period, the group can sustain the GIS without relying on experts. This model clearly requires a significant and continuing commitment of staff, time, and resources by a community organization (Ghose, 2001).

The following sections discuss some of the core elements of PPGIS and their relevance for community health applications of PPGIS. The core elements are grouped into three categories, loosely following Sieber (2006): participation, data and technology, and implementation and outcomes.

Participation

A key element of PPGIS is community involvement and participation. Participation can take on different forms as depicted in the ***ladder of participation*** (Figure 12.2). Arnstein (1969) described eight levels of participation ranging from nonparticipation to full citizen control. The lowest rungs of the ladder comprise top-down strategies in which powerful decision makers coopt or coerce citizens to agree to particular policies. In contrast, the highest rungs are associated with direct citizen control over policy decisions. Ignoring the nonparticipation sce-

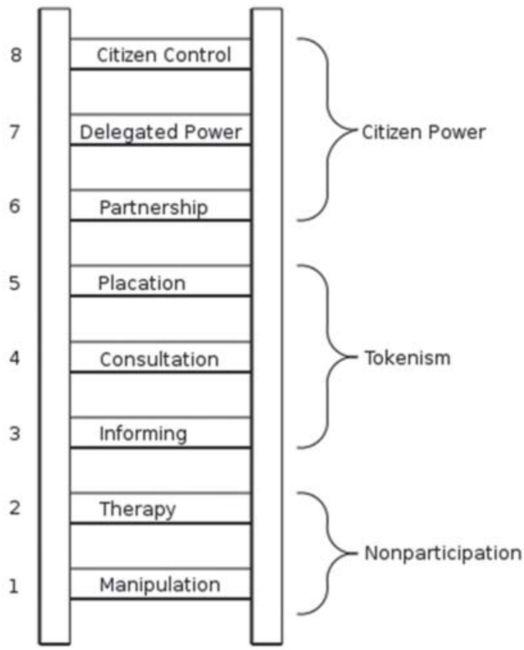


FIGURE 12.2. Arnstein's ladder of citizen participation. From Arnstein (1969). Copyright 1969 by the Tyalor & Francis Group, LLC. Reprinted by permission.

narios that contradict the very definition of PPGIS, the forms of participation displayed in Figure 12.2 suggest a range of scenarios for participatory GIS.

On the one hand, PPGIS can be a tool for informing residents about health concerns in their communities and providing education about health promotion and prevention. The informing and consultation forms of participation are emphasized. Many public health applications of PPGIS are structured this way. To engage diverse stakeholders in developing a cancer control strategy for colorectal cancer in Iowa, Beyer and Rushton (2009) created spatially smoothed maps (see Chapter 5) of cancer incidence, late-stage risk and mortality in Iowa. The maps were used in a community interaction session in which residents responded to the maps and related observed patterns of cancer incidence to their own knowledge of the area. Information gleaned from the community responses was used in targeting and planning interventions to improve cancer awareness and screening. This case study provides a good example of a PPGIS that informs and engages community residents.

On the other hand, PPGIS can be structured to promote higher levels of participation and empowerment by directly involving community groups and residents in data acquisition, mapping, analysis, and decision making (Ghose, 2001). In such systems, community stakeholders are active participants in constructing and using the GIS for their own purposes. Because of the technical expertise required in preparing GIS databases and operating GIS software, many of these PPGIS are chauffeur-driven. A PPGIS created for the West Islip, New York Breast Cancer Coalition in the mid-1990s illustrates the chauffeur-driven concept (Timander & McLafferty, 1998). Although the PPGIS was created and managed by GIS researchers, coalition members contributed data and posed community queries for investigation. For example, they asked: "Are breast cancer cases clustered among people who live at the ends of water mains?" The GIS team analyzed this query using GIS and presented the findings to the coalition.

Who participates is also a key dimension of PPGIS (Dunn, 2007). Incorporating diverse viewpoints and giving marginalized groups a voice are hallmarks of PPGIS, but how can these be accomplished? Participation is the result of a conscious decision by an individual or group to engage with a PPGIS. For individual residents, many ***barriers to participation*** exist. Willingness and ability to participate are influenced by the person's interest in the issue, as well as economic, social, and cultural factors. Time–space constraints and language and economic barriers may limit participation, particularly among individuals already marginalized by class, race or ethnicity. Research indicates that willingness to be involved in participatory decision making is often lacking, and PPGIS deliberations are frequently dominated by a vocal minority (Carver, 2003).

Effective and equitable participation also depends on the definition of the "community" of interest. Communities can be geographic, economic, occupational, social, or political (Schlossberg & Shuford, 2005). They are defined in relation to the health issue of interest, and they change over time in response to changing political and economic conditions and policy implementations. In many PPGIS, the community of interest is defined as a geopolitical unit—a state, county, or town. However, such places encompass diverse populations and environments, and health concerns spill across geopolitical boundaries. Therefore how the community is framed can affect the outcomes of PPGIS deliberations (Schlossberg & Shuford, 2005). Identifying the community of interest is a critical first step in developing PPGIS.

Some PPGIS, such as the West Islip system, emerge at the request of a particular community group. In these cases, participation in the PPGIS is completely tied to participation in the group; some populations and voices may be excluded. Moreover, for community groups there are also substantial economic, technological, and social barriers to adoption and use of GIS. Lack of knowledge and technological expertise, lack of computer resources, and day-to-day decisions about GIS implementation can limit groups' involvement in PPGIS (Elwood, 2006a).

PPGIS emerge in particular geographical and historical contexts that influence who participates and how. The health topic of concern and people and places of interest play a key role in shaping participation. At the same time, participation is the result of individual and community group decisions that reflect interest in and willingness and ability to get involved in PPGIS deliberations. Social, economic, and institutional factors constrain and/or enhance participation differently in different contexts. To be effective, a PPGIS must be positioned within community social networks in a way that gives voice to marginalized and underrepresented community interests. The design and implementation of the PPGIS are also important in efforts to promote participation.

Data and Technology

LOCAL KNOWLEDGE AND PPGIS DATA

A key element of PPGIS is community involvement in the creation, evaluation, and analysis of spatial data. Local knowledge of environmental or health problems can be an important source of information for PPGIS, providing a more detailed, grounded summary of local conditions than can be gleaned from secondary data (Harris, Weiner, Warner, & Levin, 1995). ***Local knowledge*** refers to contextual knowledge that people gain through everyday activities and experiences. It includes information about, perceptions of, and attitudes toward specific places and concerns. Local knowledge can be incorporated in PPGIS in several ways. One option is to administer a questionnaire to elicit local knowledge that is then incorporated as data layers or features in the GIS. A PPGIS from Hamilton, Ontario, discussed later in this chapter, illustrates this strategy (Maclachlan, Jerrett, Abernathy, Sears, & Bunch, 2007). A second option is to acquire local knowledge electronically. ***Volunteered geographic information***—reports submitted electronically by residents and other stakeholders—is increasingly important in PPGIS (Elwood, 2008). During Hurricane Katrina, reports of flooding from people on the ground were used in targeting rescue and relief efforts. Nuisance complaints from residents concerning, for example, noise, rubbish and rats, have long been important in public health planning. Now they can be submitted electronically and represented seamlessly on GIS maps. The maps are emergent, actively constructed through residents' participation.

A third source of local knowledge is PPGIS deliberations and discussions. Viewing GIS maps and images can encourage people to express local knowledge and pose hypotheses about the causes and consequences of health concerns. PPGIS can serve as a tool for eliciting local knowledge as people respond to maps and images of their local areas (Elwood, 2006b). In Long Island, GIS maps of breast cancer triggered hypotheses among residents about potential links between environmental contamination and cancer incidence (McLafferty, 2002).

Beyer, Comstock, and Seagren (2010) developed a participatory GIS application to ascertain local knowledge about colorectal cancer risk in a northwest Iowa community where colorectal cancer incidence, mortality, and late-stage risk were high. Focus groups and interviews were conducted with diverse community residents. Activities included a mapping exercise where participants viewed Google Earth® maps of the study area and were asked to identify features perceived as "positive" and "negative" on the map. Participants were also asked to provide descriptions about why they felt the feature was relevant for cancer risk. Through PPGIS-based conversations, participants generated hypotheses about the causes of high colorectal cancer risk and suggested policies and approaches for reducing risk.

Local knowledge is often in the form of narratives, photos, sketch maps, or audio/video clips that may be linked to and integrated with foundation spatial data (Cope & Elwood, 2009). Such data represent how a place looks and how residents' perceive and experience the place. Although these are not traditional GIS data formats, many GIS are now able to handle such diverse types of spatial information. Known as ***qualitative GIS***, these systems enable users to view and explore the multidimensional and perceptual characteristics of a place. Photos and audio and video clips can be tagged to their corresponding locations enabling users to view them as a part of GIS-based data exploration. In a study of community gardens in a Buffalo, New York, neighborhood, Knigge and Cope (2006) developed a qualitative GIS that incorporated photographs, field notes, and oral history recordings (Figure 12.3). The GIS served to represent the sociospatial contexts of community gardens and the meanings associated with them. The fusion of qualitative and quantitative information is important, because it enables nonexpert, community-based users to link traditional map-based GIS data with more familiar images and sensory information about their communities.

Local knowledge also includes participants' attitudes and preferences about the characteristics of public health policy options. As in the spatial decision support systems described in Chapter 10, preference data can be used in PPGIS to

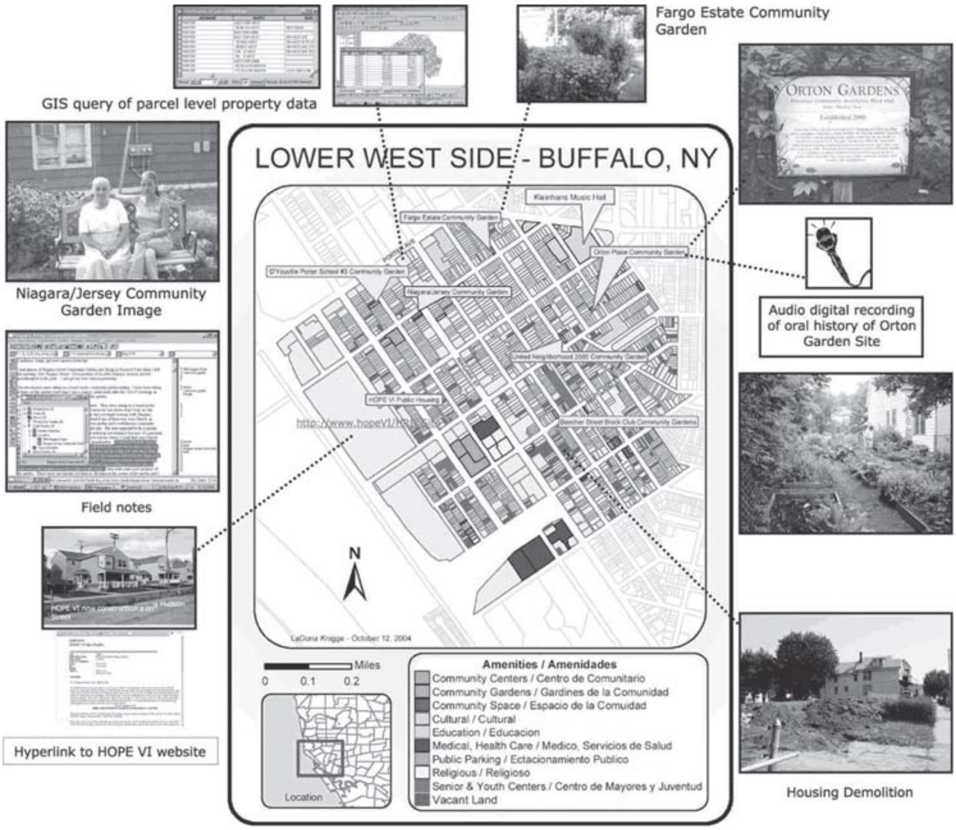


FIGURE 12.3. A qualitative GIS representing residents' local knowledge of the Lower West Side neighborhood in Buffalo, New York. The GIS combines diverse types of data, including photographs, field notes, audio recordings, and hyperlinked information. From Knigge and Cope (2006). Copyright 2006 by Pion Ltd. Reprinted by permission of LaDonna Knigge.

generate policy options that reflect the attitudes and beliefs of participants. In Bangladesh, a PPGIS was implemented to involve local communities in planning arsenic mitigation strategies (Hassan, 2005). Arsenic, a poison and known carcinogen, has been found in high concentrations in groundwater in some areas of rural Bangladesh. Impacts on human health arise when people consume contaminated groundwater, and many cases of arsenic poisoning have been reported. One way to combat this problem is to construct deep tubewells that tap drinking water below the zone of contamination. Hassan (2005) used PPGIS to integrate community residents' local knowledge in identifying locations for new deep tubewells in the study region. Much of the knowledge concerned the choice of a threshold distance—the maximum distance people were willing and able to

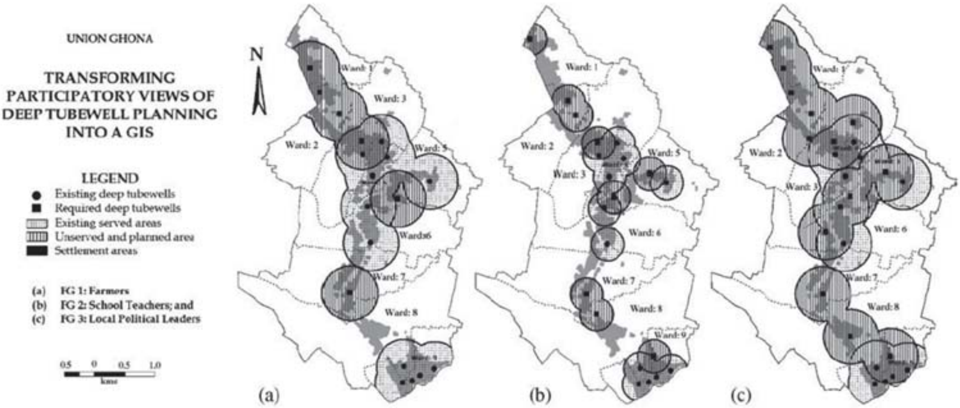


FIGURE 12.4. Participatory location plans for deep tubewells developed by three stakeholder groups: (a) farmers, (b) school teachers, and (c) local political officials. From Hassan (2005). Copyright 2005 by Elsevier. Reprinted by permission.

travel to obtain uncontaminated drinking water. Focus groups with three stakeholder groups—farmers, teachers, and local political leaders—revealed different opinions about the maximum travel distance resulting in different spatial plans for locating new deep tubewells (Figure 12.4).

PPGIS DESIGN AND TECHNOLOGY

In PPGIS, participant involvement also extends into analysis and interpretation of spatial data. This requires systems that permit multiple users and diverse forms of querying and that facilitate communication among users in the analysis process. With its emphasis on community empowerment, PPGIS needs to be designed for the nonexpert user; yet most commercial GIS require considerable knowledge and expertise. Such GIS "present major obstacles to the non-expert user in terms of navigating a language, worldview and interface that supports the system's architecture rather than the user's work view" (Haklay & Tobon, 2003, p. 577). Drawing on knowledge about human–computer interactions and tests with community participants, researchers have studied design factors that limit the usability of PPGIS for nonexpert users (Haklay & Tobon, 2003). Difficulties in navigating maps and understanding technical terms and symbology limited nonexperts' interactions with PPGIS.

Increasingly, the Internet is being used as a platform for PPGIS development. The Internet offers several advantages over traditional freestanding GIS for enhancing citizen involvement in PPGIS: It enables asynchronous communication, freeing participants from space–time constraints; it provides an effective tool for acquiring data and local knowledge from participants; and it supports new kinds of dynamic geovisualization and analysis (Nuojua, 2010).

To understand residents' perceptions and experiences of environmental health hazards in Hamilton, Ontario, Maclachlan and others (2007) developed a participatory Internet-based GIS. The GIS included standard environmental data on environmental hazards and pollution levels, but also incorporated information from a phone survey in which residents reported respiratory symptoms, asthma, and environmental exposures. Data were compiled in a web GIS that included tools for data exploration and mapping. For example, participants can zoom in to small areas and obtain local air pollution information and also perform multilayer queries. Researchers concluded that the web GIS was an effective and affordable means of providing communities with access to data and tools enabling them to conduct their own environmental investigations.

Web 2.0 holds great promise for the development of interactive, web-based PPGIS for public health that reach a broad audience. Its interactive, collaborative, and visualization capabilities make it an innovative platform for public participation (Nuojua, 2010). SafeRoadMaps is a web-based participatory GIS, developed in Web 2.0, that represents geographic variation in traffic safety across the United States (Hilton, Horan & Schooley, 2009). The system relies on GIS map-mashups to integrate diverse types of data including traffic fatality data, imagery, video clips, and news reports of traffic accidents. Users can view maps of traffic fatalities and create dynamically updated maps showing changes in public policies to improve traffic safety. In the future, users will be able to contribute volunteered geographic data such as photos, memorials, and biographies of those impacted by traffic accidents. Geocoded hyperlinks to congressional representatives enable users to e-mail elected officials about needed improvements in traffic safety. Thus, SafeRoadMaps harnesses Web 2.0 to create "a platform for mass collaboration" around issues of traffic safety (Hilton, Horan & Schooley, 2009). It demonstrates the potential for developing PPGIS-based networks of collaboration among geographically dispersed individuals and community groups to address health issues of mutual interest.

Implementation and Outcomes

PPGIS implementation involves decisions about who participates, how deliberations among participants take place, how the system is funded and administered, and how decisions are made. Many of these issues are closely tied to PPGIS design and participation, as discussed in preceding sections. This section examines PPGIS implementation in relation to policy outcomes and community impacts.

Because maintaining and operating GIS requires technical expertise, most PPGIS involve partnerships between community groups and academic researchers, students, or trained public health professionals (Ghose & Huxhold, 2001). Effective partnerships benefit all participants, providing teaching opportunities and grounded research applications for academic participants, and technical assistance and training for community groups. Stakeholders also include individuals and institutions with a direct interest in the process and outcome of the

GIS analysis—representatives of private industry, as well as representatives of the diverse groups and interests that exist in a particular community.

The very diversity of these interests, however, means that there are potential conflicts among the goals of the stakeholders involved in a partnership. The Community University Consortium for Regional Environmental Justice developed a partnership between community groups and Cornell University to document water and air quality conditions and potential hazards in neighborhoods (Larsen, 1999). The purpose of the partnership was to create a sustainable GIS website to provide demographic, commercial, and up-to-date environmental information. The maps and information produced by the GIS would assist community groups in their organizing and advocacy efforts. Because of the different institutional settings and priorities among program participants, one conclusion drawn from the project was that "the process and purpose of academic involvement in community organizations can be problematic" (Larsen, 1999, p. 147).

Questions have also been raised about the sustainability of PPGIS implemented through university-community partnerships. Such partnerships often revolve around key individuals. Involvement of academic participants is constrained by teaching, research, and service commitments that vary throughout the year (Ghose & Huxhold, 2001). Involvement of interested community participants is similarly constrained by the demands of work and family commitments. Many PPGIS partnerships diminish over time as key individuals turn to other important obligations. If the PPGIS were focused on a particular community issue, this might be appropriate. If, however, the PPGIS is intended to provide an ongoing information resource for the community, then planning for how the PPGIS will be sustained is necessary in the early stages of its development.

Concerns about participation also arise in PPGIS implemented by public health agencies to inform and educate community residents and elicit their local knowledge. Typically GIS maps and queries are shared with community participants in chauffeur-driven PPGIS sessions in which participants discuss and query maps, describe hypotheses, and comment on needed public health interventions. Because community concerns and local knowledge are expressed by participants in the session, the processes affecting participation are critically important. PPGIS sessions often attract few participants (Carver, 2003). Local politics and power relations can create formidable barriers to effective and equitable participation at PPGIS sessions. GIS exist in webs of social and political relations, and those relations shape GIS outcomes just as they affect other types of local decision making. Most communities are heterogeneous, consisting of diverse groups with differing needs, interests, and levels of political and economic clout. Little is known about the performance and outcomes of PPGIS in diverse communities and the varying capacities of diverse groups to ensure that their health interests and needs are addressed (Sheppard et al., 1999). These are very real concerns for groups disadvantaged on the basis of class, race, disability, or some other dimension of social exclusion. Without full and democratic participation, PPGIS may become an additional force of marginalization for these groups, rather than a progressive source of empowerment.

Participation is also important during PPGIS deliberations. Effective deliberations are those in which knowledge is freely and openly shared. In some contexts, participants may be reluctant to express their opinions in the presence of more powerful participants. Case studies show that PPGIS conversations are often dominated by a small number of participants (Carver, 2003). Differences in participation based on gender, race, age, class, and culture mean that some voices are less likely to be expressed than others. A well-trained facilitator can mitigate some of these issues by strategically soliciting input from nonactive participants or initiating small-group discussions, but barriers often persist.

Web-based PPGIS overcome some of these problems by diminishing social, spatial, and temporal barriers to participation. However, these systems pose a new set of challenges. Anyone, anywhere, can participate, often anonymously, so it is difficult to know whose opinions are being expressed (Dunn, 2007). The digital divide, though shrinking, continues to constrain participation in web-based initiatives. Lack of interest, knowledge, and expertise also limit web-based PPGIS use by nonexperts, and impacts on web-based collaborations and outcomes are poorly understood. In the extreme, web-based PPGIS can easily be coopted by powerful interests who stand to benefit (or lose) from the resulting decisions.

In evaluating PPGIS implementation, it is critically important to focus on outcomes: Do the systems foster improvements in community health and more effective public health policies? The research literature provides few examples of the impacts of PPGIS deliberations and planning activities on community health, because few published studies cover a sufficient time period for assessing health impacts. Case studies that track the health impacts associated with PPGIS-based public health interventions are sorely needed.

A study conducted in Flint, Michigan, illustrates the potential for linking PPGIS, public health interventions, and improvements in health outcomes. GIS and participatory research methods were used to investigate spatial and social inequalities in diabetes prevalence in Flint and to design interventions to reduce the disease burden (Kruger, Brady, & Shirey, 2008). A survey developed collaboratively between researchers and community partners was conducted to estimate diabetes risk and screening rates. Maps of the survey data revealed areas where estimated risk was high and use of screening was low. Additional focus groups and needs assessment surveys were conducted in these high-need areas to uncover behavioral and contextual risk factors, and local interventions were implemented. Data collected after the interventions showed that diabetes screening rates increased in the areas targeted for intervention.

Although PPGIS facilitate community participation in identifying and analyzing local health problems, implementing PPGIS does not guarantee that those problems will be addressed. Achieving positive health benefits requires that local knowledge and insights gained through PPGIS be translated into effective public health policies. This translation process is rarely easy or straightforward. It entails actions on the part of public health organizations and individuals that fall beyond the scope of PPGIS implementation. Moreover, many types of health problems have multiple, interrelated, and unknown causes for which there is no obvious public health solution. GIS analysis serves a useful purpose by identifying the health problem of interest and tracing its geographical distribution, but the analysis rings hollow when there are no prospects for improvement. As noted earlier, GIS maps can also lead to stigmatization or redlining, to the detriment of local residents. When health problems have solutions, the outcome of PPGIS will depend on whether the GIS analysis is tied to an effective public health response.

Conclusion

This book has demonstrated the use of GIS in creating a geographic foundation for the study of human health problems and planning effective public health interventions. GIS is much more than a tool for making maps. It involves a complex set of practices ranging from data collection and analysis to implementation and community partnerships. The design of a public health GIS raises a greater range of issues than conventional, nonspatial health data collection and analysis and involves a greater number of actors. Developing meaningful GIS applications for public health requires an understanding of geographic data and methods and the social and political contexts in which they are applied.

The geographic foundation for public health advanced in this book emphasizes the importance of "where." Where are particular health issues most concentrated? What health problems occur in particular places? How and where do diseases emerge, and how do they spread through space and time? How do social and environmental conditions in a place affect the health of local populations? How does the uneven location of health care services affect health and well-being? How can local environments and health care services be improved to enhance public health? We have shown that the geospatial data, methods, and technologies that comprise GIS have a key role in addressing these essential questions.

Wide inequalities in population health persist among countries, regions, and neighborhoods, despite advances in biomedical technologies and increased availability of health services. These health inequalities are continually shaped and reshaped by environmental, social, and economic transformations. From environmental health to infectious diseases to the delivery of health care, the diverse applications discussed in this book show that GIS can assist in providing better understandings of health inequalities and teasing out their relationships to social and environmental change.

At the same time, better understandings of health inequalities ring hollow if they are not accompanied by effective interventions to improve population health. The geographical foundation for public health advanced in this book extends to the public health policy arena where, we have argued, GIS-enabled data, visualizations, models, and interactions are key to policy formulation. Geographically targeting public health interventions to the places and populations of greatest need, geographically tailoring policies to reflect local environments, populations, and cultures, and geographically generating policies based on interactions with local stakeholders will promote policy interventions that achieve more effective and long-lasting improvements in health.

GIS applications in public health require institutional settings that foster the conceptual understanding of geographic data and methods. The lack of this understanding is at least as important a barrier to the diffusion of GIS in public health as the technological issues raised by GIS implementation. Collaborative efforts between practitioners in public health agencies, researchers based in institutions of higher education, and community groups need to address both the conceptual and the technological issues in GIS implementation, despite the different primary interests of participants. Researchers and public health practitioners who are concerned about GIS as a surveillance technology, who see manipulation of service delivery systems by for-profit providers as decreasing access to care, and who decry the absence of interventions to address some of the world's most pressing health problems can make a contribution by engaging with the forces that are shaping GIS applications in public health.

We hope that we have conveyed the amazing breadth of public health applications in GIS. While we have focused somewhat narrowly on the causes of ill health and the effective delivery of services to address health problems in most of the examples discussed in this book, the material in this concluding chapter suggests that GIS will continue to play an important and expanding role in our society. "Public health officials have the responsibility to continue to ensure that the promise of this wonderful new tool is realized" (Melnick & Fleming, 1999, p. x).

References

Aangeenbrug, R. T. (1997). Introduction. In R. T. Aangeenbrug, P. E. Leaverton, T. J. Mason, & G. A. Tobin (Eds.), *Proceedings of the International Symposium on Computer Mapping in Epidemiology and Environmental Health* (pp. iii–iv). Alexandria, VA: World Computer Graphics Foundation.

Abate, F. R. (Ed.). (1991). *OMNI gazetteer of the United States of America*. Detroit: Omnigraphics.

AbdelMalick, P., Boulos, M. N. K., & Jones, R. (2008). The perceived impact of location privacy: A web-based survey of public health perspectives and requirements in the UK and Canada. *BMC Public Health, 8*, 156.

Abel, T. D. (2007). Skewed riskscapes and environmental injustice: A case study of metropolitan St. Louis. *Environmental Management, 42*(2), 232–248.

Aberley, D., & Sieber, R. (2002). Public participation GIS (PPGIS) guiding principles. First International PPGIS Conference, URISA, held at Rutgers University, New Brunswick, NJ, July 20–22. Retrieved January 5, 2011, from *www.iapad.org/ppgis_principles.htm*.

Achabal, D. D., Moellering, H., Osleeb, J. P., & Swain, R. W. (1978). Designing and evaluating a health care delivery system through the use of interactive computer graphics. *Social Science and Medicine, 12*(1D), 1–6.

Aday, L., & Anderson, R. (1981). Equity of access to medical care: A conceptual and empirical overview, *Medical Care, 19*(suppl), 4–27.

African Geo Information Research Network. (2011). National mapping agencies in Africa. Retrieved January 5, 2011, from *nma.agirn.org/index.php/site/index*.

Agency for Healthcare Research and Quality. (2005). Emergency Preparedness Resource Inventory (EPRI). Retrieved January 5, 2011, from *www.ahrq.gov/research/epri*.

Agyeman, J., & Evans, B. (2004). "Just sustainability": The emerging discourse on environmental justice in Britain? *Geographical Journal, 170*(2), 155–164.

Alavanja, M. C. R., Hoppin, J. A., & Kamel, F. (2004). Health effects of chronic pesticide exposure: Cancer and neurotoxicity. *Annual Review of Public Health, 25*, 155–197.

Aldstadt, J. (2007). *Spatial and spatiotemporal analysis of dengue virus transmission and* Aedes aegypti *abundance (Thailand, Puerto Rico)*. Unpublished Doctoral Dissertation. Department of Geography, San Diego State University and the University of California at Santa Barbara.

Aldstadt, J., & Getis, A. (2006). Using AMOEBA to create a spatial matrix and identify spatial clusters. *Geographical Analysis, 38*(4), 327–343.

Ali, M., Emch, M., Yunus, M., & Clemens, J. (2009). Modeling spatial heterogeneity of disease risk and evaluation of the impact of vaccination. *Vaccine, 27*(28), 3724–3729.

Alter, C. F. (1988). The changing structure of elderly service delivery systems. *The Gerontologist, 28*(1), 91–98.

Amerasinghe, F. P., Breisch, N. L., Azad, A. F., Gimpel, W. F., Greco, M., Neidhardt, K., et al. (1992). Distribution, density, and Lyme disease spirochete infection in *Ixodes dammini* (Acari: Ixodidae) on white-tailed deer in Maryland. *Journal of Medical Entomology, 29*(1), 54–61.

American Geographical Society. (1944). A proposed atlas of diseases. *Geographical Review, 34*(4), 642–652.

American Hospital Association. (2009). *AHA Guide® 2010 Edition.* Chicago: American Hospital Association.

American Medical Association. (2009). *AMA directory of physicians in the United States, 41 edition*. Chicago: American Medical Association.

American Medical Association. (2011). AMA Physician Masterfile. Retrieved January 5, 2011, from *www.ama-assn.org/ama/pub/about-ama/physician-data-resources/physician-masterfile.shtml*.

American Public Health Association. (1996). Final program, American Public Health Association 124th annual meeting and exposition. Washington, DC: Author.

Anderson, J. F., Andreadis, T. G., Vossbrinck, C. R., Tirrell, S., Wakem, E. M., French, R. A., et al. (1999). Isolation of West Nile virus from mosquitoes, crows, and a Cooper's hawk in Connecticut. *Science, 286*(5448), 2331–2333.

Andreassen, H. K., Bujnowska-Fedak, M. M., Chronaki, C. E., Dumitru, R. C., Pudule, I., Santana, S., et al. (2007). European citizens' use of E-health services: A study of seven countries. *BMC Public Health, 7,* 53.

Angerer, J., Ewers, U., & Wilhelm, M. (2007). Human biomonitoring: State of the art. *International Journal of Hygiene and Environmental Health, 210*(3–4), 201–228.

Anscombe, F. J. (1973). Graphs in statistical analysis. *American Statistician, 27*(1), 17–21.

Anselin, L. (1995). Local indicators of spatial autocorrelation: LISA. *Geographical Analysis, 27*(2), 93–115.

Anselin, L. (1999). Interactive techniques and exploratory spatial data analysis. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems: Vol. 1. Principles and technical issues* (2nd ed., pp. 253–266). New York: Wiley.

Anselin, L. (2003a). *GeoDa(tm) 0.9 User's Guide.* Retrieved January 5, 2011, from *geodacenter.org/downloads/pdfs/geoda093.pdf.*

Anselin, L. (2003b). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review, 26*(2), 153–166.

Antenucci, J. C., Brown, K. Croswell, P. L., Kevany, M. J., & Archer, H. (1991). *Geographic information systems: A guide to the technology.* New York: Van Nostrand Reinhold.

Apelberg, B. J., Buckley, T. J., & White, R. H. (2005). Socioeconomnic and racial disparities in cancer risk from air toxics in Maryland. *Environmental Health Perspectives, 113*(6), 693–699.

Apparicio, P., Abdelmajid, M., Riva, M., & Shearmur, R. (2007). Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. *International Journal of Health Geographics, 7*, 7.

Apparicio, P., & Seguin, A-M. (2006). Measuring the accessibility of services and facilities for residents of public housing in Montreal. *Urban Studies, 43*(1), 187–211.

Arcury, T. Preisser, J., Gesler, W., & Powers, J. (2005). Access to transportation and health care utilization in a rural region. *Journal of Rural Health, 21*(1), 31–38.

Ardagna, C. A., Cremonini, M., & Gianini, G. (2009). Landscape-aware location-privacy protection in location-based services. *Journal of Systems Architecture, 55*(4), 243–254.

Arlinghaus, S. L. (Ed.). (1994). *Practical handbook of digital mapping: Terms and concepts*. Boca Raton, FL: CRC Press.

Armstrong, C. J. (Ed.). (1995). *World databases in geography and geology*. London: Bowker Saur.

Armstrong, M. P., Rushton, G., & Zimmerman, D. L. (1999). Geographically masking health data to preserve confidentiality. *Statistics in Medicine, 18*(5), 497–525.

Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of Planners, 35*(4), 216–224.

Aromaa, A., Koponen, P., Tafforeau, J., Vermeire, C., & the HIS/HES Core Group. (2003). Evaluation of health interview surveys and national health examination surveys in the European Union. *European Journal of Public Health, 13*(3, Suppl.), 67–72.

Ash, M., & Fetter, T. R. (2004). Who lives on the wrong side of the environmental tracks? Evidence from the EPA's Risk-Screening Environmental Indicators model. *Social Science Quarterly, 85*(2), 441–462.

Ash, M., & Robinson, D. E. (2009). Inequality, race, and mortality in U.S. cities: A political and econometric review of Deaton and Lubotsky. *Social Science and Medicine, 68*(11), 1909–1913.

Ashact, Ltd., & Dagh Watson, Spa. (1989). *INVENT: An industrial waste prediction model.* Washington, DC: United Nations.

Atkinson, N. L., Saperstein, S. L., & Pleis, J. (2009). Using the Internet for health-related activities: Findings from a national probability sample. *Journal of Medical Internet Research, 11*(1), e4.

AvRuskin, G. A., Jacquez, G. M., Meliker, J. R., Slotnick, M. J., Kaufmann, A. M., & Nriagu, J. O. (2004). Visualization and exploratory analysis of epidemiologic data using a novel space time information system. *International Journal of Health Geographics, 3*, 26.

Axinn, W., Barber, J. S., & Ghimire, D. (1997). The neighborhood history calendar: A data collection method designed for dynamic multilevel modeling. *Sociological Methodology, 27*, 355–392.

Aye, D. D., & Archambault, G. V. 1997. *Cancer Incidence in Southington, CT 1968–91 in Relation to Emissions from Solvents Recovery Services of New England*, ATSDR U50/ATU199044. Hartford, CT: Connecticut Department of Public Health.

Bader, M. D. M., Ailshire, J. A., Morenoff, J. D., & House, J. S. (2010). Measurement of the local food environment: A comparison of existing data sources. *American Journal of Epidemiology, 171*(5), 609–617.

Baer, H., & Singer, M. (2009). Global warming and the political ecology of health: Emerging crises and systemic solutions. Walnut Creek, CA: Left Coast Press.

Bailey, T. C. (1994). A review of statistical spatial analysis in geographical information systems. In S. Fotheringham & P. Rogerson (Eds.), *Spatial analysis and GIS* (pp. 13–44). Bristol, PA: Taylor & Francis.

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*. Harlow, UK: Longman Scientific & Technical.

Balas, E. A., Jaffrey, F., Kuperman, G. J., Boren, S. A., Brown, G. D., Pinciroli, F., et al. (1997). Electronic communication with patients: Evaluation of distance medicine technology. *Journal of the American Medical Association, 278*(2), 152–159.

Barcus, H., & Hare, T. (2007). Healthcare utilization, deprivation and heart-related disease in Kentucky. *Southeastern Geographer, 47*(2), 202–221.

Bartholomay, R. C., Carter, J. M., Qi, S. L., Squillace, P. J., & Rowe, G. L. (2007). Summary of selected U.S. Geological Survey Data on Domestic Well Water Quality for the Centers for Disease Control's National Environmental Public Health Tracking Program: U.S. Geological Survey Scientific Investigations Report 2007–5213. Retrieved January 5, 2011, from *pubs.usgs.gov/sir/2007/5213*.

Bartuska, T. J. (2007). The built environment: Definition and scope. In W. R. McClure & T. J. Bartuska (Eds.), *The built environment: A collaborative inquiry into design and planning* (pp. 3–14). Hoboken, NJ: Wiley.

Bassett, M. T., & Krieger, N. (1986). Social class and black-white differences in breast cancer survival. *American Journal of Public Health, 76*(12), 1400–1403.

Bausch, P., & Bumgardner, J. (2006). *Flickr hacks: Tips & tools for sharing photos online*. Sebastopol, CA: O'Reilly.

Bazilchuk, N. (2006). TRI: Corroding its original intent? *Environmental Health Perspectives, 114*(7), A420–A423.

Beach, S. L. (1987). Ground water sampling strategies for a water resources geographic information system. In D. Brown & P. Gersmehl (Eds.), *File structure design and data specifications for water resources geographic information systems, Special Report No. 10* (pp. 14-1–14-22). St. Paul, MN: Water Resources Research Center, University of Minnesota.

Beck, L. R., Lobitz, B. M., & Wood, B. L. (2000). Remote sensing and human health: New sensors and new opportunities. *Emerging Infectious Diseases, 6*(3), 217–227.

Becker, K., Glass, G., Braithwaite, W., & Zenilman, J. (1998). Geographic epidemiology of gonorrhea in Baltimore, Maryland, using a geographic information system. *American Journal of Epidemiology, 147*(7), 709–716.

Becker, N. (1994). Cancer mapping: Why not use absolute scales? *European Journal of Cancer, 30A*(4), 699–706.

Bell, N., Schuurman, N., & Hayes, M. W. (2007). Using GIS-based methods of multicriteria analysis to construct socio-economic deprivation indices. *International Journal of Health Geographics, 6*, 17.

Bennett, W. D. (1981). A location allocation approach to health care facility location: A study of the undoctored population in Lansing, Michigan. *Social Science and Medicine, 15D*(2), 305–312.

Benschop, J., Hazelton, M., Stevenson, M., Dahl, J., Morris, R., & French, N. (2008). Descriptive spatial epidemiology of subclinical Salmonella infection in finisher pig herds: Application of a novel method of spatial adaptive smoothing. *Veterinary Research, 39*(1), 2.

Bernard, P., Charafeddine, R., Forhlich, K. L., Daniel, M., Kestens, Y., & Potvin, L. (2007). Health inequalities and place: A theoretical conception of neighbourhood. *Social Science and Medicine, 65*(9), 1839–1852.

Bernstein, K. T., Curriero, F. C., Jennings, J. M., Olthoff, G., Erbelding, E. J., & Zenilman J. (2004). Defining core gonorrhea transmission utilizing spatial data. *American Journal of Epidemiology, 160*(1), 51–58.

Berry, D., & Jarrell, T. W. (1999). Ranking priorities for a state family service integration initiative by census tract in northern New Castle County, Delaware, 1997. *Journal of Public Health Management Practice, 5*(2), 57–59.

Bertin, J. (1979). Visual perception and cartographic transcription. *World Cartography, 15*, 17–27.

Besag, J., & Newell, J. (1991). The detection of clusters in rare diseases. *Journal of the Royal Statistical Society, Series A, 154*(Part 1), 143–155.

Beyer, K., Comstock, S., & Seagren, R. (2010). Disease maps as context: A methodological approach for linking confidential health information with local knowledge for community health research. *Journal of Community Health*, 35(6), 635–644.

Beyer, K., & Rushton, G. (2009). Mapping cancer for community engagement. *Public Health Research, Practice and Policy, 6*(1), 1–8.

Bhaduri, B., Bright, E., Coleman, P., & Urban, M. L. (2007). LandScan USA: A high-resolution geospatial and temporal modeling approach for population distribution dynamics. *GeoJournal, 69*(1–2), 103–117.

Bhatia, R., Brenner, B., Salgado, B., Shamasunder, B., & Prakash, S. (2004). Biomonitoring. Burden of proof: Using research for environmental justice. *Race, Poverty & the Environment, 11*(2). Retrieved January 5, 2011, from *www.urbanhabitat.org/node/172.*

Bian, L. (2004). A conceptual framework for an individual-based, spatially-explicit epidemiological model. *Environment and Planning B, 31*(3), 381–395.

Bicki, J. T., & Brown, R. B. (1991). On-site sewage disposal: The influence of system density on water quality. *Journal of Environmental Health, 53*(5), 39–42.

Billings, J., Zeitel, L., Lukomnik, J., Carey, T. S., Blank, A. E., & Newman, L. (1993). Impact of socio-economic status on hospital use in New York City. *Health Affairs, 12*(1), 162–173.

Birkhead, G. S., & Maylahn, C. M. (2000). Sources of health-related information. In S. M. Teutsch & R. E. Churchill (Eds.), *Principles and practice of public health surveillance* (2nd ed., pp. 253–286). New York: Oxford University Press.

Birkin, M., Clarke, G., Clarke, M., & Wilson, A. (1996). *Intelligent GIS: Location decision and strategic planning.* Cambridge, UK: Geoinformation International.

Bithell, J. F. (1995). The choice of test for detecting raised disease risk near a point source. *Statistics in Medicine, 14*(21–22), 2309–2322.

Bivan, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2008). *Applied spatial data analysis with R*. New York: Springer.

Blanton, J. D., Palmer, D., Christian, K. A., & Rupprecht, C. E. (2008). Rabies surveillance in the United States during 2007. *Journal of the American Veterinary Medical Association, 233*(6), 884–897.

Block, J. P., Scribner, R. A., & DeSalvo, K. B. (2004). Fast food, race/ethnicity, and income: A geographic analysis. *American Journal of Preventive Medicine, 27*(3), 211–217.

Böhme, R. (Ed.). (1989–1993). *Inventory of world topographic mapping: Vols. 1–3*. London: Elsevier Applied Science.

Boone, J. D., McGwire, K. C., Otteson, E. W., DeBaca, R. S., Kuhn, E. A., Villard, P., et al. (2000). Remote sensing and geographic information systems: Charting Sin Nombre virus infections in deer mice. *Emerging Infectious Diseases, 6*(3), 248–258.

Boone, J. E., Gordon-Larsen, P., Stewart, J. D., & Popkin, B. M. (2008a). Validation of a GIS facilities database: Quantification and implications of error. *Annals of Epidemiology, 18*(5), 371–377.

Boone, J. E., Gordon-Larsen, P., Stewart, J. D., & Popkin, B. M. (2008b). Response to: Comments on Boone et al., "Validation of a GIS facilities database: Quantification and implications of error." *Annals of Epidemiology, 18*(10), 825.

Boots, B. (1999). Spatial tessellations. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems: Vol. 1. Principles and technical issues* (2nd ed., pp. 503–526). New York: Wiley.

Borah, D. K., & Bera M. (2003). Watershed-scale hydrologic and nonpoint-source pollution models: Review of mathematical bases. *Transactions of the American Society of Agricultural Engineers, 46*(6), 1553–1566.

Bosley, M. (1997). Laboratory-based surveillance for rotavirus—United States, July 1996–June 1997. *Morbidity and Mortality Weekly Report, 46*(46), 1092–1094.

Boulos, M. N. K. (2003). Location-based health information services: A new paradigm in personalised information delivery. *International Journal of Health Geographics, 2*, 2.

Boulos, M., & Burden, D. (2007). WebGIS in practice V: 3–D interactive and real-time mapping in Second Life. *International Journal of Health Geographics, 6*, 51.

Boulos, M. N. K., Curtis, A. J., & AbdelMalick, P. (2009). Musings on privacy issues in health research involving disaggregate geographic data about individuals. *International Journal of Health Geographics, 8*, 46.

Bowen, H. J., Palmer, S. R., Fielder, H. M. P., Coleman, G., Routledge, P. A., & Fone, D. L. (2000). Community exposures to chemical incidents: Development and evaluation of the first environment public health surveillance system in Europe. *Journal of Epidemiology and Community Health, 54*(11), 870–873.

Bowman, J. D., Thomas, D. C., Jiang, L., Jiang, F., & Peters, J. M. (1999). *Bioelectromagnetics, 20*(7), 399–413.

Bowman, J. (2000). GIS model of power lines used to study EMF and childhood leukemia. In C. M. Croner (Ed.), *Public Health GIS News and Information, No. 32* (pp. 7–10). Hyattsville, MD: National Center for Health Statistics.

Boyle, P., & Duke-Williams, O. (2004). Does migration exaggerate the relationship between deprivation and self-reported limiting long-term illness? In P. Boyle, S. Curtis, E. Graham, & E. Moore (Eds.), *The geography of health inequalities in the developed world: Views from Britain and North America* (pp. 129–148). Aldershot, UK: Ashgate.

Boyle, P., Norman, P., & Rees, P. (2004). Changing places. Do changes in the relative deprivation of areas influence limiting long-term illness and mortality among non-migrant people living in non-deprived households? *Social Science and Medicine, 58*(12), 2459–2471.

Bracken, I., Higgs, G., Martin, D., & Webster, C. (1989). *A classification of geographical information systems literature and applications*, CATMOG 52. Norwich, UK: Environmental Publications, University of East Anglia.

Brassel, K., Bucher, F., Stephan, E., & Vckovski, A. (1995). Completeness. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 81–108). Oxford, UK: Elsevier Science.

Brassel, K. E., & Rief, D. (1979). A procedure to generate Thiessen polygons. *Geographical Analysis, 11*(3), 289–303.

Brauner-Otto, S. R., Axinn, W. G., & Ghimire, D. J. (2007). The spread of health services and fertility transition. *Demography, 44*(4), 747–770.

Braveman, P. A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K. S., Metzler, M., et al. (2005). Socioeconomic status in health research: One size does not fit all. *Journal of the American Medical Association, 294*(22), 2879–2888.

Braveman, P., & Tarimo, E. (2002). Social inequalities in health within countries: Not only an issue for affluent nations. *Social Science and Medicine, 54*(11), 1621–1635.

Bretsky, P. M. (1995). *Rabies epizootic in Connecticut: Analysis and human risk reduction strategies*. Master's Thesis. New Haven, CT: Yale University.

Brewer, C., MacEachren, A., Pickle, L., & Herrman, D. (1997). Mapping mortality: Evaluating color schemes for choropleth maps. *Annals of the Association of American Geographers, 87*(3), 411–438.

Brewer, C., & Pickle, L. (2002). Evaluation of methods for classifying epidemiological data on choropleth maps. *Annals of the Association of American Geographers, 92*(4), 662–681.

Brody, J., Aschengrau, W., McKelvey, W., Rudel, R., Swartz, C., & Kennedy, T. (2004). Breast cancer risk and historical exposure to pesticides from wide-area applications assessed with GIS. *Environmental Health Perspectives, 112*(8), 889–897.

Brooker, S., & Utzinger, J. (2007). Integrated disease mapping in a polyparasitic world. *Geospatial Health, 1*(2), 141–146.

Broome, F. R., & Godwin, L. S. (2003). Partnering for the people: Improving the U.S. Census Bureau's MAF/TIGER database. *Photogrammetric Engineering and Remote Sensing, 69*(10), 1119–1123.

Brown, J. S., Graham, J. A., Chen, L. C., Postlethwait, E. M., Ghio, A. J., Foster, W. M., et al. (2007). Panel discussion: Session four—assessing biological plausibility of epidemiological findings in air pollution research. *Journal of Exposure Science and Environmental Epidemiology, 17*(S2), S97–S105.

Brown, R. N., & Lane, R. S. (1992). Lyme disease in California: A novel enzootic transmission cycle of Borrelia burgdorferi. *Science, 26*(5062), 1439–1442.

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the built environment for physical activity: State of the science. *American Journal of Preventive Medicine, 36*(4S), S99–S123.

Brownstein, J. S., Cassa, C. A., & Mandl, K. D. (2006). No place to hide—Reverse identification of patients from published maps. *The New England Journal of Medicine, 355*(16), 1741–1742.

Brugge, D., Leong, A., & Lai, Z. (1999). Can a community inject public health values into transportation issues? *Public Health Reports, 114*(1), 40–47.

Brunsdon, C. (1995). Estimating probability surfaces for geographical point data: An adaptive kernel algorithm. *Computers and Geosciences, 21*(7), 877–894.

Brutsaert, W. (1982). *Evaporation into the atmosphere: Theory, history, and applications.* Boston: Reidel.

Bullen, N., Jones, K., & Duncan, C. (1997). Modelling complexity: Analyzing between-individual and between-place variation—a multilevel tutorial. *Environment and Planning A, 29*(4), 585–609.

Bullen, N., Moon, G., & Jones, K. (1996). Defining localities for health planning: A GIS approach. *Social Science and Medicine, 42*(6), 801–816.

Bunge, W. (1971). *Fitzgerald: Geography of a revolution*. Cambridge, MA: Schenkman.

Bunge, W. (2011). *Fitzgerald: Geography of a revolution* (2nd ed.). Athens: University of Georgia Press.

Bunnell, J. E., Price, S. D., Das, A., Shields, T. M., & Glass, G. E. (2003). Geographic information systems and spatial analysis of adult *Ixodes scapularis* (Acari: Ixodidae) in the Middle Atlantic Region of the U.S.A. *Journal of Medical Entomology, 40*(4), 570–576.

Burke, L. M. (1993). *Environmental quality in Los Angeles (Technical Report 93–6).* Santa Barbara, CA: National Center for Geographic Information and Analysis.

Burns, C. M., & Inglis, A. D. (2007). Measuring food access in Melbourne: Access to health and fast foods by car, bus and foot in an urban municipality in Melbourne. *Health and Place, 13*(4), 877–885.

Buttenfield, B. P., & Mackaness, W. A. (1991). Visualization. In D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.), *Geographical information systems: Principles and applications: Vol. 1. Principles* (pp. 427–443). Harlow, UK: Longman Scientific & Technical.

Caballero, B. (2007). The global epidemic of obesity: An overview. *Epidemiologic Reviews, 29*(1), 1–5.

Cagney, K. A. (2006). Neighborhood age structure and its implications for health. *Journal of Urban Health: Bulletin of the New York Academy of Medicine, 83*(5), 827–834.

Cai, Q. (2007). DMAP IV user manual. Retrieved January 5, 2011, from *www.uiowa.edu/~gishlth/DMAP4.*

Cai, Q., Rushton, G., Bhaduri, B., Bright, E., & Coleman, P. (2006). Estimating small-area populations by age and sex using spatial interpolation and statistical inference methods. *Transactions in GIS, 10*(4), 577–598.

California Department of Pesticide Regulation. (2000). Pesticide use reporting. Retrieved January 5, 2011, from *www.cdpr.ca.gov/docs/pur/purovrvw/tabofcon.htm.*

Calisher, C. H., Mills, J. N, Sweeney, W. P., Root J. J., Reeder, S. A., Jentes, E. S., et al. (2005). Population dynamics of a diverse rodent assemblage in mixed grass-shrub habitat, southeastern Colorado, 1995–2005. *Journal of Wildlife Diseases, 41*(1), 12–28.

Callahan, G. M., & Broome, F. R. (1984). The joint development of a national 1:100,000 scale digital cartographic database. In *Technical Papers of the Annual Meeting of the American Congress on Surveying and Mapping* (pp. 246–253). Falls Church, VA: American Congress on Surveying and Mapping.

Carlin, S. (2001). Community breast cancer mapping—Huntington, Long Island. ESRI User Conference Proceedings, 2001. Retrieved January 5, 2011, from *proceedings.esri.com/library/userconf/proc01/professional/papers/pap496/p496.htm.*

Carlin, S. (2002). Community breast cancer mapping—Huntington, Long Island. Retrieved January 5, 2011, from *www.hbcac.org/mapping.html.*

Carlos, H., Shi, X., Sargent, J., Tanski, & S. Berke, E. (2010). Density estimation and adaptive bandwidths: A primer for public health practitioners. *International Journal of Health Geographics, 9*, 39.

Carrel, M., Emch, M., Jobe, R.T., Moody, A., & Wan, X. (2010). Spatiotemporal structure of molecular evolution of H5N1 highly pathogenic avian influenza viruses in Vietnam. *PLoS One, 5*(1), e8631.

Carstairs, V. (1995). Deprivation indices: Their interpretation and use in relation to health. *Journal of Epidemiology and Community Health, 49*(Suppl 2), S3–S8.

Carver, S. (2003). The future of participatory approaches using geographic information: Developing a research agenda for the 21st century. *URISA Journal, 15*(APA I), 61–71.

Cassa, C. A., Grannis, S. J., Overhage, J. M., & Mandl, K. D. (2006). A context-sensitive approach to anonymizing spatial surveillance data: Impact on outbreak detection. *Journal of the American Medical Informatics Association, 13*(2), 160–165.

Cassa, C. A., Wieland, S. C., & Mandl, K. D. (2008). Re-identification of home addresses from spatial locations anonymized by Gaussian skew. *International Journal of Health Geographics, 7*, 45.

Castro, M. B., & Wright, S. A. (2007). Vertebrate hosts of Ixodes pacificus (Acari: Ixodidae) in California. *Journal of Vector Ecology, 32*(1), 140–149.

Cecchi, G., Ilemobade, A., Le Brun, Y., Hogerwerf, L., & Slingenbergh, J. (2008). Agro-ecological features of the introduction and spread of the highly pathogenic avian influenza (HPAI) H5N1 in northern Nigeria. *Geospatial Health, 3*(1), 7–16.

Cech, I., Burau, K. D., & Walston, J. (2007). Spatial distribution of orofacial cleft defect births in Harris County, Texas, 1990 to 1994, and historical evidence for the presence of low-level radioactivity in tap water. *Southern Medical Journal, 100*(6), 560–569.

Center for International Earth Science Information Network, Socioeconomic Data and Applications Center. (2011). Gridded Population of the World and the Global Rural-Urban Mapping Project. Retrieved January 5, 2011, from *sedac.ciesin.columbia.edu/gpw/documentation.jsp*.

Centers for Disease Control and Prevention (CDC). (1998). Human rabies—Texas and New Jersey, 1997. *MMWR, 47*(1), 1–5.

Centers for Disease Control and Prevention (CDC). (1999). Update: Hantavirus Pulmonary Syndrome—United States, 1999. *MMWR, 48*(24), 521–524.

Centers for Disease Control and Prevention (CDC). (2000). Guidelines for Surveillance, Prevention, and Control of West Nile Virus Infection—United States. *MMWR, 49*(2), 25–28.

Centers for Disease Control and Prevention (CDC). (2001). Lyme Disease—United States, 1999. *MMWR, 50*(10), 181–185.

Centers for Disease Control and Prevention (CDC). (2004). Lyme Disease: Reported Cases of Lyme Disease, United States, 1999. Retrieved January 5, 2011, from *www.cdc.gov/ncidod/dvbid/lyme/distribution_density1999.htm*.

Centers for Disease Control and Prevention (CDC). (2006). CDC's National Environmental Public Health Tracking Progran: National Network Implementation Plan (NNIP), Version 1.0. Retrieved January 5, 2010, from *www.cdc.gov/nceh/tracking/pdfs/nnip.pdf*.

Centers for Disease Control and Prevention (CDC). (2007a). AIDS Surveillance in the United States. Retrieved January 5, 2011, from *www.cdc.gov/hiv/topics/surveillance/resources/software/apids/manual/section1.htm#Case-Def*.

Centers for Disease Control and Prevention (CDC). (2007b). Final 2007 West Nile Virus Human Neuroinvasive Disease Incidence in the United States. Retrieved January 5, 2011, from *www.cdc.gov/ncidod/dvbid/westnile/Mapsincidence/surv&control07IncidMaps.htm*.

Centers for Disease Control and Prevention (CDC). (2008a). Mapping Public Health: The Interactive Atlas of Reproductive Health. Retrieved January 5, 2011, from *www.cdc.gov/Features/AtlasReproductiveHealth*.

Centers for Disease Control and Prevention (CDC). (2008b). Summary of notifiable diseases, United States, 2006. *MMWR, 55*(53), 1–94.

Centers for Disease Control and Prevention (CDC). (2008c). Human rabies—Minnesota. *MMWR, 57*(17), 460–462.

Centers for Disease Control and Prevention (CDC). (2009a). National Health and Nutrition Examination Survey. Retrieved January 5, 2011, from *www.cdc.gov/nchs/nhanes/about_nhanes.htm*.

Centers for Disease Control and Prevention (CDC). (2009b). BRFSS Maps. Retrieved January 5, 2011, from *www.cdc.gov/BRFSS/maps/gis_data.htm*.

Centers for Disease Control and Prevention (CDC). (2009c). Fourth national report on human exposure to environmental chemicals. Retrieved January 5, 2011, from *www.cdc.gov/exposurereport/pdf/FourthReport.pdf*.

Centers for Disease Control and Prevention (CDC). (2010a). Lyme disease (*Borrelia burgdorferi*): 2008 case definition. Retrieved January 5, 2011, from *www.cdc.gov/ncphi/disss/nndss/casedef/lyme_disease_2008.htm*.

Centers for Disease Control and Prevention (CDC). (2010b). West Nile Virus: Statistics, Surveillance, and Control, Retrieved January 5, 2011, from *www.cdc.gov/ncidod/dvbid/westnile/surv&control.htm*.

Centers for Disease Control and Prevention (CDC). (2011a). National Program of Cancer Registries (NPCR). Retrieved March January 5, 2011, from *apps.nccd.cdc.gov/cancercontacts/npcr/contacts.asp#list*.

Centers for Disease Control and Prevention (CDC). (2011b). National Health Interview Survey. Retrieved January 5, 2011, from *www.cdc.gov/nchs/nhis.htm*.

Centers for Disease Control and Prevention (CDC). (2011c). Heart disease and stroke maps. Retrieved January 5, 2011, from *apps.nccd.cdc.gov/giscvh2/*.

Chadwick, E. (1965). *Report on the sanitary condition of the labouring population of Gt. Britain* (M. W. Flinn, Ed.). Edinburgh: Edinburgh University Press.

Chakraborty, J. (2009). Automobiles, air toxics, and adverse health risks: Environmental inequities in Tampa Bay, Florida. *Annals of the Association of American Geographers, 99*(4), 674–697.

Chakraborty, J., & Armstrong, M. P. (1995). Using geographic plume analysis to assess community vulnerability to hazardous accidents. *Computers, Environment, and Urban Systems, 19*(5–6), 1–17.

Chakraborty, J., Tobin, G. A., & Montz, B. E. (2005). Population evacuation: Assessing spatial vulnerability to natural hazards. *Natural Hazards Review, 6*(1), 23–33.

Chakraborty, J., & Zandbergen, P. A. (2007). Children at risk: Measuring racial/ethnic disparities in potential exposure to air pollution at school and home. *Journal of Epidemiology and Community Health, 61*(12), 1074–1079.

Chang, K.-T. (2009). *Introduction to geographic information systems* (5th ed.). Boston: McGraw-Hill.

Chang, S. W., Katz, M. H., & Hernandez, S. R. (1992). The new AIDS case definition: Implications for San Francisco. *Journal of the American Medical Association, 267*(7), 973–975.

Charles, D. (2005). Do maps have morals? *Technology Review, 108*(6), 77–79.

Chen, J., Roth, R., Naito, A., Lengerich, E., & MacEachren, A. (2008). Geovisual analytics to enhance spatial scan statistic interpretation: An example of U.S. cervical cancer mortality. *International Journal of Health Geographics, 7*, 57.

Chen, M., Chen, L. H., & Miller-Hooks. (2007). Traffic signal timing for urban evacuation. *Journal of Urban Planning and Development—ASCE, 133*(1), 30–42.

Chernin, P. R. (1995). Demonstrating watershed protection using GIS. *Journal of the New England Water Works Association, 109*(2), 132–139.

Chicago Police Department. (2011). CLEARMAP: Citizen law enforcement analysis and reporting. Retrieved January 5, 2011 from *gis.chicagopolice.org*.

Childs, J. E. (2009). Low tech versus high tech approaches for vector-borne disease control. *Vector-borne and Zoonotic Diseases, 9*(4), 355–356.

Childs, J. E., McLafferty, S. L., Sadek, R., Miller, G. L., Khan, A. S., DuPree, E. R., et al. (1998). Epidemiology of rodent bites and prediction of rat infestation in New York City. *American Journal of Epidemiology, 148*(1), 78–87.

Cho, A. (2007). An introduction to mashups for health librarians. *Journal of the Canadian Health Libraries Association/Journal de l'Association des bibliothèques de la santé du Canada, 28*(3), 19–22.

Cho, G. (2005). *Geographic information science: Mastering the legal issues*. Chichester, UK: Wiley.

Choynowski, M. (1959). Maps based on probabilities. *Journal of the American Statistical Association, 54*(286), 385–388.

Chrisman, N. (1997). *Exploring geographic information systems*. New York: Wiley.

Christaller, W. (1933). *Central places in southern Germany* (C. W. Baskin, Trans.). Englewood Cliffs, NJ: Prentice-Hall.

Christie, S., & Fone, D. (2003). Equity of access to tertiary hospitals in Wales: A travel time analysis. *Journal of Public Health Medicine, 25*(4), 344–350.

Chuang, Y. C., Cubbin, C., Ahn, D., & Winkleby, M. A. (2005). Effects of neighbourhood socioeconomic status and convenience store concentration on individual level smoking. *Journal of Epidemiology and Community Health, 59*(7), 568–573.

Church, R. L., & Cova, T. J. (2000). Mapping evacuation risk on transportation networks using a spatial optimization model. *Transportation Research Part C—Emerging Technologies, 8*(1–6), 321–336.

Church, R., & ReVelle, C. (1974). The maximal covering location problem. *Papers of the Regional Science Association, 32*, 101–118.

Church, R. L., & ReVelle, C. S. (1976). Theoretical and computational links between the p-median, location set-covering, and the maximal covering location problem. *Geographical Analysis, 8*(4), 406–415.

Church, R. L., & Sorensen, P. (1996). Integrating normative location models into GIS: problems and prospects with the *p*-median model. In P. Longley & M. Batty (Eds.), *Spatial analysis: Modelling in a GIS environment* (pp. 167–184). Cambridge: GeoInformation International.

Clarke, K. C. (1986). Advances in geographic information systems. *Computers, Environment, and Urban Systems, 10*(3–4), 175–184.

Clayton, D., & Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics, 43*(3), 671–681.

Cliff, A., & Haggett, P. (1988). *Atlas of disease distributions: Analytical approaches to epidemiological data*. Cambridge, MA: Blackwell.

Cliff, A., & Haggett, P. (2004). Time, travel and infection. *British Medical Bulletin, 69*(1), 87–99.

Cliff, A. D., & Haggett, P. (2006). A swash-backwash model of the single epidemic wave. *Journal of Geographical Systems, 8*(3), 227–252.

Cloud, J. (2002). American cartographic transformations during the Cold War. *Cartography and Geographic Information Science, 29*(3): 261–282.

Coelho, F. C., Cruz, O. G., & Codeço, C. D. (2008). Epigrass: A tool to study disease spread in complex networks. *Source Code for Biology and Medicine, 3*, 3.

Cohen, D. A., Ghosh-Dastidar, B., Scribner, R., Miu, A., Scott, M., Robinson, P., et al.

(2006). Alcohol outlets, gonorrhea, and the Los Angeles civil unrest: A longitudinal analysis. *Social Science and Medicine, 62*(12), 3062–3071.

Cohen, R., & Martinez, M. (2009). Health insurance coverage: Early release of estimates from the National Health Interview Survey, 2008. Retrieved January 5, 2011, from *www.cdc.gov/nchs/data/nhis/earlyrelease/insur200906.htm*.

Colten, C. E. (1991). A historical perspective on industrial wastes and groundwater contamination. *The Geographical Review, 81*(2), 215–228.

Committee on the Hygiene of Housing, American Public Health Association. (1948). *Planning the neighborhood*. Chicago: Public Administration Service.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(1), 35–46.

Congdon, P. (2001). The development of gravity models for hospital patient flows under system change: A Bayesian modeling approach. *Health Care Management Science, 4*(4), 289–304.

Conley, J., Gahegan, M., & Macgill, J. (2005). A genetic approach to detecting clusters in point data sets. *Geographical Analysis, 37*(3), 286–314.

Cook, R., Royce, R., Thomas, J., & Hanusa, B. (1999). What's driving an epidemic: The spread of syphilis along an interstate highway in rural North Carolina. *American Journal of Public Health, 89*(3), 369–373.

Cope, M., & Elwood, S. (2009). *Qualitative GIS: A mixed methods approach*. Thousand Oaks, CA: Sage.

Coppock, J. T., & Rhind, D. W. (1991). The history of GIS. In D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.). *Geographical information systems: Principles and applications: Vol. 1. Principles* (pp. 21–43). Harlow, UK: Longman Scientific & Technical.

Council of State and Territorial Epidemiologists. (2009). *CSTE list of nationally notifiable conditions.* Retrieved January 5, 2011, from *www.cste.org/dnn/LinkClick.aspx?fileticket=poIrEA4NUlo%3d&tabid=36&mid=1496*.

Cova, T. J., & Church, R. L. (1997). Modelling community evacuation vulnerability using GIS. *International Journal of Geographical Information Science, 11*(8), 763–784.

Cova, T. J., & Johnson, P. (2002). Microsimulation of neighborhood evacuations in the urban-wildland interface. *Environment and Planning A, 34*(12), 2211–2229.

Cowen, D. J. (1988). GIS versus CAD versus DBMS: What are the differences? *Photogrammetric Engineering and Remote Sensing, 54*(11), 1551–1555.

Cox, K. R. (1979). *Location and public problems: A political geography of the contemporary world*. Chicago: Maaroufa Press.

Cox, M., Boyle, P. J., Davey, P. G., Feng, Z., & Morris, A. D. (2007). Locality deprivation and Type 2 diabetes incidence: A local test of relative inequalities. *Social Science and Medicine, 65*(9), 1953–1964.

Craig, W. (1995). Why we can't share geographic information. In H. Onsrud and G. Rushton (Eds.), *Sharing geographic information* (pp. 107–118). New Brunswick, NJ: Center for Urban Policy Research.

Craig, W. (2009). Governance of the NSDI. *ArcNews, 31*(3), 1, 3–5.

Cressie, N. A. C. (1993). *Statistics for spatial data* (Revised edition). New York: Wiley.

Cromley, E. (2000). Case study of the use of GIS to inventory and understand the pattern of traffic accidents in Connecticut. In K. C. Clarke, *Getting started with geographical information systems* (3rd ed., pp. 257–261). Upper Saddle River, NJ: Prentice Hall.

Cromley, E. K. (2003). GIS and disease. *Annual Review of Public Health, 24*, 7–24.

Cromley, E. K. (2007). Risk factors contributing to motor vehicle collisions in an environment of uncertainty. *Stochastic Environmental Research and Risk Assessment, 21*(5), 473–486.

Cromley, E. K., Archambault, G., Aye, D., & McGee, S. (1997). *Accuracy of residential address locations geocoded by geographic information system address-matching procedures*. Unpublished manuscript.

Cromley, E. K., Cartter, M. L., Mrozinski, R. D., & Ertel, S.-H. (1998). Residential setting as a risk factor for Lyme disease in a hyperendemic region. *American Journal of Epidemiology, 147*(5), 472–477.

Cromley, E., & Cromley, R. G. (1996). An analysis of alternative classification schemes for medical atlas mapping. *European Journal of Cancer, 32A*(9), 1551–1559.

Cromley, E. K., Cromley, R. G, & Ye, Y. (2004). On-line reporting and mapping of spatially aggregated individual records selected by user queries. *Cartographica, 39*(2): 5–13.

Cromley, E. K., & Joy, K. P. (1995). Estimating childhood exposure to electromagnetic field associated with electric power transmission lines. In S. D. Majumdar, E. W. Miller, & F. J. Brenner (Eds.), *Environmental Contaminants, ecosystems and human health* (pp. 328–340). Easton: Pennsylvania Academy of Science.

Cromley, E. K., & Shannon, G. W. (1986). Locating ambulatory medical care facilities for the elderly. *Health Services Research, 21*(4), 499–514.

Cromley, R. G. (1996). A comparison of optimal classification strategies for choroplethic displays of spatially aggregated data. *International Journal of Geographic Information Systems, 10*(4), 405–424.

Cromley, R. G., & Cromley, E. K. (2009). Choropleth map legend design for visualizing community health disparities. *International Journal of Health Geographics, 8*, 52.

Cromley, R. G., & Mrozinski, R. D. (1999). The classification of ordinal data for choropleth mapping. *The Cartographic Journal, 36*(2), 101–109.

Cromley, R. G., & Ye, Y. (2006). Ogive-based legends for choropleth mapping. *Cartography and Geographic Information Science, 33*(4), 257–268.

Croner, C., Pickle, L., Wolf, D., & White, A. (1992). A GIS approach to hypothesis generation in epidemiology. In A. W. Voss (Ed.), *ASPRS/ACSM Technical Papers: Vol. 3. GIS and cartography* (pp. 275–283). Washington, DC: ASPRS/ACSM.

Crouse, D. L., Ross, N. A., & Goldberg, M. S. (2009). Double burden of deprivation and high concentrations of ambient air pollution at the neighbourhood scale in Montreal, Canada. *Social Science and Medicine, 69*(6), 971–981.

Cummins, S., Curtis, S., Diez-Roux, A. V., & Macintyre, S. (2007). Understanding and representing "place" in health research: A relational approach. *Social Science and Medicine, 65*(9), 1825–1838.

Cummins, S., Macintyre, S., Davidson, S., & Ellaway, A. (2005). Measuring neighbourhood social and material context: Generation and interpretation of ecological data from routine and non-routine sources. *Health and Place, 11*(3), 249–260.

Cummins, S. C. J., McKay, L., & Macintyre, S. (2005). McDonald's restaurants and neighborhood deprivation in Scotland and England. *American Journal of Preventive Medicine, 29*(4), 308–310.

Cummins, S., Stafford, M., Macintyre, S., Marmot, M., & Ellaway, A. (2005). Neighourhood environment and its association with self rated health: Evidence from Scotland and England. *Journal of Epidemiology and Community Health, 59*(3), 207–213.

Current, J. R., & Schilling, D. A. (1990). Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis, 22*(2), 116–126.

Curriero, F. C., Heiner, K. S., Samet, J. M., Zeger, S. L., Strug, L., & Patz. J. A. (2002). Temperature and mortality in 11 cities of the eastern United States. *American Journal of Epidemiology, 155*(1), 80–87.

Curry, M. (1998). *Digital places: Living with geographic technologies.* New York: Routledge.

Curry, M. (1999). *In plain and open view: Geographic information systems and the problem of privacy.* Retrieved January 5, 2011, from *www.spatial.maine.edu/~onsrud/tempe/curry.html*.

Curtis, A. J., Mills, J. W., & Leitner, M. (2006). Spatial confidentiality and GIS: Re-engineering mortality locations from published maps about Hurricane Katrina. *International Journal of Health Geographics, 5*, 44.

Curtis, A., Mills, J. W., & Leitner, M. (2007). Katrina and vulnerability: The geography of stress. *Journal of Health Care for the Poor and Underserved, 18*(2), 315–330.

Cutter, S. (1993). *Living with risk: The geography of technological hazards.* London: Edward Arnold.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly, 84*(2), 242–261.

Cutter, S. L., & Finch, C. (2008). Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences of the United States of America, 105*(7), 2301–2306.

Cutter, S. L., Mitchell, J. T., & Scott, M. S. (2000). Revealing the vulnerability of people and places: A case study of Georgetown County, South Carolina. *Annals of the Association of American Geographers, 90*(4), 713–737.

Dale, P. F. (1991). Land information systems. In D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.), *Geographical information systems: Principles and applications: Vol. 2. Applications* (pp. 85–99). Harlow, UK: Longman Scientific & Technical.

Dale, P. F., & McLaren, R. A. (1999). GIS in land administration. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems: Vol.2. Management issues and applications* (2nd ed., pp. 859–875). New York: Wiley.

Dangermond, J. (1984). A classification of software components commonly used in geographical information systems. In D. Marble, H. Calkins, & D. Peuquet (Eds.), *Basic readings in geographic information systems* (pp. 1–23–1–57). Amherst, NY: SPAD Systems.

Dangermond, J., & Morehouse, S. (1987). Trends in hardware for geographic information systems. In N. R. Chrisman (Ed.), *AUTO-CARTO 8 proceedings* (pp. 380–385). Falls Church, VA: American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping.

Daskin, M. S., Haghani, A. E., Khanal, M., & Malandraki, C. (1989). Aggregation effects in maximal covering models. *Annals of Operations Research, 18*, 115–40.

Davis, K. (1973). *Cities: Their origin, growth, and human impact.* San Francisco: Freeman.

Davis, K., & Stapleton, J. (1991). Migration to rural areas by HIV patients: Impact on HIV-related healthcare use. *Infection Control and Hospital Epidemiology, 12*(9), 540–543.

Davis, R. M., & Flores, P. I. (1992). Identifying potential pollutant point sources in an area of high ground water consumption. *Ground Water Monitoring Review, 12*(2), 116–119.

Deaton, A., & Lubotsky, D. (2003). Mortality, inequality and race in American cities and states. *Social Science and Medicine, 56*(6), 1139–1153.

de Castro, M. C., & Singer, B. H. (2006). Controlling the false discovery rate: A new application to account for multiple and dependent tests. *Geographical Analysis, 38*(2), 180–208.

De Cort, M., Dubois, G., Fridman, Sh. D., Germenchuk, M. G. Izrael, Yu. A., et al. (1998). Atlas of caesium deposition on Europe after the Chernobyl accident. Luxembourg: Office for Official Publications of the European Communities.

de Gruijter, J., Brus, D., Bierkens, M., & Knotters, M. (2006). *Sampling for natural resource monitoring*. New York: Springer.

Delaware Department of Natural Resources and Environmental Control. (2004). Delaware Bay & Estuary Basin Map 2.3—1 Septic System Locations. Retrieved January 5, 2011, from *www.dnrec.delaware.gov/SiteCollectionDocuments/Whole%20Basin%20Management/2.3–1septic.pdf*.

de Lepper, M. J. C., Scholten, H. J., & Stern, R. M. (Eds.). (1995). *The added value of geographical information systems in public and environmental health*. Dordrecht, The Netherlands: Kluwer Academic.

de Marchi, S., & Hamilton, J. T. (2006). Assessing the accuracy of self-reported data: An evaluation of the Toxics Release Inventory. *Journal of Risk and Uncertainty, 35*(1), 57–76.

DeMers, M. N. (2000). *Fundamentals of geographic information systems* (2nd ed.). New York: Wiley.

Densham, P. J. (1991). Spatial decision support systems. D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.), *Geographical information systems: Principles and applications* (pp. 403–412). Harlow, UK: Longman Scientific and Technical.

Department of Public Information, United Nations. (2000). *Cartography and geographic information science: Report on the meeting of the United Nations Geographic Information Working Group*. New York: United Nations.

de Smith, M. J., Goodchild, M. F., & Longley, P. A. (2007). *Geospatial analysis: A comprehensive guide to principles, techniques and software tools* (2nd ed.). Leicester, UK: Matador.

Devasundaram, J., Rohn, D., Dwyer, D., & Israel, E. (1998). A geographic information systems application for disease surveillance. *American Journal of Public Health, 88*(9), 1406–1407.

Devine, O., & Lewis, T. (1994). A constrained empirical Bayes estimator for incidence rates in areas with small populations. *Statistics in Medicine, 13*(11), 1119–1133.

DeVries, J., Nijkamp, P., & Rietveld, P. (2009). Exponential or power distance decay for commuting? An alternative specification. *Environment and Planning A*, 41(2), 461–480.

DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., et al. (Eds.). (2006). *Geographic information science & technology: Body of knowledge*. Washington, DC: Association of American Geographers.

DiBiase, D., MacEachren, A., Krygier, J., & Reeves, C. (1992). Animation and the role of map design in scientific visualization. *Cartography and Geographic Information Systems, 19*(4), 215–227.

Diehr, P. (1984). Small area statistics, large statistical problems. *American Journal of Public Health, 74*(4), 313–314.

Diez-Roux, A. V. (2000). Multilevel analysis in public health. *Annual Review of Public Health, 21*, 171–192.
Diez Roux, A. V. (2003a). A glossary for multilevel analysis, Part I. *Epidemiological Bulletin/PAHO, 24*(3), 11–13.
Diez Roux, A. V. (2003b). A glossary for multilevel analysis, Part II. *Epidemiological Bulletin/PAHO, 24*(4), 10–13.
Diez Roux, A. V. (2004). A glossary for multilevel analysis, Part III. *Epidemiological Bulletin/PAHO, 25*(1), 14–16.
Diggle, P., Moyeed, R., Rowlingson, B., & Thomson, M. (2002). Childhood malaria in The Gambia: A case-study in model-based geostatistics. *Journal of the Royal Statistical Society Series C-Applied Statistics, 51*(Part 4), 493–506.
Diggle, P. J., Tawn, J. A., & Moyeed, R. A. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society Series C-Applied Statistics, 47*(Part 3), 299–326.
Dijkstra, E. W. (1959). A note on two problems in connection with graphs. *Numeriche mathematik, 1*, 269–271.
Do, D. P., Finch, B. K., Basurto-Davila, R., Bird, C., Escarce, J., & Lurie, N. (2008). Does place explain racial health disparities? Quantifying the contribution of residential context to the Black/white health gap in the United States. *Social Science and Medicine, 67*(8), 1258–1268.
Doa, M. J. (1992). The Toxics Release Inventory. *Hazardous Waste and Hazardous Materials, 9*(1), 61–72.
Dolinoy, D. C., & Miranda, M. L. (2004). GIS modeling of air toxics releases from TRI-reporting and Non-TRI-reporting facilities: Impacts for Environmental Justice. *Environmental Health Perspectives, 112*(17): 1717–1724.
Domagalski, J. L., Ator, S., Coupe, R., McCarthy, K., Lampe, D., Sandstrom, M., et al. (2008). Comparative study of transport processes of nitrogen, phosphorus, and herbicides to streams in five agricultural basins, USA. *Journal of Environmental Quality, 37*(3), 1158–1169.
Dorling, D., & Barford, A. (2007). Shaping the world to illustrate inequalities in health. *Bulletin of the World Health Organization, 85*(11), 890–893.
Dorling, D., Mitchell, R., Shaw, M., Orford, S., & Smith, G. D. (2000). The Ghost of Christmas Past: Health effects of poverty in London in 1896 and 1991. *British Medical Journal, 321*(7276), 1547–1551.
Drummond, J. (1995). Positional accuracy. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 31–58). Oxford, UK: Elsevier Science.
Dublin Core Metadata Intiative. (2011). *About the Initiative*. Dublin Core Metadata Initiative. Retrieved January 5, 2011, from *dublincore.org/about*.
Duncan, C., Jones, K., & Moon, G. (1998). Context, composition and heterogeneity: Using multilevel models in health research. *Social Science and Medicine, 46*(1), 97–117.
Duncan, G. J., Daly, M. C., McDonough, P., & Williams, D. R. (2002). Optimal indicators of socioeconomic status for health research. *American Journal of Public Health, 92*(7), 1151–1157.
Dunn, C. (2007). Participatory GIS: A people's GIS? *Progress in Human Geography, 31*(5), 616–637.
Dunnivant, F. M., & Anders, E. (2006). A basic introduction to pollutant fate and transport: An integrated approach with chemistry, modeling, risk assessment, and environmental legislation. Hoboken, NJ: Wiley.
Dyck, I. (1990). Context, culture and client: Geography and the health for all strategy. *Canadian Geographer, 34*(4), 338–341.

Dykes, J., MacEachern, A. M., & Kraak, M-J. (2005). *Exploring geovisualization*. Oxford, UK: Elsevier.

Eaton, D. J., Daskin, M. S., Simmons, D., Bulloch, B., & Jansma, G. (1985). Determining emergency medical service vehicle deployment in Austin, Texas. *Interfaces, 15*(1), 96–108.

Ebdon, D. (1985). *Statistics in geography* (2nd ed.). New York: Basil Blackwell.

The Eclipse Foundation. (2011). The Spatiotemporal Epidemiological Modeler (STEM) Project. Retrieved January 5, 2011, from *www.eclipse.org/stem*.

El Emam, K., Brown, A., & AbdelMalik, P. (2009). Evaluating predictors of geographic area population size cut-offs to manage re-identification risk. *Journal of the American Medical Informatics Association, 16*(2), 256–266.

Elgethun, K., Fenske, R. A., Yost, M. G., & Palcisko, G. J. (2003). Time-location analysis for exposure assessment studies of children using a novel global positioning system instrument. *Environmental Health Perspectives, 111*(1), 115–122.

Elias, J., Harmsen, D., Claus, H., Hellenbrand, W., Frosch, M., & Volgel, U. (2006). Spatiotemporal analysis of invasive meningococcal disease, Germany. *Emerging Infectious Diseases, 12*(11), 1689–1695.

Elmasri, R. E., & Navathe, S. B. (2007). *Fundamentals of database systems* (5th ed.). Reading, MA: Addison-Wesley.

Elwood, S. (2006a). Negotiating knowledge production: The everyday inclusions, exclusions and contradictions of participatory GIS research. *The Professional Geographer, 58*(2), 197–208.

Elwood, S. (2006b). Beyond cooptation or resistance: Urban spatial politics, community organization and GIS-based spatial narratives. *Annals of the Association of American Geographers, 96*(2), 323–341.

Elwood, S. 2008. Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal, 72*(3–4), 173–183.

Emch, M., Ali, M., Park, J., Yunus, M., Sack D., & Clemens, J. (2006). Relationship between neighborhood-level killed oral cholera vaccine coverage and protective efficacy: Evidence for herd immunity. *International Journal of Epidemiology, 35*(4), 1044–1050.

Evrard, O., Persoons, E., Vandaele, K., & van Wesemael, B. (2007). Effectiveness of erosion mitigation measures to prevent muddy floods: A case study in the Belgian loam belt. *Agriculture, Ecosystems and Environment, 118*(1–4), 149–158.

Ewald, P. W., & Cochran, G. (1999). Catching on to what's catching. *Natural History, 108*(1), 34–37.

Eylenbosch, W. J., & Noah, N. D. (Eds.). (1988). *Surveillance in health and disease.* Oxford, UK: Oxford University Press.

Federal Geographic Data Committee. (1998). Content standard for digital geospatial metadata. CSDGM-Version 2–FGDC-STD-001–1998. Federal Geographic Data Committee. Retrieved January 5, 2011, from *www.fgdc.gov/standards/standards_publications/index_html*.

Federal Geographic Data Committee. (2009). *Data & services*. Retrieved January 5, 2011, from *www.fgdc.gov/dataandservices*.

Ferguson, A. G., & Morris, C. N. (2007). Mapping transactional sex on the Northern Corridor highway in Kenya. *Health and Place, 13*(2), 504–519.

Field, K., & Beale, L. (2004). Using GIS to model incidence, prevalence and spread of non-legal drug use. *Transactions in GIS, 8*(4), 423–439.

Fielding, J. (2007). Environmental injustice or just the lie of the land: An investigation of the socio-economic class of those at risk from flooding in England and Wales. *Sociological Research Online, 12*, 4.

Finkel, A. M., & Golding, D. (Eds.). (1994). *Worst things first? The debate over risk-based national environmental priorities.* Washington, DC: Resources for the Future.

Fisher, E. (2007). *Opening Pandora's Box: Contextualizing the Precautionary Principle in the European Union*. University of Oxford Faculty of Law Legal Studies Research Paper Series, Working Paper No 2/2007. Retrieved January 5, 2011, from *papers.ssrn.com/sol3/papers.cfm?abstract_id=956952*.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley & Sons.

Flowerdew, R., Feng, Z., & Manley, D. J. (2007). Constructing data zones for Scottish Neighborhood Statistics. *Computers, Environment and Urban Systems, 31*(1), 76–90.

Flowerdew, R., & Green, M. (1989). Statistical methods for inference between incompatible zoning systems. In M. F. Goodchild & S. Gopal (Eds.), *Accuracy of spatial databases* (pp. 239–248). London: Taylor & Francis.

Flowerdew, R., & Green, M. (1994). Areal interpolation and types of data. In S. Fotheringham & P. Rogerson (Eds.), *Spatial analysis and GIS* (pp. 121–145). Bristol, PA: Taylor & Francis.

Flowerdew, R., Manley, D. J., & Sabel, C. E. (2008). Neighbourhood effects on health: Does it matter where you draw the boundaries? *Social Science and Medicine, 66*(6), 1241–1255.

Foley, R., & Platzer, H. (2007). Place and provision: Mapping mental health advocacy services in London. *Social Science and Medicine, 64*(3), 617–632.

Foot, D. (1981). *Operational urban models: An introduction.* New York: Methuen.

Ford, D. A., Kaufman, J. H., & Eiron, I. (2006). An extensible spatial and temporal epidemiological modelling system. *International Journal of Health Geographics, 5*, 4.

Ford, T. E., Colwell, R. R., Rose, J. B., Morse, S. S., Rogers, D. J., & Yates, T. L. (2009). Using satellite images of environmental changes to predict infectious disease outbreaks. *Emerging Infectious Diseases, 15*(9), 1341–1346.

Foresman, T. W. (Ed.). (1998). *The history of geographic information systems: Perspectives from the pioneers*. Upper Saddle River, NJ: Prentice-Hall.

Forsyth, A. (Ed.). (2007). Environment and Physical Activity: GIS Protocols, Version 4.1. Retrieved January 5, 2011, from *www.designforhealth.net/pdfs/GIS_Protocols/MinnGIS_Ver4_1_FINAL.pdf.*

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2000). *Quantitative geography: Perspectives on spatial data analysis*. London: Sage.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, UK: Wiley.

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A, 30*(11), 1905–1927.

Fotheringham, A. S., Charlton, M., & Brunsdon, C. (2009). Geographically Weighted Regression. Retrieved January 5, 2011, from *ncg.nuim.ie/ncg/GWR/index.htm.*

Fotheringham, A. S., & Rogerson, P. (1994). *Spatial analysis and GIS*. Bristol, PA: Taylor & Francis.

Fotheringham, A. S., & Zhan, F. B. (1996). A comparison of three exploratory methods for cluster detection in spatial point patterns. *Geographical Analysis, 28*(3), 200–218.

Francis, R. L., Lowe, T. J., Rayco, M. B., & Tamir, A. (2009). Aggregation error for locations models: Survey and analysis. *Annals of Operations Research, 167*(1), 171–208.

Franco-Paredes, C., Carrasco, P., & Preciado, J. (2009). The first influenza pandemic in the new millennium: Lessons learned hitherto for current control efforts and overall pandemic preparedness. *Journal of Immune Based Therapies and Vaccines, 7*, 2.

Freifield, C. C., Mandl, K. D., Reis, B. Y., & Brownstein, J. S. (2008). HealthMap: Global infections disease monitoring through automated classifications and visualization of Internet media reports. *Journal of the American Medical Informatics Association, 15*(2), 150–157.

Frickel, S., & Elliott, J. R. (2008). Tracking industrial land use conversions: A new approach for studying relict waste and urban development. *Organization and Environment, 21*(2), 128–147.

Friel, S., Chopra, M., & Satcher, D. (2007). Unequal weight: Equity oriented policy responses to the global obesity epidemic. *British Medical Journal, 335*(7632), 1241–1243.

Friis, R. H., & Sellers, T. A. (2009). *Epidemiology for public health practice* (4th ed). Sudbury, MA: Jones & Bartlett.

Fry, C. M. (1988). Maps for the physically disabled. *The Cartographic Journal, 25*(11), 20–28.

Gale, P., Drew, T., Phipps, L. P., David, G., & Wooldridge, M. (2009). The effect of climate change on the occurrence and prevalence of livestock diseases in Great Britain: A review. *Journal of Applied Microbiology, 106*(5), 1409–1423.

Galea, S., Ahern, J., Rudenstine, S., Wallace, Z., & Vlahov, D. (2005). Urban built environment and depression: A multilevel analysis. *Journal of Epidemiology and Community Health, 59*(10), 822–827.

Galea, S., Nandi, A., & Vlahov, D. (2004). The social epidemiology of substance use. *Epidemiologic Reviews, 26*(1), 36–52.

Gallo, G., & Pallottino, S. (1988). Shortest paths algorithms. *Annals of Operations Research, 13*, 3–79.

Gao, S., Mioc, D., Anton, F., Yi, X., & Coleman, D. J. (2008). Online GIS services for mapping and sharing disease information. *International Journal of Health Geographics, 7*, 8.

Garbe, E., & Blount, S. (1992). Chronic disease surveillance. In W. Halperin & E. Baker (Eds.), *Public health surveillance* (pp. 130–139). New York: Van Nostrand Reinhold.

Gatrell, A. C., Bailey, T. C., Diggle, P. J., & Rowlingson, B. S. (1996). Spatial point pattern analysis and its application in medical geography. *Transactions of the Institute of British Geographers, NS21*(1), 256–274.

Gatrell, A. C., & Löytönen, M. (Eds.). (1998). *GIS and health*. London: Taylor & Francis.

Gawande, A. (2009). The cost conundrum: What a Texas town can teach us about health care. *The New Yorker, 85*(16), 36.

General Accounting Office. (1995). *Health care shortage areas: Designations not a useful tool for directing resources to the underserved*, GAO/HEHS-95–200. Washington, DC: United States General Accounting Office.

Geoffrion, A. M. (1993). Can OR/MS evolve fast enough? *Interfaces, 13*(10), 10–25.

Geography and Map Division. (1981). *Fire insurance maps in the Library of Congress.* Washington, DC: Library of Congress.

Georgopoulos, P. G., Sasso, A. F., Isukapalli, S. S., Lioy, P. J., Vallero, D. A., Okino, M., et al. (2009). Reconstructing population exposures to environmental chemicals from biomarkers: Challenges and opportunities. *Journal of Exposure Science and Environmental Epidemiology, 19*(2), 149–171.

Geospatial One-Stop. (2011). *geodata.gov—Frequently asked questions*. Retrieved January 5, 2011, from g*os2.geodata.gov/gos/static/default/faq.htm*.

Gesler, W. M., & Meade, M. S. (1988). Locational and population factors in health care-seeking behavior in Savannah, Georgia. *Health Services Research, 23*(3), 443–462.

Getis, A., & Ord, J. (1992). The analysis of spatial association by the use of distance statistics. *Geographical Analysis, 24*(3), 189–206.

Ghose, R. (2001). Use of information technology for community empowerment: Transforming geographic information systems into community information systems. *Transactions in GIS, 5*(2), 141–63.

Ghose, R., & Huxhold, W. (2001). Role of local contextual factors in building public participation GIS: The Milwaukee experience. *Cartography and Geographic Information Science, 28*(3), 195–208.

Ghosh A., & Rushton, G. (Eds.). (1987). *Spatial analysis and location-allocation models*. New York: Van Nostrand Reinhold.

Gibbons, M. C. (2005). A historical overview of health disparities and the potential of eHealth solutions. *Journal of Medical Internet Research, 7*(5), e50.

Giebultowicz, S., Ali, M., Yunus, M., & Emch, M. (2011). A comparison of spatial and social clustering of cholera in Matlab, Bangladesh. *Health and Place, 17*(2), 490–497.

Gilbert, E. W. (1958). Pioneer maps of health and disease in England. *Geographical Journal, 124*(2), 172–183.

Gilliland, J. A., & Novak, M. (2006). On positioning the past with the present: The use of fire insurance plans and GIS for urban environmental history. *Environmental History, 11*(1), 136–139.

Ginsberg, H. S. (1993). Geographic spread of *Ixodes dammini* and *Borrelia burgdorferi*. In H. S. Ginsberg (Ed.), *Ecology and environmental management of Lyme disease*, pp. 63–81. New Brunswick, NJ: Rutgers University Press.

Ginsberg, H. S. (1994). Lyme disease and conservation. *Conservation Biology, 8*(2), 343–353.

Glanz, M. (2009). Measuring food environments: A historical perspective. *American Journal of Preventive Medicine, 36*(4S), S93–S98.

Glass, G. E., Amerasinghe, F. P., Morgan, J. M., III, & Scott, T. W. (1994). Predicting *Ixodes scapularis* abundance on white-tailed deer using geographic information systems. *American Journal of Tropical Medicine and Hygiene, 51*(5), 538–544.

Glass, G. E., Cheek, J. E., Patz, J. A., Shields, T. M., Doyle, T. J., Thoroughman, D. A., et al. (2000). Using remotely sensed data to identify areas at risk for hantavirus pulmonary syndrome. *Emerging Infectious Diseases, 6*(3), 238–247.

Glass, G. E., Morgan, J. M., III, Johnson, D. T., Noy, P. M., Israel, E., & Schwartz, B. S. (1992). Infectious disease epidemiology and GIS: A case study of Lyme disease. *Geo Info Systems, 3*(3), 65–69.

Glass, G. E., Schwartz, B. S., Morgan, J. M., Johnson, D. T., Noy, P. M., & Israel, E. (1995). Environmental risk factors for Lyme disease identified with GIS. *American Journal of Public Health, 85*(7), 944–948.

Godlund, S. (1961). *Population, regional hospitals, transportation facilities, and regions:*

Planning the location of regional hospitals in Sweden. Lund Studies in Geography Series B: Human Geography No. 21. Lund, Sweden: Department of Geography, Royal University of Lund.

Gold, C. M. (1990). Spatial data structures—the extension from one to two dimensions. In L. F. Pau (Ed.), *Mapping and spatial modeling for navigation* (pp. 11–39). Berlin: Springer.

Goldberg, S. (2008). *A geocoding best practices guide*. Springfield, IL: North American Association of Central Cancer Registries.

Golledge, R. G., & Stimson, R. J. (1997). *Spatial behavior: A geographic perspective.* New York: Guilford Press.

Goodchild, M. F. (1984). Geocoding and geosampling. In G. L. Gaile & C. Willmott (Eds.), *Spatial statistics and models* (pp. 33–53). Dordrecht, The Netherlands: D. Reidel.

Goodchild, M. F. (1992). Geographic data modeling. *Computers and Geosciences, 18*(4), 311–319.

Goodchild, M. F. (1995a). Attribute accuracy. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 59–80). Oxford, UK: Elsevier Science.

Goodchild, M. F. (1995b). GIS and geographic research. In J. Pickles (Ed.), *Ground truth: The social implications of geographic information systems* (pp. 31–50). New York: Guilford Press.

Goodchild, M. F., Fu, P., & Rich, P. (2007). Sharing geographic information: An assessment of the Geospatial One-Stop. *Annals of the Association of American Geographers, 97*(2), 250–266.

Goodchild, M. F., Haining, R., Wise, S., Arbia, G., Anselin, L., Bossard, E., et al. (1992). Integrating GIS and spatial data analysis: Problems and possibilities. *International Journal of Geographical Information Systems, 6*(5), 407–423.

Goodchild, M., & Lam, N. (1980). Areal interpolation: A variant of the traditional spatial problem. *Geo-Processing, 1*, 297–312.

Goodman, D. C., & Fisher, E. S. (2008). Physician workforce crisis? Wrong diagnosis, wrong prescription. *New England Journal of Medicine, 358*(16), 1658–1661.

Goodman, D. C., Fisher, E. S., Little, G. A., Stukel, T. A., & Chang, C. (2001). Are neonatal intensive care resources located according to need? Regional variation in neonatologists, beds, and low birth weight newborns. *Pediatrics, 108*(2), 426–431.

Goodman, D., Fisher, E., Stukel, T., & Chang, C. (1997). The distance to community medical care and the likelihood of hospitalization: Is closer always better? *American Journal of Public Health, 87*(7), 1144–1150.

Google. (2011a). *Google code-KML*. Retrieved January 5, 2011, from *code.google.com/apis/kml*.

Google. (2011b). KML Tutorial. Retrieved January 5, 2011, from code.google.com/apis/kml/documentation/kml_tut.html.

Gordis, L. (2004). *Epidemiology* (3rd ed.). Philadelphia, PA: W. B. Saunders.

Gould, P. R. (1995). *The coming plague*. New York: Blackwell.

Gould, P. R., & Leinbach, T. R. (1966). An *approach* to the geographic assignment of hospital services. *Tijdschrift voor economische in social geografie, 57*(5), 203–206.

Graham, E., Boyle, P., Curtis, S., & Moore, E. (2004). Health geographies in the developed world. In P. Boyle, S. Curtis, E. Graham, & E. Moore (Eds.), *The geography of health inequalities in the developed world: Views from Britain and North America* (pp. 3–33). Aldershot, UK: Ashgate.

Grauman, D., Tarone, R., Devesa, S., & Fraumeni, J. (2000). Alternate ranging methods for cancer mortality maps. *Journal of the National Cancer Institute, 92*(7), 534–543.

Green D., & Bossomaier, T. (2002). *Online GIS and spatial metadata.* London: Taylor & Francis.

Green, M. B., Cromley, R. G., & Semple, R. K. (1980). The bounded transportation problem. *Economic Geography, 56*(1), 30–44.

Greenberg, M. R. (1978). *Applied linear programming.* New York: Academic Press.

Greenberg, M. (2002). Should housing be built on former brownfield sites? *American Journal of Public Health, 92*(5), 703–705.

Greenberg, M. R. (2003). Reversing urban decay: Brownfield redevelopment and environmental health. *Environmental Health Perspectives, 111*(2), A74–A75.

Greenberg, M., & Lewis, M. J. (2000). Brownfields redevelopment, preferences and public involvement: A case study of an ethnically mixed neighbourhood. *Urban Studies, 37*(13), 2501–2514.

Greenberg, M., & Schneider, D. (1992). Region of birth and mortality of blacks in the United States. *International Journal of Epidemiology, 21*(2), 324–328.

Greenberg, M., & Wartenberg, D. (1991). Communicating to an alarmed community about cancer clusters: A fifty state survey. *Journal of Community Health, 16*(2), 71–82.

Gregorio, D. I., Cromley, E., Mrozinski, R., & Walsh, S. J. (1999). Subject loss in spatial analysis of breast cancer. *Health and Place, 5*(2), 173–177.

Griffith, D. A. (2008). Geographic sampling of urban soils for contaminant mapping: How many samples and from where. *Environmental Geochemistry and Health, 30*(6), 495–509.

Griffith, D. A., Doyle, P. G., Wheeler, D. C., & Johnson, D. L. (1998). A tale of two swaths: Urban childhood blood-lead levels across Syracuse, New York. *Annals of the Association of American Geographers, 88*(4), 640–665.

Griffith, D. A., Millones, M., Vincent, M., Johnson, D. L., & Hunt, A. (2007). Impacts of positional error on spatial regression analysis: A case study of address locations in Syracuse, New York. *Transactions in GIS, 11*(5), 655–679.

Grimson, R. (1999). Community breast cancer maps. Department of Preventive Medicine, State University of New York at Stony Brook. Retrieved January 5, 2011, from *www.conservationgis.org/scgis/ScgNews2/ctspsection/ct99huntington.html.*

Grineski, S. E., & Collins, T. W. (2008). Exploring patterns of environmental injustice in the Global South: *Maquiladoras* in Ciudad Juárez, Mexico. *Population and Environment, 29*(6), 247–270.

Grunwald, S., & Qi, C. (2006). GIS-based water quality modeling in the Sandusky Watershed, Ohio, USA. *Journal of the American Water Resources Association, 42*(4), 957–973.

Guagliardo, M. F. (2004). Spatial accessibility of primary care: Concepts, methods and challenges. *International Journal of Health Geographics, 3*, 3.

Guagliardo, M. F., Jablonski, K. A., Joseph, J. G., & Goodman, D. C. (2004). Do pediatric hospitalizations have a unique geography? *BMC Health Services Research, 4*, 2.

Guagliardo, M., Ronzio, C., Cheung, I., Chacko, E., & Joseph, J. (2004). Physician accessibility: An urban case study of pediatric providers, *Health and Place, 10*(3), 273–283.

Gubler, D. (2004). The changing epidemiology of yellow fever and dengue, 1900 to 2003:

Full circle? *Comparative Immunology, Microbiology, and Infectious Diseases, 27*(5), 319–330.

Gubler, D. (2008). The global threat of emergent/reemergent vector-borne diseases. In Institute of Medicine, *Vector-borne diseases: Understanding the environmental, human health, and ecological connections* (pp. 43–64). Washington, DC: National Academies Press.

Guerra, C. A., Hay, S. I., Lucioparedes, L. S., Gikandi, P. W., Tatem, A. J., Noor, A. M., et al. (2007). Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. *Malaria Journal, 6*, 17.

Guerra, M. A., Walker, E. D., & Kitron, U. (2001). Canine surveillance system for Lyme borreliosis in Wisconsin and northern Illinois: Geographic distribution and risk factor analysis. A*merican Journal of Tropical Medicine and Hygiene, 65*(5), 546–552.

Gumusay, M. U., Unal, A., & Aydin, R. (2008). Use of geographical information systems in analyzing vehicle emissions: Istanbul as a case study. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37*(Part B1), 997–1000.

Guo, Q., Kelly, M., & Graham, C. H. (2005). Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecological Modelling, 182*(1), 75–90.

Guptill, S. C. (1995). Temporal information. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 153–165). Oxford, UK: Elsevier Science.

Guthe, W. G., Tucker, R. K., & Murphy, E. A. (1992). Reassessment of lead exposure in New Jersey using GIS technology. *Environmental Research, 59*(2), 318–325.

Habicht, F. H., II. (1994). EPA's vision for setting national environemntal priorities. In A. M. Finkel & D. Golding, *Worst things first?: The debate over risk-based national environmental priorities* (pp. 33–46). Washington, DC: Resources for the Future.

Haining, R. (2003). *Spatial data analysis: Theory and practice*. Cambridge, UK: Cambridge University Press.

Haklay, M., & Tobon, C. (2003). Usability evaluation and PPGIS: Towards a user-centred design approach. *International Journal of Geographic Information Science, 17*(6), 577–592.

Hales, S., Weinstein, P., & Woodward, A. (1997). Public health impacts of global climate change. *Review of Environmental Health, 12*(3), 191–199.

Hallenbeck, W. H. (1993). *Quantitative risk assessment for environmental and occupational health* (2nd ed.). Boca Raton, FL: Lewis.

Halperin, W., & Baker, E. (Eds.). (1992). *Public health surveillance*. New York: Van Nostrand Reinhold.

Hammen, J. L., & Gerla, P. J. (1994). A geographic information systems approach to wellhead protection. *Water Resources Bulletin, 30*(5), 833–839.

Han, D., Rogerson, P. A., Bonner, M. R., Nie, J., Vena, J. E., Muti, P., et al. (2005). Assessing spatio-temporal variability of risk surface using residential history data in a case control study of breast cancer. *International Journal of Health Geographics, 4*, 9.

Han, L. L., Popovici, F., Alexander, J. P., Jr., Laurentia, V., Tengelsen, L. A., Cernescu, C., et al. (1999). Risk factors for West Nile virus infection and meningoencephalitis, Romania, 1996. *Journal of Infectious Diseases, 179*(1), 230–233.

Hanchette, C. (1999). GIS implementation of 1997 CDC guidelines for childhood lead screening in North Carolina. *GIS in public health, 3rd National Conference abstracts.* San Diego, California.

Handy, S. L., & Niemeier, D. A. (1997). Measuring accessibility: An exploration of issues and alternatives. *Environment and Planning A, 29*(7), 1175–1194.

Hanley, G. E., & Morgan, S. (2008). On the validity of area-based income measures to proxy household income. *BMC Health Services Research, 8*, 79.

Harris, T., Weiner, D., Warner, T., & Levin, R. (1995). Pursuing social goals through participatory geographic information systems: Redressing South Africa's historical political ecology. In J. Pickles (Ed.), *Ground truth: The social implications of geographic information systems* (pp. 196–222). New York: Guilford Press.

Harrower, M. (2003). Tips for designing effective animated maps. *Cartographic Perspectives, 44*(Winter), 63–65.

Harrower, M., & Brewer, C. (2003). ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal, 40*(1), 27–37.

Harrower, M., & Sheesley, B. (2005). Designing better map interfaces: A framework for panning and zooming. *Transactions in GIS, 9*(2), 77–89.

Hassan, M. M. (2005). Arsenic poisoning in Bangladesh: Spatial mitigation planning with GIS and public participation. *Health Policy, 74*(3), 244–260.

Hattis, D. (1996). Drawing the line: Quantitative criteria for risk management. *Environment, 38*(6), 10–15, 35–39.

Hawley, K., & Moellering, H. (2005). A comparative analysis of areal interpolation methods. *Cartography and Geographic Information Science, 32*(4), 411–423.

Hay, S. I., Guerra, C. A., Gething, P. W., Patil, A. P., Tatem, A. J., Noor, A. M., et al. (2009). A world malaria map: *Plasmodium falciparum* endemicity in 2007. *PLOS Medicine, 6*(3), 286–302.

Hay, S. I., & Snow, R. W. (2006). The Malaria Atlas Project: Developing global maps of malaria. *PLOS Medicine, 3*(2), 2204–2208.

Haynes, K. E., & Fotheringham, A. S. (1984). *Gravity and spatial interaction models*. Beverly Hills, CA: Sage.

Haynes, R. M., & Bentham, C. G. (1982). The effects of accessibility on general practitioner consultations, out-patient attendances and in-patient admissions in Norfolk, England. *Social Science and Medicine, 16*(5), 561–569.

Haynes, R., Bentham, G., Lovett, A., & Gale, S. (1999). Effects of distances to hospital and GP surgery on hospital inpatient episodes, controlling for needs and provision. *Social Science and Medicine, 49*(3), 425–433.

Haynes, R., Lovett, A., & Sunnenberg, G. (2003). Potential accessibility, travel time, and consumer choice: Geographical variations in general medical practice registrations in Eastern England. *Environment and Planning A, 35*(10), 1733–1750.

Health Canada Santé Canada. (2007). Biomonitoring of environmental chemicals in the Canadian Health Measures Survey. Retrieved January 5, 2011, from *www.hc-sc.gc.ca/ewh-semt/contaminants/health-measures-sante-eng.php*.

Health Resources and Services Administration. (2010). Shortage designation: HPSAs, MUAs & MUPs. Retrieved March January 5, 2011, from *bhpr.hrsa.gov/shortage/index.htm*.

Healy, R. G. (1991). Database management systems. In D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.), *Geographical information systems: Principles and applications: Vol. 1. Principles* (pp. 427–443). Harlow, UK: Longman Scientific & Technical.

Hearne, S. A. (1996). Tracking toxics: Chemical use and the public's "right-to-know." *Environment, 38*(6), 4–9, 28–34.

Hearnshaw, H. M., & Unwin, D. J. (Eds.). (1994). *Visualization in geographical information systems*. Chichester, UK: Wiley.

Heimer, R., Barbour, R., Shaboltas, A., Hoffman, I., & Kozlov, A. (2008). Spatial distribution of HIV prevalence and incidence among injection drugs users in St. Petersburg: Implications for HIV transmission. *AIDS, 22*(1), 123–130.

Henderson, D. (1980). Smallpox eradication. *Public Health Reports, 95*(5), 422–426.

Hermann, D. J., & Pickle, L. W. (1996). A cognitive subtask model of statistical map reading. *Visual Cognition, 3*(2), 165–190.

Hewko, J., Smoyer-Tomic, K., & Hodgson, J. (2002). Measuring neighborhood spatial accessibility to urban amenities: Does aggregation error matter? *Environment and Planning A, 34*(7), 1185–1206.

Higgs, G. (2004). A literature review of the use of GIS-based measures of access to health care services. *Health Services Outcomes Research Methodology, 5*(2), 119–139.

Hill, L. L. (2006). *Georeferencing*. Cambridge, MA: MIT Press.

Hillsman, E., & Rushton, G. (1975). The p-median problem with maximum distance constraints. *Geographical Analysis, 7*(1), 85–89.

Hilton, B., Horan, T., & Schooley, B. (2009). Devising and deploying SafeRoadMaps: Communicating public health issues related to rural and urban road transportation safety through GIS-based mashups, Paper presented at the Transportation Board Annual Meeting, 2009, Paper 09–2334. Retrieved January 5, 2011 from *www.saferoadmaps.org/resources*.

Hirschfield, A., Brown, P. J. B., & Bundred, P. (1995). The spatial analysis of community health services on Wirral using geographic information systems. *Journal of the Operational Research Society, 46*(2), 147–159.

Hiscock, R., Pearce, J., Blakely, T., & Witten, K. (2008). Is neighborhood access to health care provision associated with individual-level utilization and satisfaction? *Health Services Research, 43*(6), 2183–2200.

Holgate, S. T., Devlin, R. B., Wilson, S. J., & Frew, A. J. (2003). Part II: Health subjects exposed to concentrated ambient particles. *Health effects of acute exposure to air pollution*, Research Report 112, pp. 31–50. Boston: Health Effects Institute.

Holgate, S. T., Sandström, T., Frew, A. J., Stenfors, N., Nördenhall, C., Salvi, S., et al. (2003). Part I: Healthy and asthmatic subjects exposed to diesel exhaust. *Health effects of acute exposure to air pollution,* Research Report 112, pp. 1–30. Boston: Health Effects Institute.

Holick, M. F. (2009). Shining light on the vitamin D cancer connection IARC report. *Dermato-Endocrinology, 1*(1), 4–6.

Holman, M. S., Caporale, D. A., Goldberg, J., Lacombe, E., Lubelczyk, C., Rand, P. W., et al. (2004). Anaplasma phagocytophilum, Babesia microti, and Borrelia burgdorferi in Ixodes scapularis, Southern Coastal Maine. *Emerging Infectious Diseases, 10*(4), 744–746.

Holmes, E. C., & Twiddy, S. S. (2003). The origin, emergence and evolutionary genetics of dengue virus. *Infection, Genetics and Evolution, 3*(1), 19–28.

Holt, J. B., & Lo, C. P. (2008). The geography of mortality in the Atlanta metropolitan area. *Computers, Environment, and Urban Systems, 32*(2), 149–164.

Holt, J. B., Lo, C. P., & Hodler, T. W. (2004). Dasymetric estimation of population density and areal interpolation of census data. *Cartography and Geographic Information Science, 31*(2), 103–121.

Horner, M. W., & Mascarenhas, A. K. (2007). Analyzing location-based accessibility to dental services: An Ohio case study. *Journal of Public Health Dentistry, 67*(2), 113–118.

Horvath, R. J. (2006). The "Detroit Geographical Expedition and Institute" Experience. *Antipode, 3*(1), 73–85.

Hotez, P. J., Fenwick, A., & Kjetland, E. F. (2009). Africa's 32 cents solution for HIV/AIDS. *PLOS Neglected Tropical Diseases, 3*(5), e430.

Houghton, F. (2004). Looking on from the sidelines: The use of geographical information systems in health boards in Ireland 2003. *Area* 36(1), 81–83.

Hrushesky, W. J. (1985). Circadian timing of cancer chemotherapy. *Science, 228*(4965), 73–75.

Hsueh, Y., & Rajagopal, R. (1988). Modeling ground water quality sampling decisions. *Ground Water Monitoring Review, 8*(4), 121–134.

Huber, J. T., Boorkman, J., & Blackwell, J. (Eds.). (2008). *Introduction to reference sources in the health sciences* (5th ed.). New York: Neal-Schuman Publishers.

Hunter, J. M. (1974). The challenge of medical geography. In J. M. Hunter (Ed.), *The Geography of Health and Disease, Studies in Geography No. 6* (pp. 1–31). Chapel Hill: University of North Carolina Chapel Hill.

Hupert, N., Xiong, W., & Mushlin, A. (2008). The virtue of virtuality: The promise of agent-based epidemic modeling. *Translational Research, 151*(6), 273–274.

Huxhold, W., & Levinsohn, A. (1995). *Managing GIS projects*. New York: Oxford University Press.

Hyndman, J., & Holman, C.D. (2001). Accessibility and spatial distribution of general practice services in an Australian city by levels of social disadvantage. *Social Science and Medicine, 53*(12), 1599–1609.

Iglehart, J. K. (2008). Grassroots activism and the pursuit of an expanded physician supply. *The New England Journal of Medicine, 358*(16), 1741–1749.

Ihrig, M. M., Shalat, S. L., & Baynes, C. (1998). A hospital-based case-control study of stillbirths and environmental exposure to arsenic using an atmospheric dispersion model and a geographical information system. *Epidemiology, 9*(3), 290–294.

Iliffe, J. C. (2000). *Datums and map projections for remote sensing, GIS and surveying*. Caithness, Scotland: Whittles Publishing.

Indirli, M. (2009). Organization of a geographic information system (GIS) database on natural hazards and structural vulnerability for the historic center of San Guiliano di Puglia (Italy) and the city of Valparaiso (Chile). *International Journal of Architectural Heritage, 3*(4), 276–315.

Institute of Medicine (IOM) and National Research Council. (2008). *Achieving sustainable global capacity for surveillance and response to emerging diseases of zoonotic origin: Workshop summary*. Washington, DC: National Academies Press.

International Cartographic Association/Association Cartographique International. (2010). National members. Retrieved January 5, 2011, from *www.icaci.org/membership-types*.

International Organization for Standardization (ISO). (2003). *Geographic information—metadata* (ISO 19115). Retrieved January 5, 2011, from *www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26020*.

International Society for Infectious Diseases. (2011). ProMed-mail. Retrieved January 5, 2011, from *www.promedmail.org/pls/otn/f?p=2400:1000*.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. Oxford, UK: Oxford University Press.

Istre, G. (1992). Disease surveillance at the state and local levels. In W. Halperin & E. Baker (Eds.), *Public health surveillance* (pp. 42–55). New York: Van Nostrand Reinhold.

Jacquez, G. M., Kaufmann, A., Meliker, J., Goovaerts, P., AvRuskin, G., & Nriagu, J.

(2005). Global, local and focused clustering for case-control data with residential histories. *Environmental Health, 4*, 4.

Jacquez, G. M., Meliker, J. R., AvRuskin, G. A., Goovaerts, P., Kaufmann, A., Wilson, M. L., et al. (2006). Case-control geographic clustering for residential histories accounting for risk factors and covariates. *International Journal of Health Geographics, 5*, 32.

Jacquez, G. M., Meliker, J., & Kaufman, A. (2007). In search of induction and latency periods: Space-time interaction accounting for residential mobility, risk factors and covariates. *International Journal of Health Geographics, 6*, 35.

Jankowski, P., & Ewart, G. (1996). Spatial decision support system for health practitioners: Selecting a location for rural health practice. *Geographical Systems, 3*(2), 279–299.

Jankowski, P., & Nyerges, T. (2003). Toward a framework for geographic research on geographic-information supported participatory decision-making. *URISA Journal, 15*(APA I), 9–17.

Jansen, A., Frank, C., Koch, J., & Stark, K. (2008). Surveillance of vector-borne diseases in Germany: Trends and challenges in the view of disease emergence and climate change. *Parasitology Research, 103*(Suppl. 1), S11–S17.

Jarman, B. (1983). Identification of underprivileged areas. *British Medical Journal, 286*(6379), 1705–1708.

Jenks, G. F. (1977). *Optimal data classification for choropleth maps* (Occasional Paper No. 2). Lawrence: Department of Geography, University of Kansas.

Jenks, G. F., & Caspall, F. C. (1971). Error on choroplethic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers, 61*(2), 217–244.

Jennings, J. M., Curriero, F. C., Celentano, D., & Ellen, J. M. (2005). Geographic identification of high gonorrhea transmission areas in Baltimore, Maryland. *American Journal of Epidemiology, 161*(1), 73–80.

Jensen, J. R. (2005). *Introductory digital image processing: A remote sensing perspective* (3rd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Jensen, J. R. (2007). *Remote sensing of the environment: An earth resource perspective* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B, Potoglou, D., Sahsuvaroglu, T., et al. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology, 15*(2), 185–204.

Jerrett, M., & Finkelstein, M. (2005). Geographies of risk in studies linking chronic air pollution exposure to health outcomes. *Journal of Toxicology and Environmental Health, Part A, 68*(13–14), 1207–1242.

Jia, X. Y., Briese, T., Jordan, I., Rambaut, A., Ci, C., Mackenzie, J. S., et al. (1999). Genetic analysis of West Nile New York 1999 encephalitis virus. *Lancet, 354*(9194), 1971–1972.

Johnston, K. M., & Maguire, D. (2007). *Agent-based modeling position statement*. Workshop on Agent-Based Modeling of Complex Spatial Systems, April 14–16, 2007. Retrieved January 5, 2011, from *www.ncgia.ucsb.edu/projects/abmcss/docs/maguire_paper.pdf*.

Jones, C. G., Ostfeld, R. S., Richard, M. P, Schauber, E. M., & Wolff, J. O. (1998). Chain reactions linking acorns to gypsy moth outbreaks and Lyme disease risk. *Science, 279*(13), 1023–1026.

Jones, K., Duncan, C., & Twigg, L. (2004). Evaluating the absolute and relative income hypotheses in an exploratory analysis of deaths in the Health and Lifestyle Survey. In P. Boyle, S. Curtis, E. Graham, & E. Moore (Eds.), *The geography of health inequalities in the developed world: Views from Britain and North America* (pp. 219–244). Aldershot, UK: Ashgate.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & et al. (2008). Global trends in emerging infectious disease. *Nature, 451*(7181), 990–993.

Jones, N., Bloomfield, A., Rainger, W., & Taylor, P. (1998). Epidemiology and control of the 1997 measles epidemic in Auckland. *The New Zealand Public Health Report, 5*(8), 57–60.

Jones, R. C., Liberatore, M., Fernandez, J. R., & Gerber, S. I. (2006). Use of a prospective space-time scan statistic to prioritize shigellosis case investigations in an urban jurisdiction. *Public Health Reports, 121*(2), 133–139.

Jones, R. L., Homa, D. M., Meyer, P. A., Brody, D. J., Caldwell, K. L., Pirkle, J. L., et al. (2009). Trends in blood lead levels and blood lead testing among US children aged 1 to 5 years, 1988 2004. *Pediatrics, 123*(3), e376–e385.

Joseph, A., & Phillips, D. (1984). *Accessibility and utilization: Geograhical perspectives on health care delivery.* New York: Harper and Row.

Joy, K. P. (1994). *Estimating child exposure to electromagnetic fields associated with transmission lines in Hartford County, Connecticut.* Unpublished Master's Thesis. Department of Geography, University of Connecticut.

Juarez, P. D., Robinson, P. L., & Matthews-Juarez, P. (2002). 100% access, zero health disparities, and GIS: An improved methodology for designating health professions shortage areas. *Journal of Health and Social Policy, 16*(1–2), 155–167.

Justice and Environment. (2011). Access to Justice in Environmental Matters under the Aarhus Convention. Retrieved January 5, 2011, from *www.justiceandenvironment.org/_files/file/J&E-Aarhus-position-paper-2010–05–24.pdf.*

Kainz, W. (1995). Logical consistency. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 109–135). Oxford, UK: Elsevier Science.

Kairouz, S., Glicksman, L., Demers, A., & Adlaf, E. M. (2002). For all these reasons, I do … drink: A multilevel analysis of contextual reasons for drinking among Canadian undergraduates. *Journal of Studies on Alcohol, 63*(5), 600–608.

Kairouz, S., & Greenfield, T. K. (2007). A comparative multi-level analysis of contextual drinking in American and Canadian adults. *Addiction, 102*(1), 71–80.

Kamadjeu, R. (2009). Tracking the polio virus down the Congo River: A case study on the use of Google Earth in public health planning and mapping. *International Journal of Health Geographics, 8*, 4.

Kamphuis, C. B. M., Giskes, K., Kavanagh, A. M., Thornton, L. E., Thomas, L. R., van Lenthe, F. J., et al. (2008). Area variation in recreational cycling in Melbourne: A compositional or contextual effect? *Journal of Epidemiology and Community Health, 62*(10), 890–898.

Kaplan, E., & Hegarty, C. (Eds.). (2006). *Understanding GPS: Principles and applications* (2nd ed.). Norwood, MA: Artech House.

Kaufman, J. S. (2001). Commentary: Social epidemiology? Way! *International Journal of Epidemiology, 30*(1), 48–49.

Kawachi, I., & Kennedy, B. P. (1999). Income inequality and health: Pathways and mechanisms. *Health Services Research, 34*(1), 215–227.

Kearns, R. A. (1993). Place and health: Towards a reformed medical geography. *The Professional Geographer, 45*(2), 139–147.

Keating, J. (1993). *The geopositioning selection guide for resource management*. Cheyenne, WY: Bureau of Land Management.

Keeling, M., & White, P. (2011). Targeting vaccinations against novel infections: Risk, age and spatial structure for pandemic influenza in Great Britain. *Journal of the Royal Society Interface, 8*(58), 661–670.

Keister, K. (1993). Charts of change. *Historic Preservation, 45*(3), 42–48, 91–92.

Kelsey, J. L., Whittemore, A. S., Evans, A. S., & Thompson, W. D. (1996). *Methods in observational epidemiology* (2nd ed.). New York: Oxford University Press.

Kennedy, H. (1999). Mapping health care networks. *ArcUser, 2*(2), 30–31.

Kessler, B. L. (1992). Glossary of GIS terms. *Journal of Forestry, 90*(11), 37–45.

Khan, O., Davenhall, W., Ali, M., Castillo-Salgado, C., Vazquez-Prokopec, G., Kitron, U., et al. (2010). Geographical information systems and tropical medicine. *Annals of Tropical Medicine and Parasitology, 1104*(4), 303–318.

Khumawala, B. M. (1973). An efficient algorithm for the *p*-median problem with maximum distance constraints. *Geographical Analysis, 5*(4), 309–321.

Killea, M. E., Swei, A., Lane, R. S., Briggs, C. J., & Ostfeld, R. S. (2008). Spatial dynamics of Lyme disease: A review. *EcoHealth, 5*(2), 167–195.

Kim, A. A., Martinez, A. N., Klausner, J. D., Goldenson, J., Kent, C., Liska, S., et al. (2008). Use of sentinel surveillance and geographic information systems to monitor trends in HIV prevalence, incidence, and related risk behavior among women undergoing syphilis screening in a jail setting. *Journal of Urban Health, 86*(1), 79–92.

King, L. J. (1984). *Central place theory*. Scientific Geography Series Vol. 1. Beverly Hills-Newbury Park, CA: Sage.

Kingham, S., A. Gatrell, & Rowlingson, B. (1995). Testing for clustering of health events within a geographical information systems framework. *Environment and Planning A, 27*(5), 809–821.

Kirby, R. (1996). Toward congruence between theory and practice in small area analysis and local public health data. *Statistics in Medicine, 15*(17–18), 1859–1866.

Kistemann, T., Munzinger, A., & Dangendorf, F. (2002). Spatial patterns of tuberculosis incidence in Cologne (Germany). *Social Science and Medicine, 55*(1), 7–19.

Kitagawa, E. M. (1966). Theoretical considerations in the selection of a mortality index, and some empirical comparisons. *Human Biology, 38*(3), 293–308.

Kitchen, R., & Dodge, M. (2007). Rethinking Maps. *Progress in Human Geography, 31*(3), 331–344.

Kitron, U., Clennon, J. A., Cecere, M. C., Gürtler, R. E., King, C. H., & Vazquez-Prokopec, G. (2006). Upscale or downscale: Applications of fine scale remotely sensed data to Chagas disease in Argentina and schistosomiasis in Kenya. *Geospatial Health, 1*(1), 49–58.

Kitron, U., Jones, C. J., Bouseman, J. K., Nelson, J. A., & Baumgartner, D. L. (1992). Spatial analysis of the distribution of *Ixodes dammini* (Acari: Ixodidae) on white-tailed deer in Ogle County, Illinois. *Journal of Medical Entomology, 29*(2), 259–266.

Kline, J. A., & Walthall, J. D. H. (2010). Patient Protection and Affordable Care Act of 2010: Summary, analysis, and opportunities for advocacy for the academic emergency physician. *Academic Emergency Medicine, 17*(7), e69–e74.

Knigge, L., & Cope, M. (2006). Grounded visualization: Integrating the analysis of qualitative and quantitative data through grounded theory and visualization. *Environment and Planning A, 38*(11), 2021–2037.

Knox, P. (1978). The intraurban ecology of primary care: Patterns of accessibility and their policy implications. *Environment and Planning A, 10*(4), 415–435.

Kohli, S., Noorlind Brage, H., & Löfman, O. (2000). Childhood leukaemia in areas with different radon levels: A spatial and temporal analysis using GIS. *Journal of Epidemiology and Community Health, 54*(11), 822–826.

Kohli, S., Sahlen, K., Lofman, O., Sivertun, A., Foldevi, M., Trell, E., et al. (1997). Individuals living in areas with high background radon: A GIS method to identify populations at risk. *Computer Methods and Programs in Biomedicine, 53*(2), 105–112.

Kotton, C. N. (2007). Zoonoses in solid-organ and hematopoietic stem cell transplant recipients. *Clinical Infectious Disease, 44*(6), 857–866.

Kovner, A. R., & Knickman, J. R. (Eds.). (2008). *Jonas & Kovner's health care delivery in the United States* (9th ed.). New York: Springer.

Kraak, M.-J. (1999). Visualising spatial distribution. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems: Vol. 1. Principles and technical issues* (2nd ed., pp. 157–173). New York: Wiley.

Kraak, M.-J. (2004). The role of the map in a web-GIS environment. *Journal of Geographical Systems, 6*(2), 83–93.

Kraak, M.-J., & Brown, A. (Eds.). (2000). *Web cartography: Development and prospects.* London: Taylor & Francis.

Krebs, J. R., Anderson, R. M., Clutton-Brock, T., Donnelly, C. A., Frost, S., Morrison, W. I., et al. (1998). Badgers and bovine TB: Conflicts between conservation and health. *Science, 279*(5352), 817–818.

Krebs, J. W., Strine, T. W., & Childs, J. E. (1993). Rabies surveillance in the United States during 1992. *Journal of the American Veterinary Association, 203*(12), 1718–1726.

Krebs, J. W., Wilson, M. L., & Childs, J. E. (1995). Rabies—Epidemiology, prevention, and future research. *Journal of Mammalogy, 76*(3), 681–694.

Krieger, N. (1994). Epidemiology and the web of causation: Has anyone seen the spider? *Social Science and Medicine, 39*(7), 887–903.

Krieger, N. (2001). Theories for social epidemiology in the 21st century: An ecosocial perspective. *International Journal of Epidemiology, 30*(4), 668–677.

Krieger, N. (2002). A glossary for social epidemiology, Part I. *Epidemiological Bulletin/PAHO, 23*(1), 7–11.

Krieger, N., Chen, J. T., Waterman, P. D., Rehkopf, D. H., & Subramanian, S. V. (2003). Race/ethnicity, gender, and monitoring socioeconomic gradients in health: A comparison of area-based socioeconomic measures—The Public Health Disparities Geocoding Project. *American Journal of Public Health, 93*(10), 1655–1671.

Krieger, N., Waterman, P., Chen, J. T., Soobader, M-J., Subramanian, S. V., & Carons, R. (2002). Zip code caveat: Bias due to spatiotemporal mismatches between zip codes and US Census-defined geographic areas—The Public Health Geocoding Disparities Project. *American Journal of Public Health, 92*(7), 1100–1102.

Krieger, N., & Zierler, S. (1996). What explains the public's health—a call for epidemiologic theory. *Epidemiology, 7*(1), 107–109.

Krieger, N., & Zierler, S. (1997). The need for epidemiologic theory. *Epidemiology, 8*(2), 212–214.

Kronick, R., Goodman, D. C., Wennberg, J., & Wagner, E. (1993). Special report: The markeplace in health care reform: The demographic limitations of managed competition. *The New England Journal of Medicine, 328*(2), 148–152.

Kruger, D., Brady, J., & Shirey, L. (2008). Using GIS to facilitate community-based planning of diabetes intervention efforts. *Health Promotion Practice, 9*(1), 76–81.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods, 25*(6), 1481–1496.

Kulldorff, M. (1998). Statistical methods for spatial epidemiology. In A. Gatrell & M. Loytonen (Eds.), *GIS and Health* (pp. 49–62). London: Taylor & Francis.

Kulldorff, M. (2010). SaTScan User Guide for Version 9.1. Retrieved January 5, 2011, from *www.satscan.org*.

Kumar, N. (2004). Frequency histogram legend in the choropleth map: A substitute to traditional legends. *Cartography and Geographic Information Science, 31*(4), 217–236.

Kunst, A. E., Groenhof, F., Mackenbach, J. P., & EU Working Group on Socioeconomic Inequalities in Health. (1998). Occupational class and cause specific mortality in middle aged men in 11 European countries: Comparison of population based studies. *British Medical Journal, 316*(7145), 1636–1641.

Kwan, M.-P. (1998). Space–time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis, 30*(3), 191–216.

Kwan, M. P. (1999). Gender and individual access to urban opportunities: A study using space–time measures. *The Professional Geographer, 51*(2), 210–227.

Kwan, M. P. (2000). Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration of a large data set. *Transportation Research Part C-Emerging Technologies, 8*(1–6), 185–203.

Lacroix, E-M., & Backus, J. E. B. (2006). Organizing electronic information to serve the needs of health practitioners and consumers. *Library Trends, 54*(4), 605–617.

Lake, I. R., Harrison, F. C., Chalmers, R. M., Bentham, G., Nichols, G., Hunter, P. R., et al. (2007). Case-control study of environmental and social factors influencing cryptosporidiosis. *European Journal of Epidemiology, 22*(11), 805–811.

Lam, D., & Poropatich, R. K. (2008). Telemedicine deployments within NATO military forces: A data analysis of current and projected capabilities. *Telemedicine and e-Health, 14*(9), 946–951.

Lam, N. S-N. (1983). Spatial interpolation methods: A review. *The American Cartographer, 10*(2), 129–149.

Lampin-Maillet, C., Jappiot, M., Long, M., Morge, D., & Ferrier, J. P. (2009). Characterization and mapping of dwelling types for forest fire prevention. *Computers Environment and Urban Systems, 33*(3), 224–232.

Lancelot, R., Poncon, N., Hendrickx, G., & Fontenille, D. (2009). Environmental changes and emergence of vector-borne diseases in Europe: Improving risk surveillance and management. *Bulletin de l'Académie Vétérinaire de France, 162*(1), 89–96.

Lanciotti, R. S., Roehrig, J. T., Deubel, V., Smith, J., Parker, M., Steele, K., et al. (1999). Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science, 286*(5448), 2333–2337.

Landesman, L. Y. (2005). *Public health management of disasters: The practice guide* (2nd ed.). Washington, DC: American Public Health Association.

Langford, I. (1994). Using empirical Bayes estimates in the geographical analysis of disease risk. *Area, 26*(2), 142–149.

Langford, M., & Higgs, G. (2006). Measuring potential access to primary healthcare services: The influence of alternative spatial representations of population. *Professional Geographer, 58*(3), 294–306.

Langford, M., Higgs, G., Radcliffe, J., & White, S. (2008). Urban population distribution models and service accessibility estimation. *Computers Environment and Urban Systems, 32*(1), 66–80.

Larsen, S. (1999). Community–university partnerships using Internet based GIS to communicate air and water quality conditions. In E. Kendy (Ed.), Science into policy: Water in the public realm (Technical Publication Series TPS-99–3, pp. 147–152). Herndon, VA: American Water Resources Association.

Larsgaard, M. L. (2005). Metaloging of digital geospatial data. *The Cartographic Journal, 42*(3), 231–237.

Laurini, R. (2001). *Information systems for urban planning: A hypermedia co-operative approach*. London: Taylor & Francis.

Laurini, R., & Thompson, D. (1992). *Fundamentals of spatial information systems*. San Diego, CA: Academic Press.

Laveissiere, C., & Meda, A. H. (1999). Incidence of sleeping sickness and settlement density in the forests of Cote d'Ivoire: Predicting high risk areas for the establishment of an efficient surveillance network. *Tropical Medicine and International Health, 4*(3), 199–206.

Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G., & Shmoys, D. B. (Eds.). (1987). *The traveling salesman problem: A guided tour of combinatorial optimization.* Chichester, UK: Wiley.

Lawson, A. B. (2006). *Statistical methods in spatial epidemiology* (2nd ed.). Chichester, UK: Wiley.

Lawson, J., & Floyd, J. (1996). The future of epidemiology: A humanist response. *American Journal of Public Health, 86*(7), 1029.

Leatherdale, S. T., & Strath, J. M. (2007). Tobacco retailer density surrounding schools and cigarette access behaviors among underage smoking students. *Annals of Behavioral Medicine, 33*(1), 105–111.

Lebel, A., Pampalon, R., & Villenueve, P. Y. (2007). A multi-perspective approach for defining neighbourhood units in the context of a study on health inequalities in the Quebec City region. *International Journal of Health Geographics, 6*, 27.

Lee, B. Y., Bedford, V. L., Roberts, M. S., & Carley, K. M. (2008). Virtual epidemic in a virtual city: Simulating the spread of influenza in a U. S. metropolitan area. *Translational Research, 151*(6), 275–287.

Lemen, R. M. (1999). The evolution of topographic mapping in the U.S. Geological Survey's National Mapping Program, U.S. Geological Survey Open-File Report 99–386, Retrieved January 5, 2011, from *pubs.usgs.gov/of/1999/of99–386/lemen.html*.

Lemon, S. M., Sparling, P. F., Hamburg, M. A., Relman, D. A., Choffnes, E. R., & Mack, A. (Eds.). (2008). *Vector-borne diseases: Understanding the environmental, human health, and ecological connections: Workshop summary* (pp. 1–40). Washington, DC: National Academies Press.

Lepofsky, M., Abkowitz, M., & Cheng, P. (1993). Transportation hazard analysis in integrated GIS environment. *Journal of Transportation Engineering, 119*(2), 239–254.

Lewis-Michl, E. L., Melius, J. M., Kallenbach, L. R., Ju, C. L., Talbot, T. O., Orr, M. F., et al. (1996). Breast cancer risk and residence near industry or traffic in Nassau and Suffolk counties, Long Island, New York. *Environmental Health Perspectives, 51*(4), 255–265.

Li, L., Bian, L., Yakob, L., Zhou, G., & Yan, G. (2009). Temporal and spatial stability of Anopheles gambiae larval habitat distribution in Western Kenya highlands. *International Journal of Health Geographics, 8*, 70.

Lichter, D. T., Parisi, D., Grice, S. M., & Taquino, M. (2007). Municipal underbounding:

Annexation and racial exclusion in small Southern towns. *Rural Sociology, 72*(1), 47–68.

Lin, G., Allan, D., & Penning, M. (2002). Examining distance effects on hospitalizations using GIS: A study of three health regions in British Columbia, Canada. *Environment and Planning A, 34*(11), 2037–2053.

Linder, S. H., Marko, D., & Sexton, K. (2008). Cumulative cancer risk from air pollution in Houston: Disparities in risk burden and social disadvantage. *Environmental Science and Technology, 42*(12), 4312–4322.

Lipowicz, A. (2006). Map the money: GIS providers face down advances by Microsoft, Google. *Washington Technology, 21*(19), 10, 12.

List, G. F., & Turnquist, M. A. (1998). Routing and emergency-response team siting for high-level radioactive waste shipments. *IEEE Transactions on Engineering Management, 45*(2), 141–152.

Litt, J. S., & Burke, T. A. (2002). Uncovering the historic environmental hazards of urban brownfields. *Journal of Urban Health: Bulletin of the New York Academy of Medicine, 79*(4), 464–481.

LizardTech. (2004). *MrSID technology primer*. Retrieved January 5, 2011, from *www.lizardtech.com/files/geo/techinfo/MrSID_Tech_Primer.pdf.*

Llobera, M. (2003). Extending GIS-based visual analysis: The concept of visual landscapes. *International Journal of Geographical Information Science, 17*(1), 25–48.

Loader, C. R. (1999). Bandwidth selection: Classical or plug-in? *Annals of Statistics, 27*(2), 415–438.

Loaiciga, H. A. (1989). An optimization approach for groundwater quality monitoring network design. *Water Resources Research, 25*(8), 1771–1782.

Long, S. K., & Stockley, K. (2010). Sustaining health reform in a recession: an update on Massachusetts as of Fall 2009. *Health Affairs, 29*(6), 1234–1241.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2001). *Geographic information systems and scien*ce. Chichester, UK: Wiley.

Loomis, D., Schulman, M. D., Bailer, A. J., Stainback, K., Wheeler, M., Richardson, D. B., et al. (2009). Political economy of US states and rates of fatal occupational injury. *American Journal of Public Health, 99*(8), 1400–1408.

Loomis, W. F. (1970). Rickets. *Scientific American, 223*(6), 76–91.

Lovasi, G. S., Weiss, J. C., Hoskins, R., Whitsel, E. A., Rice, K., Erickson, C. F., et al. (2007). Comparing a single-stage geocoding method to a multi-stage geocoding method: How much and where do they disagree? *International Journal of Health Geographics, 6*, 12.

Lovett, A., Haynes, R., Sunnenberg, G., & Gale, S. (2002). Car travel time and accessibility by bus to general practitioner services: A study using patient registers and GIS. *Social Science and Medicine, 55*(1), 97–111.

Lowe, J., & Sen, A. (1996). Gravity model applications in health planning: Analysis of an urban hospital market. *Journal of Regional Science, 36*(3), 437–461.

Lowry, J. H., Jr., Miller, H. J., & Hepner, G. F. (1995). A GIS-based sensitivity analysis of community vulnerability to hazardous contaminants on the Mexico/U.S. border. *Photogrammetric Engineering & Remote Sensing, 61*(11), 1347–1359.

Löytönen, M., & Sabel, C. E. (2004). Mobile phone positioning systems and the accessibility of health services. In R. Maheswaran & M. Craglia (Eds.), *GIS in public health practice* (pp. 277–286). Boca Raton, FL: CRC Press.

Luke, D. A. (2005). Getting the big picture in community science: Methods that capture context. *American Journal of Community Psychology, 35*(3/4), 185–200.

Lundstrom, J. O. (1999). Mosquito-borne viruses in western Europe: A review. *Journal of Vector Ecology, 24*(1), 1–39.
Luo, W. (2004). Using a GIS-based floating catchment method to assess areas with shortage of physicians. *Health and Place, 10*(1), 1–11.
Luo, W., & Qi, L. (2009). An enhanced two-step floating catchment area method for measuring spatial accessibility to primary care physicians. *Health and Place, 15*(4), 1100–1107.
Luo, W., & Wang, F. (2003). Measures of spatial accessibility to healthcare in a GIS environment: Synthesis and a case study in the Chicago region. *Environment and Planning B, 30*(6), 865–884.
Luzzader-Beach, S. (1995). Evaluating the effects of spatial monitoring policy on groundwater quality portrayal. *Environmental Management, 19*(3), 383–392.
Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.
Lytle, L. A. (2009). Measuring the food environment: State of the science. *American Journal of Preventive Medicine, 36*(4S), S134–S144.
Maantay, J. A., Maroko, A. R., & Porter-Morgan, H. (2008). Research note—A new method for mapping population and understanding the spatial dynamics of disease in urban areas: Asthma in the Bronx, New York. *Urban Geography, 29*(7), 724–738.
Macdonald, K. I. (2002). Commentary: Social epidemiology. A way? *International Journal of Epidemiology, 30*(1), 46–47.
MacDowell, M., Glasser, M., Fitts, M., Nielsen, K., & Hunsaker, M. (2010). A national view of rural health workforce issues in the USA. *Rural Remote Health, 10*(3), 1531.
MacEachren, A. M. (1994). *Some truth with maps: A primer on symbolization and design*. Washington, DC: Association of American Geographers.
MacEachren, A. (1995). *How maps work: Representation, visualization and design.* New York: Guilford Press.
MacEachren, A. M., Buttenfield, B., Campbell, J., DiBiase, D., & Monmonier, M. (1992). Visualization. In R. Abler, M. Marcus, & J. Olson (Eds.), *Geography's inner worlds: Pervasive themes in contemporary American geography* (pp. 99–137). New Brunswick, NJ: Rutgers University Press.
MacEachren, A. M., Crawford, S., Akella, M., & Lengerich, E. (2008). Design and implementation of a model, web-based, GIS-enabled cancer atlas. *The Cartographic Journal, 45*(4), 246–260.
Macer, D. (2005). Ethical, legal and social issues of genetically modifying insect vectors for public health. *Insect Biochemistry and Molecular Biology, 35*(7), 649–660.
Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: How can we conceptualise operationalise and measure them? *Social Science and Medicine, 55*(1), 125–139
Macintyre, S., Hiscock, R., Ellaway, A., & Kearns, A. (2004). The fallacy of a range of household and area-based indicators of material resources in the geography of health inequalities. In P. Boyle, S. Curtis, E. Graham, & E. Moore (Eds.), *The geography of health inequalities in the developed world: Views from Britain and North America* (pp. 309–329). Aldershot, UK: Ashgate.
Macintyre, S., & Hunt, K. (1997). Socio-economic position, gender, and health: How do they interact? *Journal of Health Psychology, 2*(3), 315–334.
Macintyre, S., Macdonald, L., & Ellaway, A. (2008). Do poorer people have poorer access to local resources and facilities? The distribution of local resources by area deprivation in Glasgow, Scotland. *Social Science and Medicine, 67*(6), 900–914.
Mackenzie, J. S., Gubler, D. J., & Petersen, L. R. (2004). Emerging flaviviruses: The

spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nature Medicine Supplement, 10*(12), S98–S109.

Maclachlan, J., Jerrett, M., Abernathy, T., Sears, M., & Bunch, M. (2007). Mapping health on the Internet: A new tool for environmental justice and public health research. *Health and Place, 13*(1), 72–86.

Macurek, A. (2010). Tracking the progress of H1N1 swine flu. Retrieved January 5, 2011, from *flutracker.rhizalabs.com*.

Mahoney, R. P. (1991). GIS and utilities. In D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.). *Geographical information systems: Principles and applications, Vol. 2. Applications* (pp. 101–114). Harlow, UK: Longman Scientific & Technical.

Mair, C., Diez Roux, A. V., & Galea, S. (2008). Are neighbourhood characteristics associated with depressive symptoms? A review of evidence. *Journal of Epidemiology and Community Health, 62*(11), 940–946.

Malaria Atlas Project. (2011). MAP: Malaria Atlas Project. Retrieved January 5, 2011, from *www.map.ox.ac.uk*.

Malczewski, J. (1999). *GIS and multicriteria decision analysis.* New York: Wiley.

Maling, D. H. (1992). Coordinate systems and map projections (2nd ed). Oxford, UK: Pergamon Press.

Mandl, K. D., Overhage, M., Wagner, M. M., Lober, W. B., Sebastiani, P., Mostashari, F., et al. (2004). Implementing syndromic surveillance: A practical guide informed by the early experience. *Journal of the American Medical Informatics Association, 11*(2), 141–150.

MAPresso. (2011). MAPresso: Free Java applet for unclassed choropleth maps and cartograms. Retrieved January 5, 2011, from *www.mapresso.com.*

Marks, A. P., Thrall, G. I., & Arno, M. (1992). Siting hospitals to provide cost-effective health care. *Geo Info Systems, 2*(8), 58–66.

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*(Part 3), 421–441.

Martens, P. (2000). Malaria and global warming in perpsective? *Emerging Infectious Diseases, 6*(3), 313–314.

Martin, D., & Williams, H. (1992). Market area analysis and accessibility to primary healthcare centres. *Environment and Planning A, 24*(7), 1009–1019.

Marx, R. W. (1986). The TIGER system: Automating the geographic structure of the United States Census. *Government Publications Review, 13*(2), 181–201.

Masser, I., Campbell, H., & Craglia, M. (Eds.). (1996). *GIS diffusion: The adoption and use of geographical information systems in local government in Europe*. Bristol, PA: Taylor & Francis.

Mason, T. J., McKay, F. W., Hoover, R. N., Blot, W. J., & Fraumeni, Jr., J. F. (1975). *Atlas of cancer mortality for U.S. counties, 1950–1969*. Bethesda, MD: U.S. Department of Health, Education, and Welfare.

Mather, F. J., White, L. E., Langlois, E. C., Shorter, C. F., Swalm, C. M., Shaffer, J. G., et al. (2004). Statistical methods for linking health, exposure, and hazards. *Environmental Health Perspectives, 112*(14), 1440–1445.

Mather, T. N., Nicholson, M. C., Donnelly, E. F., Matyas, B. T. (1996). Entomologic index for human risk of Lyme disease. *American Journal of Epidemiology, 144*(11), 1066–1069.

Matthews, S., Detwiler, J., & Burton, L. (2005). Viewing people and places: Coupling geographic information analysis and ethnographic research. *Cartographica, 40*(4), 75–90.

Mattos, N. M., Meyer-Wegener, K., & Mitschang, B. (1993). Grand tour of concepts for object-orientation from a database point of view. *Data and Knowledge Engineering, 9*(3), 321–352.

Maxcy, K. F. (1926). An epidemiological study of endemic typhus (Brill's disease) in the Southeastern United States with special reference to its mode of transmission. *Public Health Reports, 41*(52), 2967–2995.

Mayer, J. D. (2000). Geography, ecology and emerging infectious diseases. *Social Science and Medicine, 50*(7–8), 937–952.

Mayer, M. (2006). Are we there yet? Distance to care and relative supply among pediatric subspecialties. *Pediatrics, 118*(6), 2313–2321.

Mayhew, L. D., Gibbard, R. W., & Hall, H. (1986). Predicting patient flows and hospital case-mix. *Environment and Planning A, 18*(5), 619–638.

Mazumdar, S., Rushton, G., Smith, B. J., Zimmerman, D. I., & Donham, K. J. (2008). Geocoding accuracy and the recovery of relationships between environmental exposures and health. *International Journal of Health Geographics, 7*, 13.

McCarthy, W. J., Mistry, R., Lu, Y., Patel, M., Zheng, H., & Dietsch, B. (2009). Density of tobacco retailers near schools: Effects on tobacco use among students. *American Journal of Public Health, 99*(11), 2006–2013.

McCormack, E. (1999). Using a GIS to enhance the value of travel diaries. *ITE Journal, 69*(1), 38–39, 40–43.

McCormick, B. H., DeFanti, T. A., & Brown, M. D. (Eds.). (1987). Visualization in scientific computing. *ACM SIGGRAPH Computer Graphics, 21*(6) (Special Issue).

McDonnell, W. F., Kehrl, H. R., Abdul Salaam, S., Ives, P. J., Folinsbee, L. J., Devlin, R. B., et al. (1991). Respiratory response of humans exposed to low levels of ozone for 6.6 hours. *Archives of Environmental Health, 46*(3), 145–150.

McGeady, D., Kujala, J., & Ilvonen, K. (2008). The impact of patient-physician web messaging on healthcare service provision. *International Journal of Medical Informatics, 77*(1), 17–23.

McGrail, M., & Humphreys, J. (2009). The index of rural access: An innovative approach for measuring primary care access. *BMC Health Services Research, 9*, 124.

McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: Assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization, 19*(1), 17–37.

McKone, T. E., Castorina, R., Harnly, M. E., Kuwabara, Y., Eskenazi, B., & Bradman, A. (2007). Merging models and biomonitoring data to characterize sources and pathways of human exposure to organophosphorus pesticides in the Salinas Valley of Califonia. *Environmental Science and Technology, 41*(9), 3233–3240.

McKone, T. E., Ryan, P. B., & Özkaynak, H. (2009). Exposure information in environmental health research: Current opportunities and future directions for particulate matter, ozone, and toxic air pollutants. *Journal of Exposure Science and Environmental Epidemiology, 19*(1), 30–34.

McLafferty, S. (2002). Mapping women's worlds: Knowledge, power and the bounds of GIS. *Gender, Place and Culture, 9*(3), 263–269.

McLafferty, S. (2003). GIS and health care. *Annual Review of Public Health, 24*, 25–42.

McLafferty, S. (2009). Placing substance use: Geographical perspectives on substance use and addiction. In Y. Thomas, D. Richardson, & I. Cheung (Eds.), *Geography and drug addiction* (pp. 1–16). Dordrecht, The Netherlands: Springer.

McLafferty, S. (2010). Placing pandemics: Geographical dimensions of vulnerability and spread. *Eurasian Geography and Economics, 51*(2), 1–19.

McLafferty, S., & Broe, D. (1990). Patient outcomes and regional planning of coronary care services: A location-allocation approach. *Social Science and Medicine, 30*(3), 297–304.

McLafferty, S., & Grady, S. (2004). Prenatal care need and access: A GIS analysis. *Journal of Medical Systems, 28*(3), 321–333.

McLauglin, C., & Boscoe, F. (2007). Effects of randomization methods on statistical inference in disease cluster detection. *Health and Place, 13*(1), 152–163.

McManus, J. (1993). Mapping helps Met track its health care networks. *National Underwriter, 97*(23), 38.

McPherson, K. (2001). Commentary: Epidemiology? Keep it broad and deep. *International Journal of Epidemiology, 30*(1), 48.

Medlock, J. M., Avenell, D., Barrass, I., & Leach, S. (2006). Analysis of the potential for survival and seasonal activity of *Aedes albopictus* (Diptera: Culicidae) in the United Kingdom. *Journal of Vector Ecology, 31*(2), 292–304.

Meersman, S. C., Breen, N., Pickle, L. W., Meissner, H. I., & Simon, P. (2009). Access to mammography screening in a large urban population: A multi-level analysis. *Cancer Causes and Control, 20*(8), 1469–1482.

Mellow, J., Schlager, M. D., & Caplan, J. M. (2008). Using GIS to evaluate post-release prisoner services in Newark, New Jersey. *Journal of Criminal Justice, 36*(5), 416–425.

Melnick, A. L., & Fleming, D. W. (1999). Modern geographic information systems: Promise and pitfalls. *Journal of Public Health Management and Practice, 5*(2), viii–x.

Meltzer, E., Artom, G., Marva, E., Assous, M. V., Rahav, G., & Schwartz, E. (2006). Schistosomiasis among travelers: New aspects of an old disease. *Emerging Infectious Diseases, 12*(11), 1696–1700.

Mennis, J. (2003). Generating surface models of population distribution using dasymetric mapping. *The Professional Geographer, 55*(1), 32–42.

Mennis, J. L., & Jordan, L. (2005). The distribution of environmental equity: Exploring spatial nonstationarity in multivariate models of air toxic releases. *Annals of the Association of American Geographers, 95*(2), 249–268.

Mennis, J., & Mason, M. (2011). People, places, and adolescent substance use: Integrating activity space and social network data for analyzing health behavior. *Annals of the Association of American Geographers, 101*(2), 272–291.

Meyers, J. (1999). GIS in the utilities. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems: Vol. 2. Management issues and applications* (2nd ed., pp. 801–818). New York: Wiley.

Miller, H. J., & Shaw, S.-L. (2001). *Geographic information systems for transportation: Principles and applications*. New York: Oxford University Press.

Miller, L., Lemke, L. D., Xu, X., Molaroni, S. M., You, H., Wheeler, A. J., et al. (2010). Intra-urban correlation and spatial variability of air toxics across an international airshed in Detroit, Michigan (USA) and Windsor, Ontario (Canada). *Atmospheric Environment, 44*(9), 1162–1174.

Mitka, M. (2009). Health care workers, others offer ideas on health care reform to Obama team. *Journal of the American Medical Association, 301*(9), 922–923.

Moellering, H. (Ed.). (1987). *A draft proposed standard for digital cartographic data* (American Congress on Surveying and Mapping Report #8). Bethesda, MD:

National Committee for Digital Cartographic Standards, American Congress on Surveying and Mapping.
Moellering, H. (Ed.). (1991). *Spatial database transfer standards: Current international status*. Oxford, UK: Elsevier Science.
Moffitt, F. H., & Bossler, J. D. (1998). *Surveying* (10th ed.). Menlo Park, CA: Addison-Wesley.
Mohan, J. (1983). Location-allocation models, social science and health service planning: An example from North East England. *Social Science and Medicine, 17*(8), 493–499.
Monmonier, M. (1996). *How to lie with maps* (2nd ed.). Chicago: University of Chicago Press.
Moochhala, S. M., Shahi, G. S., & Cote, I. L. (1997). The role of epidemiology, controlled clinical studies, and toxicology in defining environmental risks. In G. S. Shahi, B. S. Levy, A. Binger, T. Kjellstrom, & R. Lawrence (Eds.), *International perspectives on environment, development, and health: Toward a sustainable world* (pp. 341–352). New York: Springer.
Moon, G., & Gillespie, R. (Eds.). (1995). *Society and health: An introduction to social science for health professionals*. London: Routledge.
Moore, J. E., Sandquist, G. M., & Slaughter, D. M. (1994). A route-specific system for risk assessment of radioactive materials transportation accidents. *Nuclear Technology, 112*(1), 63–78.
Moore, M., Gould, P., & Keary, B. S. (2003). Global urbanization and impact on health. *International Journal of Hygiene and Environmental Health, 206*(4–5), 269–278.
Morello-Frosch, R., Brody, J. G., Brown, P., Altman, R. G., Rudel, R. A., & Perez, C. (2009). Toxic ignorance and right-to-know in biomonitoring results communication: A survey of scientists and study participants. *Environmental Health, 8*, 6.
Morello-Frosch, R., & Jesdale, B. M. (2006). Separate and unequal: Residential segregation and estimated cancer risks associated with ambient air toxics in U.S. metropolitan areas. *Environmental Health Perspectives*, 114(3), 386–393.
Moreno, M. A., Ralston, J. D., & Grossman, D. C. (2009). Adolescent access to online health services: Perils and promise. *Journal of Adolescent Health, 44*(3), 244–251.
Morens, D. M., Folkers, G. K., & Fauci, A. S. (2004) The challenge of emerging and reemerging infectious diseases. *Nature, 430*(6996), 242–249.
Morland, K., Diez Roux, A. V., & Wing, S. (2006). Supermarkets, other food stores, and obesity: The Atherosclerosis Risk in Communities Study. *American Journal of Preventive Medicine, 30*(4), 333–339.
Morrill, R., & Earickson, R. (1968). Hospital variation and patient travel distances. *Inquiry, 5*(4), 26–34.
Morrill, R. L., & Symons, J. (1977). Efficiency and equity aspects of optimum location. *Geographical Analysis, 9*(3), 215–225.
Morrison, D., Alexander, D., Fisk, J., & McGuire, J. (1999). Improving delivery of health and community services to welfare recipients, Columbia, SC, 1997. *Journal of Public Health Management and Practice, 5*(2), 49–50.
Morrison, J. L. (1995). Spatial data quality. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 1–12). Oxford, UK: Elsevier Science.
Morrow, B. H. (1999). Identifying and mapping community vulnerability. *Disasters, 23*(1), 1–18.
Murray, C. J. L., Kulkarni, S. C., Michaud, C., Tomijima, N., Bulzacchelli, M. T., Iandiorio, T. J., et al. (2006). Eight Americas: Investigating mortality disparities across

races, counties, and race-counties in the United States. *PLoS Medicine, 3*(9), 1513–1524.

Murray, C. J. L., & Lopez, A. D. (Eds.). (1996). *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard School of Public Health on behalf of the World Health Organization and the World Bank.

Nadelman, R. B., Horowitz, H. W., Hsieh, T-C.,Wu, J. M., Aguero-Rosenfeld, M. E., Schwartz, I., et al. (1997). Simultaneous human granulocytic ehrlichiosis and Lyme borreliosis. *New England Journal of Medicine, 37*(1), 27–30.

Nair, S., Richardson, E. A., Thompson, W. R. E., Shortt, N. K., & Stewart, K. J. (2009). The influence of geography on uptake of plastic surgery services—analysis based on bilateral breast reduction data. *Journal of Plastic Reconstructive and Aesthetic Surgery, 63*(4), 666–672.

Napier, B. (2003). National geospatial preparedness for all-hazard emergency management. *Photogrammetric Engineering and Remote Sensing, 69*(10), 1137–1139.

Nasca, P. C. (1997). Current problems that are likely to affect the future of epidemiology. *American Journal of Epidemiology, 146*(11), 907–911.

National Academy of Sciences, Mapping Science Committee. (1995). *A data foundation for the National Spatial Data Infrastructure*. Washington, DC: National Academy Press.

nationalatlas.gov. (2010). The Public Land Survey System (PLSS). Retrieved January 5, 2011, from *www.nationalatlas.gov/articles/boundaries/a_plss.html*.

National Cancer Institute. (2011a). Surveillance Epidemiology and End Results (SEER). Retrieved January 5 2011, from *seer.cancer.gov*.

National Cancer Institute. (2011b). Geographic Information System for Breast Cancer Studies on Long Island. Retrieved January 5, 2011, from *li-gis.cancer.gov*.

National Cancer Institute. (2011c). Cancer mortality maps and graphs. Retrieved January 5, 2011, from *www3.cancer.gov/atlasplus*.

National Center for Health Statistics. (1979). *Proceedings of the 1976 Workshop on Automated Cartography and Epidemiology* (DHEW Publication No. [PHS] 79–1254). Hyattsville, MD: Department of Health, Education, and Welfare.

National Center for Health Statistics. (1995). *Proceedings, 25th Public Health Conference on Records and Statistics and the National Committee on Vital and Health Statistics 45th Anniversary Symposium* (DHHS Publication No. [PHS] 96–1214. Hyattsville, MD: Department of Health and Human Services.

National Center for Health Statistics. (2006a). *CDC GIS resources*. Retrieved January 5, 2011, from *www.cdc.gov/gis/resources.htm*.

National Center for Health Statistics. (2006b). Health, United States, 2006 with chartbook on trends in the health of Americans. Retrieved January 5, 2011, from *www.cdc.gov/nchs/data/hus/hus06.pdf*.

National Institute of Standards and Technology. (1994). *Federal information processing standard publication 173 (Spatial data transfer standard part I, version I.I)*. Gaithersburg, MD: Author.

National Library of Medicine. (2010). TOXMAP®: Environmental health e-maps. Retrieved January 5, 2011, from *toxmap.nlm.nih.gov/toxmap/main/index.jsp*.

National Research Council Panel on Distributed Geolibraries. (1999). *Distributed geolibraries: Spatial information resources*. Washington, DC: National Academy Press.

Natural Resources Canada. (2009). The Atlas of Canada. Retrieved January 5, 2011, from *atlas.nrcan.gc.ca/site/english/maps/index.html*.

National Science and Technology Council, Office of Science and Technology Policy, Office of the President of the United States. (1995). *Infectious disease: A global health threat*. Washington, DC: U.S. Government Printing Office.

Navaie-Waliser, M., Spriggs, A., & Feldman, P. (2002). Informal caregiving: Differential experiences by gender. *Medical Care, 40*(12), 1249–1259.

Needham, L. L., Barr, D. B., & Calafat, A. M. (2004). Characterizing children's exposures: Beyond NHANES. *NeuroToxicology, 26*(4), 547–533.

Nelson, J. D., & Ward, R. C. (1981). Statistical considerations and sampling techniques for ground-water quality monitoring. *Ground Water, 19*(6), 617–625.

Neumayer, E., & Plumpter, T. (2007). The gendered nature of natural disasters: The impact of catastrophic events on the gender gap in life expectancy, 1981–2002. *Annals of the Association of American Geographers, 97*(3), 551–566.

Neutra, R., Swan, S., & Mack, T. (1992). Clusters galore: Insights about environmental clusters from probability theory. *The Science of the Total Environment, 127*(1–2), 187–200.

Nicholson, M. C., & Mather, T. N. (1996). Methods for evaluating Lyme disease risks using geographic information systems and geospatial analysis. *Journal of Medical Entomology, 33*(5), 711–720.

Niggebrugge, A., Haynes, R., Jones, A., Lovett, A., & Harvey, I. (2005). The index of multiple deprivation 2000 access domain: A useful indicator for public health? *Social Science and Medicine, 60*(12), 2743–2753.

Nizeyimana, E., Petersen, G. W., Anderson, M. C., Evans, B. M., Hamlett, J. M., & Baumer, G. M. (1996). Statewide GIS/census data assessment of nitrogen loadings from septic systems in Pennsylvania. *Journal of Environmental Quality, 25*(2), 346–354.

Nolen, R. S. (2000). Multistate surveillance system in place for West Nile virus. *Journal of the American Veterinary Medical Association, 216*(1), 11.

Nonprofit Knowledge Works. (2011). Greater New Orleans Community Data Center. Retreived January 5, 2011, from *www.gnocdc.org*.

Norman, P., Boyle, P., & Rees, P. (2004). Selective migration, health and deprivation: A longitudinal analysis. *Social Science and Medicine, 60*(12), 2755–2771.

North American Association of Central Cancer Registries. (2011). About NAACCR. Retrieved January 5, 2011, from *www.naaccr.org*.

Nuojua, J. (2010). WebMapMedia: A map-based Web application for facilitating participation in spatial planning. *Multimedia Systems, 16*(1), 3–21.

Nuwer, M. R. (2008). The US health care system: Part 2: Proposal for improvement and comparison to other systems. *Neurology, 71*(23), 1914–1920.

Nyerges, T. (1993). Understanding the scope of GIS: Its relationship to environmental modeling. In M. F. Goodchild, B. O. Parks, & L. T. Steyaert (Eds), *Environmental modeling with GIS* (pp. 75–93). Oxford, UK: Oxford University Press.

Nyerges, T., Jankowski, P., Tuthill, D., & Ramsey, K. (2006). Collaborative water resource decision support: Results from a field experiment. *Annals of the Association of American Geographers, 96*(4), 699–725.

Oak Ridge National Laboratory. (2008). *LandScan(tm)*. Retrieved January 5, 2011, from *www.ornl.gov/sci/landscan*.

Obermeyer, N. (1995). Reducing inter-organizational conflict to facilitate sharing geographic information. In H. Onsrud & G. Rushton (Eds.), *Sharing geographic information* (pp. 138–148). New Brunswick, NJ: Center for Urban Policy Research.

O'Brien, M. (1994). A proposal to address, rather than rank, environmental problems. In

A. M. Finkel & D. Golding (Eds.), *Worst things first?: The debate over risk-based national environmental priorities* (pp. 87–106). Washington, DC: Resources for the Future.

O'Carroll, P. W., Yasnoff, W. A., Ward, M. E., Ripp, L. H., & Martin, E. L. (2003). *Public health informatics and information systems*. New York: Springer-Verlag.

Office of Environmental Health Hazard Assessment, State of California. (2008). California Environmental Contaminant Monitoring Program frequently asked questions. Retrieved January 5, 2011, from *oehha.ca.gov/multimedia/biomon/pdf/CECBP-FAQs032508.pdf.*

Office of Geographic and Environmental Information (MassGIS). (2011). *Massachusetts Geographic Information System*. Commonwealth of Massachusetts Executive Office of Environmental Affairs. Retrieved January 5, 2011, from *www.mass.gov/mgis/massgis.htm*.

Office of Policy, Planning and Evaluation. (2002). One Hundred Fifty-First Connecticut Registration Report 1998. Retrieved January 5, 2011, from *www.ct.gov/dph/lib/dph/hisr/xls/rr1998.xls.*

Ogden, N. H., St.-Onge, L., Barker, I. K., Brazeau, S., Bigras-Poulin, M., Charron, D. F., et al. (2008). Risk maps for range expansion of the Lyme disease vector, Ixodes scapularis, in Canada now and with climate change. *International Journal of Health Geographics, 7*, 24.

Olea, R. A. (1984). Sampling design optimization for spatial functions. *Mathematical Geology, 16*(4), 365–391.

Oliver, L., Ferber, U., Grimski, D., Millar, K., & Nathanail, P. (2005). The scale and nature of European brownfields. Retrieved January 5, 2011, from *www.cabernet.org.uk/resourcefs/417.pdf.*

Onega, T., Duell, E., Shi, X., Wang, D., Demidenko, E., & Goodman, D. (2008). Geographic access to cancer care in the U.S. *Cancer, 112*(4), 909–918.

Onsrud, H., & Rushton, G. (1995). *Sharing geographic information*. New Brunswick, NJ: Center for Urban Policy Research.

Onyebuchi, A. A., Ogbu, U. C., & Okeke, C. E. (2008). Too poor to leave, too rich to stay: Developmental and global health correlates of physician migration to the United States, Canada, Australia, and the United Kingdom. *American Journal of Public Health, 98*(1), 148–154.

Open Geospatial Consortium. (2011a). *Geography markup language*. Open Geospatial Consortium, Inc. Retrieved March 30, 2010, from *www.opengeospatial.org/standards/gml*.

Open Geospatial Consortium, Inc. (2011b). GML in JPEG 2000 for Geographic Imagery Encoding. Retrieved March 30, 2010, from *www.opengeospatial.org/standards/gmljp2*.

Openshaw, S. (1984). *The modifiable areal unit problem, Concepts and techniques in modern geography* (Number 38). Norwich, UK: Geo Books.

Openshaw, S., Charlton, M., & Craft, A. (1988). Searching for leukemia clusters using a geographical analysis machine. *Papers of the Regional Science Association, 64*, 95–106.

Oppong, J. R. (1996). Accommodating the rainy season in Third World location-allocation applications. *Socio-Economic Planning Sciences, 30*(2), 121–137.

Ordnance Survey. (1999). *Gazetteer of Great Britain: All names from the 1:50,000 scale Landranger maps* (4th ed.). London: Macmillan.

Ordnance Survey. (2011a). *1:50 000 scale gazetteer*. Ordnance Survey. Retrieved January

5, 2011, from *www.ordnancesurvey.co.uk/oswebsite/products/50kgazetteer_open/techinfo.html*.

Ordnance Survey. (2011b). *Using the National Grid*. Ordnance Survey. Retrieved January 5, 2011, from *www.ordnancesurvey.co.uk/oswebsite/gi/nationalgrid/nationalgrid.pdf*.

Ordnance Survey. (2011c). Get-a-map. Retrieved March 30, 2010, from *www.ordnancesurvey.co.uk/oswebsite/getamap/*.

Osleeb, J., & Kahn, S. (1999). Integration of geographic information. In V. Dale & M. English (Eds.), *Tools to aid environmental decision making* (pp. 161–189). New York: Springer-Verlag.

Ostfeld, R. S. (2009). Biodiversity loss and the rise of zoonotic pathogens. *Clinical Microbiology and Infection, 15*(1 Suppl. 1), 40–43.

O'Sullivan, D. (2008). Geographical information science: Agent-based models. *Progress in Human Geography, 32*(4), 541–550.

Overton, W. S., & Stehman, S. V. (1993). Properties of designs for sampling continuous spatial resources from a triangular grid. *Communications in Statistics: Theory and Methods, 22*(9), 251–264.

Owen, I. (1889). Geographical distribution of rickets, acute and subacute rheumatism, chorea, cancer, and urinary calculus in the British Islands. *British Medical Journal, 1*(1464), 113–116.

Pang, T. T. P., & Lee, S. S. (2008). Measuring the geographic coverage of methadone maintenance programme in Hong Kong by using geographic information system (GIS). *International Journal of Health Geographics, 7*, 5.

Pappas, G., Hadden, W., Kozak, L., & Fisher, G. (1997). Potentially avoidable hospitalization: Inequalities in rates between US socioeconomic groups. *American Journal of Public Health, 87*(5), 811–816.

Parker, E., & Campbell, J. (1998). Measuring access to primary medical care: Some examples of the use of geographical information systems. *Health and Place, 4*(2), 183–193.

Parrish, R. G., II, & McDonnell, S. M. (2000). Sources of health-related information. In S. M. Teutsch & R. E. Churchill (Eds.), *Principles and practice of public health surveillance*, 2nd ed. (pp. 30–75). New York: Oxford University Press.

Parrott, R., & Stutz, F. P. (1991). Urban GIS applications. In D. J. Maguire, M. F. Goodchild, & D. W. Rhind (Eds.), *Geographical information systems: Principles and applications: Vol. 2. Applications* (pp. 247–260). Harlow, UK: Longman Scientific & Technical.

Parry, R. B., & Perkins, C. R. (Eds.). (1987). *World mapping today*. London: Butterworths.

Partridge, E., & Fouad, M. (2010). Community-driven approaches to for reducing health disparities in cancer. *Journal of the American Medical Association, 303*(11), 1090–1091.

Patz, J. A., & Olson, S. H. (2008). Climate change and health: Global to local influences on disease risk. In S. M. Lemon, P. F. Sparling, M. A. Hamburg, D. A. Relman, E. R. Choffnes, & A. Mack (Eds.), *Vector-borne diseases: Understanding the environmental, human health, and ecological connections: Workshop summary* (pp. 88–103). Washington, DC: National Academies Press.

Patz, J. A., Strzepek, K., Lele, S., Hedden, M., Greene, S., Noden, B., Hay, S. I., Kalkstein, L., & Beier, J. C. (1998). Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya. *Tropical Medicine and International Health, 3*(10), 818–827.

Pearce, J., Hiscock, R., Blakely, T., & Witten, K. (2009). A national study of the association between neighborhood access to fast-food outlets and the diet and weight of local residents. *Health and Place, 15*(1), 193–197.

Pearce, J., Witten, K., Hiscock, R., & Blakely, T. (2007). Are socially disadvantaged neighborhoods deprived of community health-related resources? *International Journal of Epidemiology, 36*(2), 348–355.

Pearce, N. (1996). Traditional epidemiology, modern epidemiology, and public health. *American Journal of Public Health, 86*(5), 678–683.

Pearson, F., II. (1990). *Map projections: Theory and applications*. Boca Raton, FL: CRC Press.

Peek-Asa, C., Ramirez, M. R., Shoaf, K., Seligson, H., & Kraus, J. F. (2000). GIS mapping of earthquake-related deaths and hospital admissions from the 1994 Northridge, California, earthquake. *Annals of Epidemiology, 10*(1), 5–13.

Penchansky, R., & Thomas, J. W. (1981). The concept of access: Definition and relationships to consumer satisfaction. *Medical Care, 19*(2), 127–140.

Peng, Z-R., & Tsou, M-H. (2003). *Internet GIS: Distributed geographic information services for the Internet and wireless networks*. Hoboken, NJ: Wiley.

Pennington, D. W., Margni, M., Ammann, C., & Jolliet, O. (2005). Multimedia fate and human intake modeling: Spatial versus nonspatial insights for chemical emissions in Western Europe. *Environmental Science & Technology, 39*(4), 1119–1128.

Peters, A., & MacDonald, H. (2004). *Unlocking the Census with GIS*. Redlands, CA: ESRI.

Peters, J., & Hall, G. B. (1999). Assessment of ambulance response performance using a geographic information system. *Social Science and Medicine, 49*(11), 1551–1566.

Peterson, M. P. (2005). Maps and the Internet: An introduction. In M. P. Peterson (Ed.), *Maps and the Internet* (pp. 1–16). Amsterdam: Elsevier.

Peuquet, D. J. (1984). A conceptual framework and comparison of spatial data models. *Cartographica, 21*(4), 66–113.

Phair, M. (Ed.). (1997). Standards open GIS data access. *ENR, 239*(9), 15.

Phillips, R., Kinman, E., Schnitzer, P., Lindblom, E., & Ewigman, B. (2000). Using geographic information systems to understand health care access. *Archives of Family Medicine, 9*(10), 971–978.

Pickett, K. E., & Pearl, M. (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: A critical review. *Journal of Epidemiology and Community Health, 55*(2), 111–122.

Pickle, L. W. (2009). A history and critique of U.S. mortality atlases. *Spatial and Spatio-temporal Epidemiology, 1*(1), 3–17.

Pickles, J. (1995a). Representations in an electronic age: Geography, GIS, and democracy. In J. Pickles (Ed.), *Ground truth: The social implications of geographic information systems* (pp. 1–30). New York: Guilford Press.

Pickles, J. (1995b). *Ground truth: The social implications of geographic information systems.* New York: Guilford Press.

Pistocchi, A. (2008). A GIS-based approach for modeling the fate and transport of pollutants in Europe. *Environmental Science and Technology, 42*(10), 3640–3647.

Plewe, B. (1997). *GIS online: Information retrieval, mapping, and the Internet*. Santa Fe, NM: OnWord Press.

Plewe, B. (2007). Web cartography in the United States. *Cartography and Geographic Information Science, 34*(2), 133–136.

Poiker, T. K. (1985). Geographic information systems in the geographic curriculum. *Operational Geographer, 3*(8), 38–41.

Pollack, L. A., Gotway, C. A., Bates, J. H., Parikh-Patel, A., Richards, T. B., Seeff, L. C., et al. (2006). Use of the spatial scan statistic to identify geographic variations in late stage colorectal cancer in California (United States). *Cancer Causes and Control*, 17(4), 449–457.

Porta, M., Puigdomènech, E., Ballester, F., Selva, J., Ribas-Fitó, N., Llop, S., et al. (2008). Monitoring concentrations of persistent organic pollutants in the general population: The international experience. *Environment International, 34*(4), 546–561.

Poulstrup, A., & Hansen, H. L. (2004). Use of GIS and exposure modeling as tools in a study of cancer incidence in a population exposed to airborned dioxin. *Environmental Health Perspectives, 112*(9), 1032–1036.

Price, M. (2008). *Mastering ArcGIS*. New York: McGraw-Hill.

Probst, J., Laditka, J., & Laditka, S. (2009). Association between community health center and rural health clinic presence and county-level hospitalization rates for ambulatory care sensitive conditions: An analysis across eight US states. *BMC Health Services Research, 9*, 134.

Probst, J., Laditka, S., Wang, J., & Johnson, A. (2007). Effects of residence and race on burden of travel for care: Cross sectional analysis of the 2001 U.S. National Household Travel Survey. *BMC Health Services Research, 7*, 40.

Public Health Agency of Canada. (2007). *GIS for public health practice*. Catalogue # HP5–48/2007E-PDF. Retrieved January 5, 2011, from *www.phac-aspc.gc.ca/php-psp/gis_e.html*.

Puckett, L. J. (1994). *Nonpoint and point source of nitrogen in major watersheds of the United States. Water Resources Investigations Report 94-001*. Reston, VA: U.S. Geological Survey.

Puett, R. C., Lawson, A. B., Clark, A. B., Aldrich, T. E., Porter, D. W., Feigley, C. E., et al. (2005). Scale and shape issues in focused cluster power for count data. *International Journal of Health Geographics, 4*, 8.

Rainham, D., McDowell, I., Krewski, D., & Sawada, M. (2010). Conceptualizing the healthscape: Contributions of time geography, location technologies, and spatial ecology to place and health research. *Social Science and Medicine, 70*(5), 668–676.

Rand, P. W., Lacombe, E. H., Dearborn, R., Cahill, B., Elias, S., Lubelczyk, C. B., et al. (2007). Passive surveillance in Maine, an area emergent for tick-borne diseases. *Journal of Medical Entomology, 44*(6), 1118–1129.

Ratliff, Evan. (2007). The whole earth cataloged: How Google Maps is changing the way we see the world. *Wired, 15*(7), 155–159.

Reinhardt, M., Elias, J., Albert, J., Frosch, M., Harmsen, D., & Vogel, U. (2008). EpiScanGIS: An online geographic surveillance system for meningococcal disease. *International Journal of Health Geographics, 7*, 33.

Reisig, V., & Hobbiss, A. (2000). Food deserts and how to tackle them: A study of one city's approach. *Health Education Journal, 59*(2), 137–149.

ReVelle, Charles S., & Swain, R. W. (1970). Central facilities location. *Geographical Analysis, 2*(1), 30–42.

Rezza, G., Nicoletti, L., Angelini, R., Romi, R., Finarelli, A. C., Panning, M., et al. (2007).

Infection with chikungunya virus in Italy: An outbreak in a temperate region. *Lancet, 370*(9602), 1840–1846.

Rhodes, T., Singer, M., Bourgois, P., Friedman, S. R., & Strathdee, S. A. (2005). The social structural production of HIV risk among injecting drug users. *Social Science and Medicine, 61*(5), 1026–1044.

Richards, T. B., Croner, C. M., Rushton, G., Brown, C. K., & Fowler, L. (1999). Geographic information systems and public health: Mapping the future. *Public Health Reports, 114*(4), 359–373.

Richardson, M., van Lill, S. W. P., van der Spuy, G. D., Munch, Z., Booysen, C. N., Beyers, N., et al. (2002). Historic and recent events contribute to the disease dynamics of Beijing-like Mycobacterium tuberculosis isolates in a high incidence region. *International Journal of Tuberculosis and Lung Disease, 6*(11), 1001–1011.

Ricketts, T. C., Goldsmith, L. J., Holmes, G. M., Randolph, R., Lee, R., Taylor, D. H., et al. (2007). Designating places and populations as medically underserved: A proposal for a new approach. *Journal of Health Care for the Poor and Underserved, 18*(3), 567–589.

Rifai, H. S., Hendricks, L. A., Kilborn, K., & Bedient, P. B. (1993). A geographic information system (GIS) user interface for delineating wellhead protection areas. *Ground Water, 31*(3), 480–488.

Rigau-Pérez, J. G. (2008). Severe dengue: The need for new case definitions. *The Lancet Infectious Diseases*, 6(5), 297–302.

Rigaux, P., Scholl, M., & Voisard, A. (2002). *Spatial databases with application to GIS*. San Francisco, CA: Morgan Kaufmann.

Ritz, B., & Costello, S. (2006). Geographic model and biomarker-derived measures of pesticide exposure and Parkinson's disease. *Annals of the New York Academy of Sciences, 1076*, 378–387.

Riva, M., Curtis, S., Gauvin, L., & Fagg, J. (2009). Unravelling the extent of inequalities in health across urban and rural areas: Evidence from a national sample in England. *Social Science and Medicine, 68*(4), 654–663.

Riva, M., Gauvin, L., & Barnett, T. A. (2007). Toward the next generation of research into small area effects on health: A synthesis of multilevel investigations published since July 1998. *Journal of Epidemiology and Community Health, 61*(10), 853–861.

Rockland County Planning Department GIS. (2000). *Rockland County, New York, Number of dead wildlife per square mile*. Rockland County, NY: Author.

Rogerson, P. A. (2001). Monitoring point patterns for the development of space-time clusters. *Journal of the Royal Statistical Society* A, 164(1), 87–96.

Rogerson, P., & Yamada, I. (2009). *Statistical detection and surveillance of geographic clusters*. Boca Raton, FL: Chapman & Hall/CRC.

Rollin, B. E. (2003). Toxicology and new social ethics for animals. *Toxicologic Pathology, 31*(Suppl. 1), 128–131.

Rosenberg, M., & Hanlon, N. (1996). Access and utilization: A continuum of health service environments. *Social Science and Medicine, 43*(6), 975–983.

Ross, N., Wolfson, M., Berthelot, J-M., & Dunn, J. R. (2004). Why is mortality higher in unequal societies? Interpreting income inequality and mortality in Canada and the USA. In P. Boyle, S. Curtis, E. Graham, & E. Moore (Eds.), *The geography of health inequalities in the developed world: Views from Britain and North America* (pp. 103–128). Aldershot, UK: Ashgate.

Rossi, E., Rinaldi, L., Musella, V., Veneziano, V., Carbone, S., Gradoni, L., et al. (2007).

Mapping the main *Leishmania* phlebotomine vector in the endemic focus of the Mt. Vesuvius in southern Italy. *Geospatial Health, 1*(2), 191–198.

Rothenberg, R. (1983). The geography of gonorrhea transmission: empirical demonstration of core group transmission. *American Journal of Epidemiology, 117*(6), 688–694.

Rouhani, S. (1985). Variance reduction analysis. *Water Resources Research, 21*(6), 837–846.

Roush, S., Birkhead, G. Koo, D. Cobb, A., & Fleming, D. (1999). Mandatory reporting of disease and conditions by health care professionals and laboratories. *Journal of the American Medical Association, 282*(2), 164–170.

Rubes, J., Selevan, S. G., Evenson, D. P., Zudova, D., Vozdova, M., Zudova, Z., et al. (2005). Episodic air pollution is associated with increased DNA fragmentation in human sperm without other changes in semen quality. *Human Reproduction, 20*(10), 2776–2783.

Rubes, J., Selevan, S. G., Sram, R. J., Evenson, D. P., & Perreault, S. D. (2007). *GSTM1* genotype influences the susceptibility of men to sperm DNA damage associated with exposure to air pollution. *Mutation Research, 625*(1–2), 20–28.

Ruckleshaus, W. D. (1983). Science, risk, and public policy. *Science, 221*(4615), 1026–1028.

Ruiz, M. O. (2005). Geocoded photography. G*eospatial Solutions, 15*(5): 37–40.

Ruiz, M. O., Walker, E. D., Forster, E. S., Haramis, L. D., & Kitron, U. D. (2007). Association of West Nile virus illness and urban landscapes in Chicago and Detroit. *International Journal of Health Geographics, 6*, 10.

Rull, R. P., & Ritz, B. (2003). Historical pesticide exposure in California using pesticide use reports and land-use surveys: An assessment of misclassification error and bias. *Environmental Health Perspectives, 111*(13), 1582–1589.

Rupprecht, C. E., Smith, J. S., Fekadu, M., & Childs, J. E. (1995). The ascension of wildlife rabies: A cause for public health concern or intervention? *Emerging Infectious Diseases, 1*(4), 107–114.

Rushton, G. (1975). *Planning primary health services for rural Iowa: An interim report.* Technical Report No. 39. Iowa City, Iowa: University of Iowa, Center for Locational Analysis, Institute of Urban and Regional Research.

Rushton, G. (1997). *Improving public health through geographical information systems: An instructional guide to major concepts and their implementation*. Department of Geography, University of Iowa, Iowa City. Retrieved January 5, 2011, from *www.uiowa.edu/~geog/health*.

Rushton, G. (1999). Methods to evaluate geographic access to health services. *Journal of Public Health Management Practice, 5*(2), 93–100.

Rushton, G. (2003). Public health, GIS, and spatial analytic tools. *Annual Review of Public Health, 24*, 43–56.

Rushton, G., Armstrong, M. P., Gittler, J., Greene, B. R., Pavlik, C. E., West, M. M., et al. (Eds.). (2008). *Geocoding health data: The use of geographic codes in cancer prevention and control, research, and practice*. Boca Raton, FL: CRC Press.

Rushton, G., & Lolonis, P. (1996). Exploratory spatial analysis of birth defect rates in an urban population. *Statistics in Medicine, 15*(7–9), 717–726.

Ruth, M. (2010). GeoTIFF FAQ Version 2.4. Retrieved January 5, 2011, from *www.remotesensing.org/geotiff/faq.html*.

Ryan, P. H., & LeMasters, G. K. (2007). A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicology, 19*(Suppl. 1), 127–133.

Sabel, C. E., Boyle, P. J., Löytönen, M., Gatrell, A. C., Jokelainen, M., Flowerdew, R., et al. (2003). Spatial clustering of Amyotrophic Lateral Sclerosis in Finland at place of birth and place of death. *American Journal of Epidemiology, 157*(10), 898–905.

Sallis, J. F. (2009). Measuring physical activity environments: A brief history. *American Journal of Preventive Medicine, 36*(4S), S86–S92.

Sallis, J. F., Saelens, B. E., Frank, L. D., Conway, T. L., Slymen, D. J., Cain, K. L., et al. (2009). Neighborhood built environment and income: Examining multiple health outcomes. *Social Science and Medicine, 68*(7), 1285–1293.

Sandrick, K. (1998). Mapping the market: Managed care and geographic systems. *Health Management Technology, 19*(11), 17.

Sasaki, S., Suzuki, H., Fujino, Y., Kimura, Y., & Cheelo, M. (2009). Impact of drainage networks on outbreaks of cholera in Lusaka, Zambia. *American Journal of Public Health, 99*(11), 1982–1987.

SASI Group. (2008). The Grim Reaper's road map: An atlas of mortality in Britain. Retrieved January 5, 2011 from *www.sasi.group.shef.ac.uk/publications/reaper.html*.

Saul, S. (2006). U.S. doctors object to data-mining. *The New York Times*, May 4, 2006. Retrieved January 5, 2011, from *www.nytimes.com/2006/05/04/business/worldbusiness/04iht-Drugs.html*.

Saul, S. (2008). Federal court upholds drug privacy law. *The New York Times*, November 19, 2008, Section B, p. 10.

Savitz, D. A. (1997). The alternative to epidemiologic theory: Whatever works. *Epidemiology, 8*(2), 210–212.

Schapiro, M. (2007). *Exposed: The toxic chemistry of everday products and what's at stake for American power.* White River Junction, VT: Chelsea Green Publishing.

Schlossberg, M., & Shuford, E. (2005). Delineating "public" and "participation" in PPGIS. *URISA Journal, 16*(2), 15–26.

Schulz, C., Conrad, A., Becker, K., Kolossa-Gehring, M., Seiwert, M., & Seifert, B. (2007). Twenty years of the German Environmental Survey (GerES): Human biomonitoring—Temporal and spatial (West Germany/East Germany) differences in population exposure. *International Journal of Hygiene and Environmental Health, 210*(3–4), 271–297.

Schuurman, N. (2004). *GIS: A short introduction*. Malden, MA: Blackwell.

Schuurman, N., Bell, N. J., L'Heureux, R., & Hameed, S. M. (2009). Modelling optimal location for pre-hospital helicopter emergency medical services. *BMC Emergency Medicine, 9*, 6.

Schuurman, N., Fiedler, R., Grzybowski, S., & Grund, D. (2006). Defining rational hospital catchments for non-urban areas based on travel-time. *International Journal of Health Geographics, 5*, 43.

Schuurman, N., Leight, M., & Berube, M. (2008). A Web-based graphical user interface for evidence-based decision making for health care allocations in rural areas. *International Journal of Health Geographics, 7*, 49.

Schuurman, N., & Pratt, G. (2002). Care of the subject: Feminism and critiques of GIS, *Gender, Place and Culture, 9*(3), 291–299.

Scotch, M., Odofin, L., & Rabinowitz, P. (2009). Linkages between animal and human health sentinel data. *BMC Veterinary Research, 5*, 15.

Scott, A. J. (1970). Location-allocation systems: A review. *Geographical Analysis, 2*(2), 95–119.

Scott, D., & Horner, M. (2008). Examining the role of urban form in shaping people's accessibility to opportunities: An exploratory spatial data analysis. *Journal of Transport and Land Use, 1*(2), 89–119.

Scott, M., Cutter, S. L., Menzel, C., Minhe, J., & Wagner, D. (1997). Spatial accuracy of the EPA's environmental hazards databases and their use in environmental equity analyses. *Applied Geographic Studies, 1*(1), 45–61.

Scott, R. E., McCarthy, F. G., Jennett, P. A., Perverseff, T., Lorenzetti, D., Saeed, A., et al. (2007). Telehealth outcomes: A synthesis of the literature and recommendations for outcome indicators. *Journal of Telemedicine and Telecare, 13*(Suppl. 2), 1–38.

Scott, T. W., & Morrison, A. C. (2008). Longitudinal field studies will guide a paradigm shift in dengue prevention. In S. M. Lemon, P. F. Sparling, M. A. Hamburg, D. A. Relman, E. R. Choffnes, & A. Mack (Eds.), *Vector-borne diseases: Understanding the environmental, human health, and ecological connections: Workshop summary* (pp. 132–149). Washington, DC: National Academies Press.

Scribner, R., Mason, K., Theall, K., Simonsen, N., Schneider, S. K., Towvim, L. G., et al. (2008). The contextual role of alcohol outlet density in college drinking. *Journal of Studies on Alcohol and Drugs, 69*(1), 112–120.

Scribner, R. A., Johnson, S. A., Cohen, D. A., Robinson, W., Farley T. A., & Gruenewald, P. (2008). Geospatial methods for identification of core groups for HIV/AIDS. *Substance Use and Misuse, 403*(2), 203–221.

Seidel, J., Beck, C., Pocobelli, G., Lemaire, J., Bugar, J., Quan, H., et al. (2006). Location of residence associated with the likelihood of patient visit to the preoperative assessment clinic. *BMC Health Services Research, 6*, 13.

Selvin, S. (2004). *Statistical analysis of epidemiologic data* (3rd ed.). New York: Oxford University Press.

Sepai, O., Collier, C., Van Tongelen, B., & Casteleyn, L. (2008). Human biomonitoring data interpretation and ethics: Obstacles or surmountable challenges? *Environmental Health, 7*(Suppl. 1), S13.

Serebriakova, T. N. (2005). *Mapping radiation from external and internal sources in Belarus after the Chernobyl disaster: Implications for epidemiological analysis*. Unpublished Doctoral Dissertation. Department of Geography, University of Connecticut.

Service, R. F. (2003). World Trade Center: Chemical studies of 9/11 disaster tell complex tale of "bad stuff." *Science, 301*(5640), 1649.

Servigne, S., Lesage, N., & Libourel, T. (2006). Quality components, standards, and metadata. In R. Devillers & R. Jeansoulin (Eds.), *Fundamentals of spatial data quality* (pp. 179–210). London: ISTE.

Seto, E. Y. W., Knapp, F., Zhong, B., & Yang, C. (2007). The use of a vest equipped with a global positioning system to assess water-contact patterns associated with schistosomiasis. *Geospatial Health, 1*(2), 233–241.

Shannon, G. W. (1980). The utility of medical geography research. *Social Science and Medicine, 14D*(1), 1–2.

Shannon, G. W. (1981). Disease mapping and early theories of yellow fever. *The Professional Geographer, 33*(2), 221–227.

Shannon, K., Rusch, M., Shoveller, J., Alexson, D., Gibson, K., Tyndall, M.W., et al.

(2007). Mapping violence and policing as an environmental-structural barrier to health service and syringe availability among substance-using women in street-level sex work. *International Journal of Drug Policy, 19*(2), 140–147.

Sharp, R. R. (2003). Ethical issues in environmental health research—Ethics and environmental health. *Environmental Health Perspectives, 111*(14): 1786–1788.

Shaw, M., Thomas, B., Smith, G. D., & Dorling, D. (2008). *The Grim Reaper's road map: An atlas of mortality in Britain*. Bristol, UK: Policy Press.

Sheppard, E. (2005). Knowledge production through critical GIS: Genealogy and prospects. *Cartographica, 40*(4), 5–20.

Sheppard, E., Couclelis, H., Graham, S., Harrington, J. W., & Onsrud, H. (1999). Geographies of the information society. *International Journal of Geographical Information Science, 13*(8), 797–823.

Sherman, J., Spencer, J., Preisser, J., Gesler, W., & Arcury, T. (2005). A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics, 4*, 24.

Shi X., & Berke, E. (2009). Computing travel time when the exact address is unknown: A comparison of point and polygon ZIP code approximation methods. *International Journal of Health Geographics, 8*, 23.

Shortt, N. K., Moore, A., Coombes, M., & Wymer, C. (2005). Defining regions for locality health care planning: A multidimensional approach. *Social Science and Medicine, 60*(12), 2715–2727.

Shoultz, J. (2005). *Enterprise real-time geocoding for public health analysis and visualization*. 2005 ESRI Health GIS Conference Paper 1060. South Carolina Department of Health and Environmental Control. Retrieved January 5, 2011, from *www.scdhec.gov/co/phsis/biostatistics/index.asp?page=award*.

Sicotte, D., & Swanson, S. (2007). Whose risk in Philadelphia? Proximity to unequally hazardous industrial facilities. *Social Science Quarterly, 88*(2), 515–534.

Sieber, R. (2000). Conforming (to) the opposition: The social construction of geographical information systems in social movements. *International Journal of Geographical Information Science, 14*(8), 775–793.

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers, 96*(3), 491–507.

Siegrist, J. (2001). Social epidemiology—a promising field. *International Journal of Epidemiology, 30*(1), 50.

Silverman, R. A., Galea, S., Blaney, S., Freese, J., Prezant, D. J., Park, R., et al. (2007). The "vertical response time": Barriers to ambulance response in an urban area. *Academic Emergency Medicine, 14*(9), 772–778.

Sims, L. D. (2007). Lessons learned from Asian H5N1 outbreak control. *Avian Diseases, 51*(s1), 174–181.

Singer, M., Stopka, T., Siano, C., Springer, K., Barton, G., Khoshnood, K., et al. (2000). The social geography of AIDS and hepatitis risk: Qualitative approaches for assessing local differences in sterile-syringe access among injection drug users. *American Journal of Public Health, 90*(7), 1049–1056.

Sloan, C. D., Duell, E. J., Shi, X., Irwin, R., Andrew, A. S., Williams, S. M., et al. (2009). Ecogeographic genetic epidemiology. *Genetic Epidemiology, 33*(4), 281–289.

Slocum, T. A., McMaster, R. B., Kessler, F. C, & Howard, H. H. (2009). *Thematic cartography and visualization* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Slocum, T., Yoder, S. C. Kessler, F. C., & Sluter, R. S. (2000). MapTime: Software for exploring spatiotemporal data associated with point locations. *Cartographica, 37*(1), 15–32.

Smith, D. L., Dushoff, J., Snow, R. W., & Hay, S. I. (2005). The entomological inoculation rate and Plasmodium falciparum infection in African children. *Nature, 438*(7067), 492–495.

Smith, D. M., Cummins, S., Taylor, M., Dawson, J., Marshall, D., Sparks, L., et al. (2010). Neighbourhood food environment and area deprivation: Spatial accessibility to grocery stores selling fresh fruit and vegetables in urban and rural settings. *International Journal of Epidemiology, 39*(1), 277–284.

Smith, D. P., Gould, M. I., & Higgs, G. (2003). (Re)surveying the uses of geographical information systems in health authorities (1991–2001). *Area*, 36(1), 81–83.

Smith, N. (1992). Real wars, theory wars. *Progress in Human Geography, 16*(2), 257–271.

Smolders, R., Casteleyn, L., Joas, R., & Shoeters, G. (2008). Human biomonitoring and the INSPIRE directive: Spatial data as link for environmental and health research. *Journal of Toxiocology and Environmental Health-Part B-Critical Reviews*, 11(8), 646–659.

Smolders, R., & Schoeters, G. (2007). Identifying opportunities and gaps for establishing an integrated EDR-triad at a European level. *International Journal of Hygiene and Environmental Health, 210*(3–4), 253–257.

Snyder, J. P. (1987). *Map projections: A working manual* (U.S. Geological Survey Professional Paper 1395). Washington, DC: U.S. Government Printing Office.

Sol, V. M., Lammers, P. E. M., Aiking, H., De Boer, J., & Feenstra, J. F. (1995). Integrated environmental index for application in land-use zoning. *Environmental Management, 19*(3), 456–467.

Somers, C. M., & Cooper, D. N. (2009). Air pollution and mutations in the germline: Are humans at risk? *Human Genetics, 125*(2), 119–130.

Soumare, B., Tempia, S., Cagnolati, V., Mohamoud, A., Van Huylenbroeck, G., & Berkvens, D. (2007). Screening for Rift Valley fever infection in northern Somalia: A GIS based survey method to overcome the lack of sampling frame. *Veterinary Microbiology, 121*(3–4), 249–256.

South Carolina Vital Record and Statistics Integrated Information Systems Project Team. (2005). *Understanding the South Carolina Vital Record and Statistics Integrated Information Systems*. South Carolina Department of Health and Environmental Control. Retrieved January 5, 2011, from *www.scdhec.gov/co/phsis/biostatistics/index.asp?page=vrsiis*.

Sperling, J. (1995). Development of the TIGER database: Experiences in spatial data sharing at the U.S. Bureau of Census. In H. Onsrud & G. Rushton (Eds.), *Sharing geographic information* (pp. 377–396). New Brunswick, NJ: Center for Urban Policy Research.

Spielman, A. (1994). Why entomological antimalaria research should not focus on transgenic mosquitoes. *Parasitology Today, 10*(10), 374–376.

Spielman, S. E., & Yoo, E-H. (2009). The spatial dimensions of neighborhood effects. *Social Science and Medicine, 68*(6), 1098–1105.

St. Martin, K., & Wing, J. (2007). The discourse and discipline of GIS. *Cartographica, 42*(3), 235–248.

Stafford, K. C., III, Denicola, A. J., Pound, J. M., Miller, J. A., & George, J. E. (2009). Topical treatment of white-tailed deer with an acaricide for the control of Ixodes scapularis (Acari: Ixodidae) in a Connecticut Lyme borreliosis hyperendemic community. *Vector-borne and Zoonotic Diseases, 9*(4), 371–379.

Stafford, M., Cummins, S., Macintyre, S., Ellaway, A., & Marmot, M. (2005). Gender differences in the associations between health and neighbourhood environment. *Social Science and Medicine, 60*(8), 1681–1692.

Stahler, G. J., Mazzella, S., Mennis, J., Chakravorty, S., Rengert, G., & Spiga, R. (2007). The effect of individual, program, and neighborhood variables on continuity of treatment among dually diagnosed individuals. *Drug and Alcohol Dependence, 87*(1), 54–62.

Stallones, L., Nuckols, J. R., & Berry, J. K. (1992). Surveillance around hazardous waste sites: geographic information systems and reproductive outcomes. *Environmental Research, 59*(2), 81–92.

Stark, S. L, Nuckols, J. R., & Rada, J. (1999). Using GIS to investigate septic system sites and nitrate pollution potential. *Journal of Environmental Health, 61*(8), 15–20, 64.

Steere, A. C., Broderick, T. F., & Malawista, S. E. (1978). Erythema chronicum migrans and Lyme arthritis: Epidemiological evidence for a tick vector. *American Journal of Epidemiology, 108*(4), 312–321.

Stern, J. E. (1989). *State plane coordinate system of 1983* (NOAA Manual NOS NGS 5). Rockville, MD: National Geodetic Information Center, NOAA.

Stirling, T. D. (1978). Does smoking kill workers or working kill smokers? OR The mutual relationship between smoking, occupation, and respiratory disease. *International Journal of Health Services, 8*(3), 437–452.

Stockwell, J. R., Sorensen, J. W., Eckert, J. W., Jr., & Carreras, E. M. (1993). The U.S. EPA geographic information system for mapping environmental releases of toxic chemical release inventory (TRI) chemicals. *Risk Analysis, 13*(2), 155–164.

Subramanian, S. V., & Kawachi, I. (2003). Response: In defense of the income inequality hypothesis. *International Journal of Epidemiology, 32*(6), 1037–1040.

Subramanian, S. V., & Kawachi, I. (2004). Income inequality and health: What have we learned so far? *Epidemiologic Reviews, 26*(1), 78–91.

Sui, D. Z. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment, and Urban Systems, 32*(1), 1–5.

Sumilo, D., Asokliene, L., Bormane, A., Vasilenko, V., Golovljova, I., & Randolph, S. E. (2007). Climate change cannot explain the upsurge of tick-borne encephalitis in the Baltics. *PLOS One, 2*(6), e500.

Sundby, J. (1999). Are women disfavoured in the estimation of Disability Adjusted Life Years and the Global Burden of Disease? *Scandinavian Journal of Public Health, 27*(4), 279–285.

Susser, M., & Susser, E. (1996a). Choosing a future for epidemiology: I. eras and paradigms. *American Journal of Public Health, 86*(5), 668–673.

Susser, M., & Susser, E. (1996b). Choosing a future for epidemiology: II. from black box to Chinese boxes and eco-epidemiology. *American Journal of Public Health, 86*(5), 674–677.

Tai, W., Porell, F., & Adams, E. K. (2004). Hospital choice of rural Medicare beneficiaries: Patient, hospital attributes, and the patient-hospital relationship. *Health Services Research, 39*(6), 1903–1922.

Talbot, T. O., Kulldorff, M., Forand, S. P., & Haley, V. B. (2000). Evaluations of spatial filters to create smoothed maps of health data. *Statistics in Medicine, 19*(17–18), 2399–2408.

Talen, E. (1998). Visualizing fairness: Equity maps for planners. *Journal of the American Planning Association, 64*(1), 22–38.

Talen, E. (2000). Bottom-up GIS: A new tool for individual and group expression in participatory planning. *Journal of the American Planning Association, 66*(3), 279–294.

Tango, T., & Takahashi, K. (2005). A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics, 4*, 11.

Tanser, F. (2006). Methodology for optimizing location of new primary health care facilities in rural communities: A cast study of KwaZulu-Natal, South Africa. *Journal of Epidemiology and Community Health, 60*(10), 846–850.

Tanser, F., Gijsbertsen, B., & Herbst, K. (2006). Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system. *Social Science and Medicine, 63*(3), 691–705.

Tarlov, E., Zenk, S., Campbell, R., Warnecke, R., & Block, R. (2009). Characteristics of mammography facility locations and stage of breast cancer at diagnosis in Chicago. *Journal of Urban Health, 86*(2), 196–213.

Taubman, D. S., & Marcellin, M. W. (Eds.). (2002). *JPEG 2000: Image compression fundamentals, standards, and practice*. Norwell, MA: Kluwer Academic Publishers.

Terry, T. (2004). The United States National Grid. *Professional Surveyor Magazine, 24*(10), 12–13, 16, 18, 20.

Thacker, S. B., Redmond, S., Rothernberg, R. B., Spitz, S. B., Choi, K., & White, M. (1986). A controlled trial of disease surveillance strategies. *American Journal of Preventive Medicine, 2*(6), 345–350.

Thacker, S. B., Stroup, D. F., Parrish, R. G., & Anderson, H. A. (1996). Surveillance in environmental public health: Issues, systems, and sources. *American Journal of Public Health, 86*(5), 633–641.

Theophilides, C. N., Ahearn, S. C., Binkowski, E. S., Paul, W. S., & Gibbs, K. (2006). First evidence of West Nile Virus amplification and relationship to human infections. *International Journal of Geographical Information Science, 20*(1), 103–115.

Thill, J., & Kim, M. (2005). Trip-making, induced travel demand and accessibility. *Journal of Geographical Systems, 7*(2), 229–248.

Thomas, D. C., Bowman, J. D., Jiang, L., Jiang, F., & Peters, J. M. (1999). *Bioelectromagnetics, 20*(7), 414–422.

Thomas, J., & Tucker, M. (1996). The development and use of the concept of a sexually transmitted disease core. *The Journal of Infectious Diseases, 174*(Suppl. 2), S134–S143.

Thomas, R. (1992). *Geomedical systems: Intervention and control*. New York: Routledge.

Thurston, J., Poiker, T. K., & Moore, J. P. (2003). *Integrated geospatial technologies: A guide to GPS, GIS, and data logging*. Hoboken, NJ: Wiley.

Thurston, R. C., Kubzansky, L. D., Kawachi, I., & Berkman, L. F. (2005). Is the association between socioeconomic position and coronary heart disease stronger in women than in men? *American Journal of Epidemiology, 162*(1), 57–65.

Tim, U. S. (1995). The application of GIS in environmental health sciences: opportunities and limitations. *Environmental Research, 71*(2), 75–88.

Timander, L., & McLafferty, S. (1998). Breast cancer in West Islip, NY: A spatial clustering analysis with covariates. *Social Science and Medicine, 46*(12), 1623–1635.

Timyan, J., Griffey Brechin, S., Measham, D., & Ogunleye, B. (1993). Access to care: More than a problem of distance. In M. Koblinsky, J. Timyan, & J. Gay (Eds.), *The health of women: A global perspective* (pp. 217–234). Boulder, CO: Westview Press.

Tobler, W. R. (1969). Geographical filters and their inverses. *Geographical Analysis, 1*(3), 234–253.

Tobler, W. R. (1973). Choropleth maps without classes intervals? *Geographical Analysis, 5*(3), 262–265.

Tomlin, C. D. (1990). *Geographic information systems and cartographic modeling.* Englewood Cliffs, NJ: Prentice Hall.

Tomlinson, R. F. (1987). Current and potential uses of geographical information systems: The North American experience. *International Journal of Geographical Information Systems, 1*(3), 203–218.

Tomlinson, R. F., & Boyle, A. R. (1981). The state of development of systems for handling national resources inventory data. *Cartographica, 18*(4), 65–95.

Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations Research, 19*(5), 1363–1373.

Torok, T. J., Kilgore, P. E., Clarke, M. J., Holman, R. C., Bresee, J. S., & Glass, R. I. (1997). Visualizing geographic and temporal trends in rotavirus activity in the United States, 1991–1996. *Pediatric Infectious Disease Journal, 16*(10), 941–946.

Townsend, P., Phillimore, P., & Beattie, A. (1988). *Health and deprivation: Inequality and the north*. Beckenham, UK: Croom Helm.

Toyabe, S., & Kouhei, A. (2006). Referral from secondary care and to aftercare in a tertiary care university hospital in Japan. *BMC Health Services Research, 6*, 11.

Treno, A. J., Alaniz, M. L., & Gruenewald, P. J. (2000). The use of drinking places by gender, age, and ethnic groups: An analysis of routine drinking activities. *Addiction, 95*(4), 537–551.

Trooskin, S. B., Hadler, J., St. Louis, T., & Navarro, V. J. (2005). Geospatial analysis of hepatitis C in Connecticut: A novel application of a public health tool. *Public Health, 119*(11), 1042–1047.

Troped, P. J., Wilson J., Matthews, C. E., Cromley, E. K., & Melly, S. J. (2010). Associations between the built environment and location-based physical activity. *American Journal of Preventive Medicine, 38*(4), 429–438.

Tsai, T. F., Popovici, F., Cernescu, C., Campbell, G. L., & Nedelcu, N. I. (1998). West Nile encephalitis epidemic in southeastern Romania. *Lancet, 352*(9130), 767–771.

Tufte, E. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

U.K. Department of Health. (2011). Health Survey for England. Retrieved January 5, 2011, from *www.dh.gov.uk/en/Publicationsandstatistics/PublishedSurvey/HealthSurveyForEngland/index.htm*.

Urbat, M., Lehndorff, E., & Schwark, L. (2004). Biomonitoring of air quality in the Cologne conurbation using pine needles as a passive sampler: Part I. Magnetic properties. *Atmospheric Environment, 38*(23), 3781–3792.

U.S. Census Bureau. (1992). *TIGER/Line (TM) Files, 1992*. Chapter 5: Data Quality. Retrieved January 5, 2011, from *www.census.gov/geo/www/tiger/chapter5.asc*.

U.S. Census Bureau. (2005). Spatial data storage and topology in the redesigned MAF/TIGER system. Retrieved January 5, 2011, from *www.census.gov/geo/mtep_obj2/topo_and_data_stor.html*.

U.S. Census Bureau. (2008a). Appendix A. Geographic terms and concepts. Retrieved January 5, 2011, from *www.census.gov/geo/www/geoareas/GTC_08.pdf*.

U.S. Census Bureau. (2008b). Census dates for countries and areas of the world: 1945 to 2014. Retrieved January 5, 2011, from *www.census.gov/ipc/www/cendates*.

U.S. Census Bureau. (2009a). Technical Documentation: 2009 Tiger/Line Shapefiles. Retrieved January 5, 2011, from *www.census.gov/geo/www/tiger/tgrshp2009/TGRSHP09.pdf*.

U.S. Census Bureau. (2010a). International statistical agencies. Retrieved January 5, 2011, from *www.census.gov/aboutus/stat_int.html*.

U.S. Census Bureau. (2010b). American National Standards Institute (ANSI) Codes. Retrieved January 5, 2011, from *www.census.gov/geo/www/ansi/ansi.html*.

U.S. Department of Agriculture. (2008). Qualitative and quantitative synopsis on NAIP usage from 2004–2008. Retrieved January 5, 2011, from *www.fsa.usda.gov/Internet/FSA_File/naip_usage.pdf*.

U.S. Department of Agriculture. (2010). Imagery Programs: NAIP Imagery. Retrieved January 5, 2011, from *www.fsa.usda.gov/FSA/apfoapp?area=home&subject=prog&topic=nai*.

U.S. Department of Health and Human Services. (2000). *Healthy People 2010*, 2nd ed., Vol. II. Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (2003). *Summary of the HIPAA Privacy Rule*. Retrieved January 5, 2011, from *www.hhs.gov/ocr/privacy/hipaa/understanding/summary/privacysummary.pdf*.

U.S. Department of Health and Human Services. (2006). *Healthy People 2010 Midcourse Review*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Transportation, Pipeline and Hazardous Materials Safety Administration. (2011). Incident reporting. Retrieved January 5, 2011, from *www.phmsa.dot.gov/hazmat/incident-report*.

U.S. Environmental Protection Agency. (1992). Environmental equity: Reducing risks for all communities. Retreived January 5, 2011, from *www.epa.gov/compliance/ej/resources/reports/annual-project-reports/reducing_risk_com_vol1.pdf*.

U.S. Environmental Protection Agency. (2006). Toxics Release Inventory Burden Reduction Final Rule. *Federal Register, 71*(246), 76932–76045.

U.S. Environmental Protection Agency. (2007). Assessing health risks from pesticides. Retrieved January 5, 2011, from *www.epa.gov/pesticides/factsheets/riskassess.htm*.

U.S. Environmental Protection Agency. (2008a). 2006 TRI public data release. Retrieved January 5, 2011, from *www.epa.gov/TRI/tridata/tri06*.

U.S. Environmental Protection Agency. (2008b). Unused pharmaceuticals in the health care industry: Interim report. EPA-821-R-08-010. Retrieved January 5, 2011, from *www.epa.gov/waterscience/ppcp/hcioutreach.pdf*.

U.S. Environmental Protection Agency. (2009a). Using North American Industrial Classification System (NAICS) Codes for Toxics Release Inventory (TRI) reporting. Retrieved January 5, 2011, from *www.epa.gov/tri/lawsandregs/naic*.

U.S. Environmental Protection Agency. (2009b). Facility Registry System (FRS). Retrieved January 5, 2011, from *www.epa.gov/enviro/html/facility.html*.

U.S. Environmental Protection Agency. (2010a). Toxics Release Inventory program: Laws, regulations and executive orders. Retrieved January 5, 2010, from *www.epa.gov/tri/lawsandregs/index.htm*.

U.S. Environmental Protection Agency. (2010b). Risk-Screening Environmental Indicators (RSEI). Retrieved January 5, 2011, from *www.epa.gov/opptintr/rsei*.

U.S. Environmental Protection Agency. (2010c). National Air Toxics Assessments. Retrieved January 5, 2011, from *www.epa.gov/ttn/atw/natamain*.

U.S. Geological Survey. (2009a). A new generation of maps. Retrieved January 5, 2011, from *nationalmap.gov/digital_map/index.html*.

U.S. Geological Survey. (2010a). EarthExplorer. Retrieved January 5, 2011, from *edcsns17.cr.usgs.gov/EarthExplorer*.

U.S. Geological Survey. (2010b). Digital line graphs (DLGs). Retrieved January 5, 2011, from *eros.usgs.gov/#/Find_Data/Products_and_Data_Available/DLGs*.

U.S. Geological Survey. (2010c). Disease maps 2010. Retrieved January 5, 2011, from *diseasemaps.usgs.gov/index.html*.

U.S. Geological Survey. (2011). *Geographic Names Information System (GNIS)*. U.S. Geological Survey. Retrieved January 5, 2011, from geonames.usgs.gov/pls/gnispublic.

Vallin, J., & Meslé, F. (2004). Convergences and divergences in mortality. A new approach to health transition. *Demographic Resesearch, 2*(2), 11–43.

Van Eldere J., Mera, R. M., Miller, L. A., Poupard, J.A., & Amrine-Madsen, H. (2007). Risk factors for development of multiple-class resistance to Streptococcus pneumoniae strains in Belgium over a 10–year period: Antimicrobial consumption, population density, and geographic location. *Antimicrobial Agents and Chemotherapy, 51*(10), 3491–3497.

Van Zandt, S., & Mhatre, P. C. (2009). Growing pains: Perpetuating inequality through the production of low-income housing in the Dallas/Fort Worth metroplex. *Urban Geography, 30*(5), 490–513.

Veregin, H., & Hargitai, P. (1995). An evaluation matrix for geographical data quality. In S. C. Guptill & J. L. Morrison (Eds.), *Elements of spatial data quality* (pp. 167–188). Oxford, UK: Elsevier Science.

Viel, J-F., Arveux, P., Baverel, J., & Cahn, J-Y. (2000). Soft-tissue sarcoma and non-Hodgkin's lymphoma clusters around a municipal solid waste incinerator with high dioxin emission levels. *American Journal of Epidemiology, 152*(1), 13–19.

Vimo. (2011). About Vimo at Vimo-Comparison shopping for health. Retrieved January 5, 2011, from *www.vimo.com/html/about.php*.

Vogt, R. L. (1992). National survey of state epidemiologists to determine the status of Lyme disease surveillance. *Public Health Reports, 107*(6), 644–646.

Vogt, R. L., LaRue, D., Klaucke, D. N., & Jillson, D. A. (1983). Comparison of an active and passive surveillance system of primary care providers for hepatitis, measles, rubella and salmonellosis in Vermont. *American Journal of Public Health, 73(7)*, 795–797.

Voiland, A. (2008). New satellite data reveal impact of Olympic pollution controls. NASA News, December 16, 2008. Retrieved January 5, 2011, from *www.nasa.gov/topics/earth/features/olympic_pollution.html*.

Wagenet, R. J., & Hutson, J. L. (1996). Scale-dependency of solute transport modeling/GIS applications. *Journal of Environmental Quality, 25*(3), 499–510.

Walker, D. H., Barbour, A. G., Oliver, J. H., Lane, R. S., Dumler, J. S., Dennis, D., T., et al. (1996). Emerging bacterial zoonotic and vector-borne diseases: Ecological and epidemiological factors. *Journal of the American Medical Association, 275*(6), 463–469.

Walker, R. E., Keane, C. R., & Burke, J. G. (2010). Disparities and access to healthy food in the United States: A review of the food deserts literature. *Health and Place, 16*(5), 876–884.

Wallace, D., & Wallace, R. (1998). *A plague on your houses: The war on the urban poor.* London: Verso.

Waller, L. A., & Gotway, C. A. (2004). *Applied spatial statistics for public health data.* Hoboken, NJ: Wiley.

Walsh, M. P. (2008). Ancillary benefits for climate change mitigation and air pollution control in the world's motor vehicle fleets. *Annual Review of Public Health, 29,* 1–9.

Walsh, S. J., Page, P. H., & Gesler, W. M. (1997). Normative models and healthcare palnning: Network-based simulations within a geographic information system environment. *Health Services Research, 32*(2), 243–260.

Wang, F. (2006). *Quantitative methods and applications in GIS.* Boca Raton, FL: CRC Press, Taylor & Francis Group.

Wang, F., & Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health and Place, 11*(2), 131–146.

Wang, L. (2007). Immigration, ethnicity and accessibility to culturally diverse family physicians. *Health and Place, 13*(3), 656–671.

Wang, L., Rosenberg, M., & Lo, L. (2008). Ethnicity and utilization of family physicians: a case study of Mainland Chinese immigrants in Toronto, Canada. *Social Science and Medicine, 67*(9), 1410–1422.

Wang, M. C., Gonzalez, A. A., Ritchie, L. D., & Winkleby, M. A. (2006). The neighborhood food environment: Sources of historical data on retail food stores. *International Journal of Behavioral Nutrition and Physical Activity, 3,* 15.

Wardak, A., Gorman, M. E., Swami, N., & Rejeski, D. (2007). Environmental regulation of nanotechnology and the TSCA. *IEEE Technology and Society Magazine, 27*(2), 48–56.

Waring, S., Zakos-Feliberti, A., Wood, R., Stone, M., Padgett, P, & Arafat, R. (2005). The utility of geographic information systems (GIS) in rapid epidemiological assessments following weather-related disasters: Methodological issues based on the Tropical Storm Allison experience. *International Journal of Hygiene and Environmental Health, 208*(1–2), 109–116.

Warnecke, L. Johnson, J. M., Marshall, K., & Brown, R. S. (1992). *State geographical information activities compendium.* Lexington, KY: Council of State Government.

Warrell, M. J., & Warrell, D. A. (2004). Rabies and other lyssavirus diseases. *The Lancet, 363*(9413), 959–969.

Wartenberg, D., & Greenberg, M. (1990). Detecting disease clusters: the importance of statistical power. *American Journal of Epidemiology, 132*(Suppl. 1), S156–S166.

Waxman, M. (2009). Infectious disease/CDC update: Commentary. *Annals of Emergency Medicine, 53*(5), 645–646.

Weis, B. K., Balshaw, D., Barr, J. R., Brown, D., Ellisman, M., Lioy, P., et al. (2005). Personalized exposure assessment: Promising approaches for human environmental health research. *Environmental Health Perspectives, 113*(7), 840–848.

Weise, F. (Ed). (1997). *Health statistics: An annotated bibliographic guide to information resources.* Lanham, MD: Medical Library Association and the Scarecrow Press.

Weisner, C. G. (1994). *Pediatric asthma and outdoor air pollution in Brooklyn and Queens: A geographical analysis.* Unpublished Master's thesis, Department of Geography, Hunter College, City University of New York.

Wennberg, J. (Ed.). (1998). *The Dartmouth atlas of health care*. Chicago: American Hospital Publishing.

Wennberg, J., Fisher, E., & Skinner, J. (2004). Geography and the debate over Medicare reform. *Health Affairs, 23*(2004 Variations Supplement), W96–W114.

Wennberg, J., & Gittelsohn, A. (1982). Variations in medical care among small areas. *Scientific American, 246*(4), 120–134.

Westervelt, J. D. (2002). Geographic information systems and agent-based modelling. In H. R. Gimblett (Ed.), *Integrating geographic information systems and agent-based modelling techniques for simulating social and ecological processes*. Oxford, UK: Oxford University Press.

Wharton, M., Chorba, T. L., Vogt, R. L., Morse, D. L., & Buehler, J. W. (1990). Case definitions for public health surveillance. *MMWR, 39*(RR-13), 1–43.

Wheeler, D. W. (2007). A comparison of spatial clustering and cluster detection techniques for childhood leukemia incidence in Ohio, 1996–2003. *International Journal of Health Geographics, 6*, 13.

Whitsel, E. A. (2008). Error and bias in geocoding school and students' home addresses. *Environmental Health Perspectives, 116*(8), A330–A330.

Wieland, S. C., Cassa, C. A., Mandl, K. D., & Berger, B. (2008). Revealing the spatial distribution of a disease while preserving privacy. *Proceedings of the National Academy of Sciences of the United States of America, 105*(46), 17608–17613.

Williams, D. R., & Collins, C. (2001). Racial residential segregation: A fundamental cause of racial disparities in health. *Public Health Reports, 116*(5), 404–416.

Wilson, J. D. (1998). GIS goes mobile. *GeoWorld, 11*(12), 54–57.

Wilson, J. L., & Branigan, A. (1999). Location of East Carolina University School of Medicine telemedicine sites in relation to primary care health professional shortage areas for North Carolina's Health Service Area VI, 1997. *Journal of Public Health Management and Practice, 5*(2), 45–46.

Wilson, M. E. (2003). The traveler and emerging infections: Sentinel, courier, transmitter. *Journal of Applied Microbiology, 94*(Suppl.), 1S-11S.

Wilson, M. L., Bretsky, P. M., Cooper, G. H.., Egbertson, S. H., Van Kruiningen, H. J., & Cartter, M. L. (1997). Emergence of raccoon rabies in Connecticut, 1991–1994: Spatial and temporal characteristics of animal infection and human contact. *American Journal of Tropical Medicine and Hygiene, 57*(4), 457–463.

Winch, P. (1998). Social and cultural responses to emerging vector-borne diseases. *Journal of Vector Ecology, 23*(1), 47–53.

Wing, S., Freedman, S., & Band, L. (2002). The potential impact of flooding on confined animal feeding operations in Eastern North Carolina. *Environmental Health Perspectives, 110*(4), 387–391.

Winn, D. W. (2005). Science and society: The Long Island Breast Cancer Study Project, *Nature Reviews Cancer, 5*(12), 986–994.

Wolch, J. (1996). Community-based human service delivery. *Housing Policy Debate, 7*(4), 649–670.

Wong, D. (2009). The modifiable area unit problem. In A. S. Fotheringham & P. Rogerson (Eds.), *The SAGE Handbook of Spatial Analysis* (pp. 105–123). Los Angeles: Sage.

Wong, D. W. S., & Lee, J. (2005). *Statistical analysis of geographic information with ArcView GIS and ArcGIS*. Hoboken, NJ: Wiley.

Wong, D. W. S., & Meyer, J. W. (1993). A spatial decision support system approach to evaluate the efficiency of a meals-on-wheels program. *The Professional Geographer, 45*(3), 332–341.

Wood, M. (1994). The traditional map as a visualization technique. In H. M. Hearnshaw & D. J. Unwin (Eds.), *Visualization in geographical information systems* (pp. 9–17). Chichester, UK: Wiley.

Woodward, M. (2004). *Epidemiology: Study design and data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Worboys, M. (1995). *GIS: A computing perspective*. Bristol, PA: Taylor & Francis.

Worboys, M. F. (1999). Relational databases and beyond. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems: Vol. 1. Principles and technical issues* (2nd ed., pp. 373–384). New York: Wiley.

Worboys, M., & Duckham, M. (2004). *GIS: A computing perspective* (2nd ed.). Boca Raton, FL: CRC Press.

World Health Organization. (2006). *World health report 2006*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2008). The global burden of disease: 2004 update. Retrieved January 5, 2011, from *www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf*.

World Health Organization. (2011). Global Health Atlas. Retrieved January 5, 2011, from *apps.who.int/globalatlas*.

Xiao, N., Armstrong, M., & Bennett, D. (2005). ChoroWare: A software toolkit for choropleth map classification. *Geographical Analysis, 38*(1), 102–122.

Yang, Y., & Atkinson, P. M. (2008). Individual space–time activity-based model: A model for the simulation of airborne infectious-disease transmission by activity-bundle simulation. *Environment and Planning B, 35*(1), 80–99.

Yasenovskiy, V., & Hodgson, J. (2007). Hierarhical location-allocation with spatial choice interaction modeling. *Annals of the Association of American Geographers, 97*(3), 496–511.

Yearsley, C., Worboys, M., Story, P., Jayawardena, P., & Bofakos, P. (1994). Computational support for spatial information systems: Models and algorithms. In M. Worboys (Ed.), *Innovations in GIS 1* (pp. 75–88). Bristol, PA: Taylor & Francis.

Yeh, A. G. (1999). Urban planning and GIS. In P. A. Longley, M. F. Goodchild, & D. J. Maguire (Eds.), *Geographical information systems: Vol. 2. Management issues and applications* (2nd ed., pp. 877–888), New York: Wiley.

Young, R. (1999). Prioritising family health needs: A time–space analysis of women's health-related behaviours. *Social Science and Medicine, 48*(6), 797–813.

Young, S., Balluz, L., & Malilay, J. (2004). Natural and technologic hazardous material releases during and after natural disasters: A review. *Science of the Total Environment, 322*(1–3), 3–20.

Yu, D. (2008). Taking an exposure history. Agency for Toxic Substances & Disease Registry, Case Studies in Environmental Medicine. Retrieved January 5, 2011, from *www.atsdr.cdc.gov/csem/exphistory/ehcover_page.html*.

Yuan, M. (2009). Space-time modeling 2: GIScience approaches. In R. Kitchen & N. Thrifts (Eds.), *International encyclopedia of human geography* (Vol. 10, pp. 286–295). Oxford, UK: Elsevier.

Zandbergen, P. A. (2007). Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. *BMC Public Health, 7*, 37.

Zandbergen, P. A. (2008). Comments on Boone et al., "Validation of a GIS facilities database: Quantification and implications of error." *Annals of Epidemiology, 18*(10), 823–824.

Zandbergen, P. A. (2008). A comparison of address point, parcel and street geocoding techniques. *Computers, Environment and Urban Systems, 32*(3), 214–232.

Zandbergen, P. A., & Chakraborty, J. (2006). Improving environmental exposure analysis using cumulative distribution functions and individual geocoding. *International Journal of Health Geographics, 5*, 23.

Zandbergen, P. A., & Green, J. W. (2007). Error in determining exposure potential of children at school locations using proximity-based GIS techniques. *Environmental Health Perspectives, 115*(9), 1363–1370.

Zarcadoolas, C., Boyer, J., Krishnaswami, A., & Rothenberg, A. (2007). GIS maps to communicate emergency preparedness: How useable are they for inner city residents? *Journal of Homeland Security and Emergency Management, 4*(3), 16.

Zeileis, A., Hornik, K., & Murrell, P. (2009). Escaping RGBland: Selecting colors for statistical graphics. *Computational Statistics and Data Analysis, 53*(9), 3259–3270.

Zenilman, J., Ellish, N., Fresia, A., & Glass, G. (1999). The geography of sexual partnerships in Baltimore: Applications of core theory dynamics using a geographic information system. *Sexually Transmitted Diseases, 26*(2), 75–81.

Zhan, F. B., & Noon, C. E. (1998). Shortest path algorithms: An evaluation using real road networks. *Transportation Science, 32*(1), 65–73.

Zielhus, G. A., & Kiemeney, L. A. L. M. (2001). Social epidemiology? No way. *International Journal of Epidemiology, 30*(1), 43–44.

Zimmerman, D. L., & Pavlik, C. (2008). Quantifying the effects of mask metadata disclosure and multiple releases on the confidentiality of geographically masked health data. *Geographical Analysis, 40*(1), 52–76.

Index

D

F

G

H

M

N

O

P

Q

R

S

T

U

About the Authors

Ellen K. Cromley, PhD, is Assistant Clinical Professor in the Department of Community Medicine and Health Care, University of Connecticut School of Medicine, Farmington. Previously, she was Professor of Geography at the University of Connecticut, Storrs. Dr. Cromley's research is primarily concerned with geographical patterns of health and disease, health facility location, and GIS design to support public health surveillance and intervention programs. She has published articles in geography, epidemiology, public health, and health services journals.

Sara L. McLafferty, PhD, is Professor of Geography at the University of Illinois at Urbana–Champaign. In her research she explores the use of spatial statistical methods and GIS for understanding health and social issues in cities, as well as gender and racial disparities in geographical access to services and employment opportunities. Dr. McLafferty has published articles in geography, epidemiology, and urban studies journals and is an Associate Editor of the journal *Health and Place*.

GEOGRAPHY

GIS and Public Health, Second Edition

Ellen K. Cromley and Sara L. McLafferty

"The authors have outdone themselves with the clarity, completeness, and relevance of this book. Whether you're new to the practice of public health GIS or an experienced GIS user, you will benefit from the wealth of material provided."
—Carol Hanchette, Department of Geography and Geosciences, University of Louisville

"Extensive coverage of important information about GIS applications in public health research and practice makes this an ideal text for my graduate-level spatial methods class."
—Robin Puett, School of Public Health, University of Maryland

"A much-needed introduction to GIS for students, faculty, and professionals preparing to research or practice in public health. The second edition illustrates many new developments in the use of geospatial data, including spatial analysis methods now widely available on the Internet."
—Gerard Rushton, Department of Geography, University of Iowa

"I have found this text more rewarding to use in teaching and to consult for my own work than any other single book in the field."
—Stephen A. Matthews, Population Research Institute, The Pennsylvania State University

Authoritative and comprehensive, this is the leading text and professional resource on using geographic information systems (GIS) to analyze and address public health problems. Basic GIS concepts and tools are explained, including ways to access and manage spatial databases. The book presents state-of-the-art methods for mapping and analyzing data on population, health events, risk factors, and health services, and for incorporating geographical knowledge into planning and policy. Numerous maps, diagrams, and real-world applications are featured. The companion Web page (*www.guilford.com/p/cromley*) provides lab exercises with data that can be downloaded for individual or course use.

New to This Edition

• Incorporates major technological advances, such as Internet-based mapping systems and the rise of data from cell phones and other GPS-enabled devices. • Chapter on health disparities. • Expanded coverage of public participation GIS. • Companion Web page has all-new content. • Goes beyond the United States to encompass an international focus.

Ellen K. Cromley is Assistant Clinical Professor in the Department of Community Medicine and Health Care, University of Connecticut School of Medicine, Farmington.

Sara L. McLafferty is Professor of Geography at the University of Illinois at Urbana–Champaign.

Cover design by Robin Lenowitz; images courtesy of the authors

ISBN 978-1-60918-750-7

9 781609 187507

The Guilford Press
72 Spring Street
New York, NY 10012
www.guilford.com