Preface

Why Write this Book?

This book has been written for anyone involved in public or community health with an interest in learning about a geographic information system (GIS). Although the focus is on infant health, many of the topics covered are transferable to other community health issues. It is also written for the upper division undergraduate or graduate student interested in GIS. Having worked with both of these audiences for several years, I understood the need for a breezy, readable text that complements what is already known and what is being taught. I wanted a book that was readable but at the same time included a lot of detail. The hope is that the overall message comes through, but that there is a point in keeping the book around as a reference when it is needed.

This book is written to fill a gap, as many other community health groups, especially local Healthy Start organizations at regional and national meetings, have asked me questions about how to incorporate a GIS into their program. So far, little has been written on this subject. Equally so, my GIS students have a plethora of "texts" that adequately describe the right buttons to push and how metadata should be written, but again, relatively little exists in terms of a detailed, contextual work that helps one understand why one would want to use a GIS. I hope these two audiences will start to meet up at some point, maybe even as a result of this book. A few of my GIS students have already been "turned onto" public health and medical geography. The demand for their services will certainly grow over the next few years.

The language of the book is meant to be more readable than that of a typical GIS text; indeed, one reviewer commented on the "subjective…informal approach to GIS…" and this is exactly the spirit I have hoped to capture: to demystify a subject for those who are not entirely comfortable in the world of computing.

But first, let me introduce you to our group at Louisiana State University...

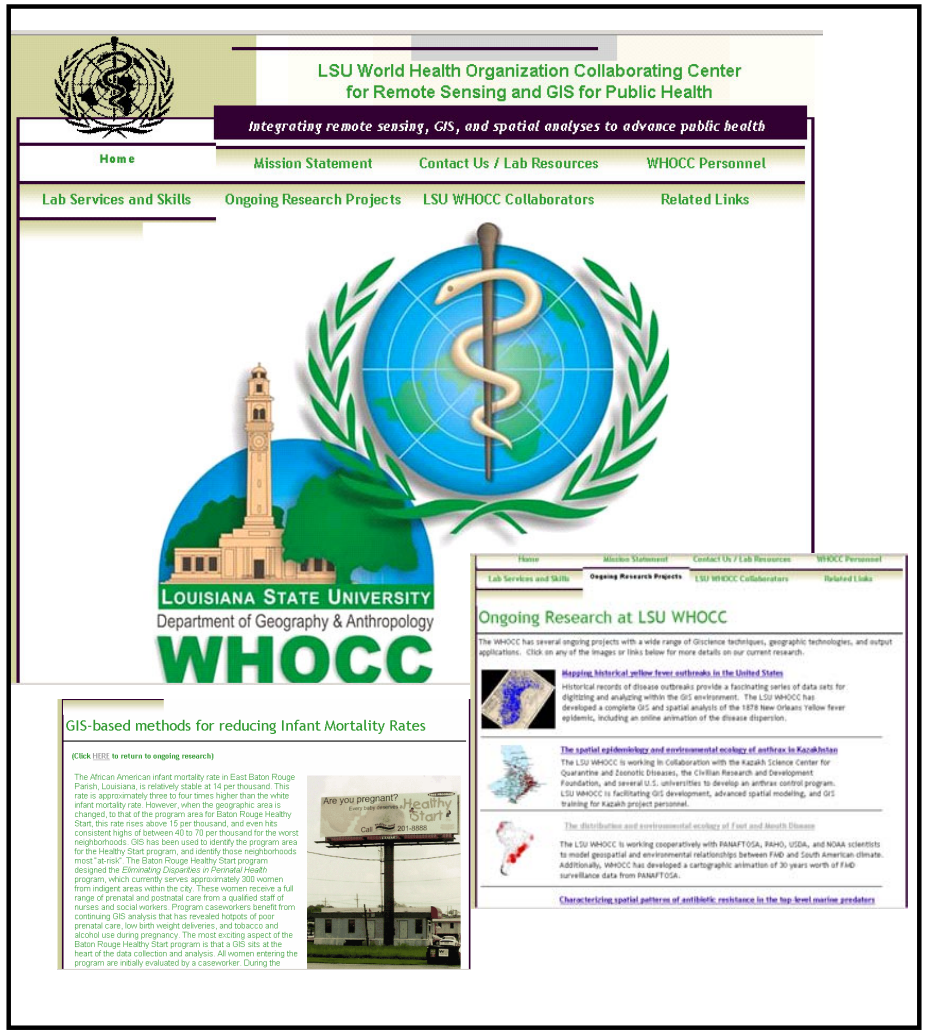
The World Health Organization's Collaborating Center (WHOCC) for Remote Sensing and GIS for Public Health is housed in the Department of Geography and Anthropology at LSU. This center is involved in several health initiatives, ranging from understanding the ecological and anthropological dimensions of Chagas disease in Mexico, antibiotic resistance found in sharks in the Gulf of Mexico, and anthrax in Kazakhstan, to the spatial modeling of diseases in history. The common factor in all these projects is the use of geographic information science (GISc) in solving health-related issues. The use of spatial technologies is also reflected in the mission statement of the center, and can be found on its Web site (www.whocc.lsu.edu):

- *To apply geographic technologies and techniques to public health problems, whether they be local, regional, national, or international in nature.*
- *To provide geographic information science (GISc) training and guidance to those in health related fields.*
- *To promote the use of GISc in public health.*
- *To advance GISc in association with public health through research initiatives with scientists and public health workers at other institutions.*
- *To promote and advance quantitative, geospatial modeling, and mapping techniques with special applications to public health questions.*

The Web site includes synopses about our research projects, as well as a few more "fun" items, like a cartographic animation of the 1878 yellow fever epidemic. One of those projects is the ongoing relationship between the WHOCC and the Healthy Start program of Baton Rouge (see Figure 1). This book has been 4 years in the writing, and it attempts to capture the essence of this relationship, from the initial understanding of the problem to the development of a Healthy Start initiative to eliminate racial disparities in infant mortality. GISc has been used on all stages of this journey. This book attempts to provide a working guide, first introduction, and insight into research possibilities for others about to embark on a similar journey. Chapter IX details this journey, drawing on the approaches described in other chapters to show how GIS was used in the initial grant writing, and how it is implemented today.

This book meets every one of the WHOCC mission statements, one of which is training and guidance, though this book should not be viewed as a training manual. Although techniques are discussed in depth, these are framed within the context of infant health. The complexity of working on health issues in some of our poorest neighborhoods is not skirted. It is not enough to know how to press the

Figure 1. The WHOCC Web site



right buttons, it is more important to know how to ask the right questions. The user should make the GIS work for him or her and the local community, and not the other way round.

As previously mentioned, this book has been written for the non-skilled GIS user. The first few chapters cover some of the GIS basics, including data sources and types, spatial analysis, and how to make maps. These chapters are not written to guide the user through a step-by-step learning process; rather, they are meant to provide an overview, a flavor, to those taking their first steps down the GIS path. It is more important to understand why and how we use GIS rather than knowing one particular software package. Once these initial building blocks have been gained, subsequent chapters expand and illustrate how more involved questions

relevant to a community health unit can be addressed. Although these chapters may appear too advanced for a first-time user, it is hoped that once interest is generated, the reader will want to progress on to the next level of implementation and inquiry. Several detailed research examples are provided, and it may be advisable for the first-time reader to skim through these until he or she feels more comfortable with GIS and geographical analysis. These more involved topics range from a selection of advanced spatial analysis methods to the preservation of confidential data within the GIS environment.

This book is also a suitable companion piece for the first-time GIS student taking a university course, whether specializing in the public health field or not. Many of the chapters follow the traditional course outline used in the introductory GIS classes taught at Louisiana State University. Scattered throughout the book are also examples of our current research initiatives, again written in an informal way, but giving the topic substance beyond just another GIS manual.

Overview of the Chapters

Many of the topics initially presented in the first three chapters are revisited in subsequent sections of the book. Indeed, spatial analysis is revisited twice more, such is its importance. In this way the book is somewhat hierarchical, allowing the reader to choose the depth of material required.

Chapter I

The goal of this chapter is to present an initial understanding of the role space plays in health analysis. Ask yourself, what disease or health outcome does not create a spatial pattern? The answer: none.

This chapter provides an introduction to many of the themes that run through the book, and should be seen as a springboard, a means to jump off into the subsequent sections depending on where the interest lies. The topic of negative infant health is first discussed here, along with an initial appreciation of the geography of health. Summary statistics for infant health outcomes are provided at a variety of geographic scales, descending from the national level to eventually end up at the zip codes served by the Baton Rouge Healthy Start. This geographic area will act as the canvas on which most of the book's examples are drawn. In this way it is hoped the reader will gain a degree of familiarity with the situation in Baton Rouge. This should also help clarify why the GIS approaches and solutions presented were applied.

This chapter also teases the reader with the first sense of how important geography can be in terms of understanding health surfaces. This is achieved by giving an initial insight into the complexities of working with neighborhood-level health data. Three main "risks" usually associated with pregnancy outcomes — smoking, poverty, and stress — are presented to illustrate this point. Understanding risks such as these are presented as simple "overlays" onto a city map in order to identify neighborhood vulnerability. This common GIS technique allows us to combine multiple datasets, often drawn from multiple disciplines, into one holistic impression. The topic of neighborhood vulnerability is again broached in Chapter IV, where the full weight of its complexity is described.

The chapter ends with what is hoped to be an interesting example of how space/geography and epidemiology/public health have long been bedfellows. An example from the 1800s is used to give the reader an insight into how the collection, display, and exploratory analysis of data can be used to understand a public health problem and begin the search for a mitigation solution. Apart from being an interesting case study, it is hoped these descriptions from the 1800s will excite the reader, in the same way they do the authors, about what can be achieved by approaching a health problem spatially.

Chapter II

The goal of this chapter is to provide the non-GIS reader (a community health investigator or even a student taking a GIS class for the first time) with an overview of what a GIS is, what data sources are available, and how these data can be manipulated to answer questions.

This chapter is not meant for those skilled in GIS use. An overview is presented about the capabilities of a GIS, how different spatial layers can be manipulated to answer different questions, and how the power of a simple query can reduce data so that specific neighborhood level investigations can be made. The reader is then given a tour of typical data types and how to bring them into the GIS, with particular emphasis placed on geocoding addresses, as this is probably the most useful data input procedure for community health units. Background "map" data such as 1:24,000 Scanned Topographic Maps (DRGs) and Digital Orthophoto Quarter Quadrangles (DOQQs) are also discussed.

As most forms of neighborhood analysis will involve socioeconomic information, the reader is briefly introduced to census data, in all their aggregations (and where to find them). Woven within these data concepts are other related issues such as the power and pitfalls of spatial analysis and visualization. *Power* includes the ability for local communities to control their own data, analysis, and understanding of local problems, while *pitfalls* range from the variations that

occur with different spatial aggregations to the issue of confidentiality and data access. Finally, four important pregnancy-related data manipulations are outlined, three of which can be created within a GIS: the Kessner and Kotelchuck Indices, which judge acceptable levels of prenatal care, and Perinatal Periods of Risk (PPOR), which link key components of the birth outcome to a fetal or infant demise. The fourth type of data are typically aspatial: Pregnancy Risk Assessment Monitoring System (PRAMS). These data can still be used as a template for other neighborhood level survey collections. These four data sources are probably unfamiliar to most GIS users, while most infant health researchers might not have thought about them in a GIS context.

Chapter III

The goal of this chapter is to introduce the reader to arguably the two most powerful uses of a GIS, the ability to extract understanding from spatial data in the form of maps and spatial analysis.

One of the first uses of a GIS for the newly initiated is the production of spatial displays or "maps." This chapter briefly covers the errors that can be faced in this process, from overreliance on spatial identifiers (zip codes are often filled in incorrectly), to how to avoid basic cartographic mistakes. Choropleth (graduated color) maps, dot maps, isoline surfaces (contours), and proportional symbol maps are all discussed. Color schemes, appropriate data manipulation, and the difference between quantitative and qualitative displays are also presented. Although this chapter is not a substitute for a good computer cartography class, or even a dedicated book on the topic, it should provide a reasonable starting point to making legible and acceptable maps.

The reader is then introduced to the power of spatial analysis. Although there has been an upswing in GIS use, especially in government agencies, the ability to ask the right spatial questions and know how to look for the answers is still sadly lacking. Yet having the ability to ask and answer spatial questions is probably the most powerful role a GIS can play for a community health organization. Questions could include where are our risks, what are our risks, how do our risks change over time and space? And one of the most important questions of course is, do our neighborhoods meet the thresholds set for federal funding?

Several available online spatial analysis packages are briefly reviewed. Important geographic concepts are also discussed that are relevant to most spatial analyses, such as how does the spatial aggregation of data affect results, are data spatial autocorrelated (events are not independent of each other), and what is distance decay? Again these concepts are kept simple, the goal being to inform rather than overwhelm.

The chapter concludes with a discussion as to how a neighborhood should be defined. This is important in terms of creating the geographies analyzed in the GIS. It is doubly important for a new form of analysis that combines both individual health outcome data (a low-birth-weight delivery), and neighborhood context (the education level or income base of the mother's neighborhood).

Chapter IV

The goal of this chapter is to provide an insight to a GIS user, whether newly introduced or skilled, to the complexities involved in working with community health at the neighborhood level.

A common criticism of GIS and spatial analysis is that people are "reduced" to numbers, that the analyst does not understand the context of the situation, and by extension the results have little validity in solving real-world problems. This chapter hopes to address this criticism by presenting the complexity faced in working with neighborhood level GIS analyses. It is heavy with citations, offering the reader many examples of further reading assignments, or even template research ideas. The focus of the chapter is to identify some of the risks that can lead to a negative birth outcome. In so doing, risks that may not directly, traditionally, affect a delivery are also covered. This is for two reasons: First, because they are still stress inducing which is related to pregnancy outcome, and second, because a broad spectrum of academics believe that the social context determines many health outcomes. This social context does not begin on the first day of a pregnancy.

Following the disease ecology approach often championed in medical geography, risks are broken into biological, behavioral, social, cultural, economic, and environmental categories. Many of these risks combine to create stress-inducing neighborhoods. Simple "choices" such as smoking during pregnancy, or attending a prenatal visit, soon become more complex, and are as much to do with the living environment as any individual behavior.

Neighborhood risks, such as crime, poor building materials, presence of lead, proximity to busy roads are all presented as being inputs into a complex milieu. It is important to appreciate this complexity. GIS is a tool that can begin to build a holistic vulnerability surface by layering one risk on top of another. However, the complexity may be missed if only GIS analysis is used. Qualitative approaches, such as focus groups, are needed to help explain why hot spots occur where they do. Paraphrasing Stan Openshaw, a famous geographer and one of the original GIS modelers, a cluster is only a data anomaly until it is given context.

Chapter V

The goal of this chapter is to introduce the novice GIS user to a series of spatial analysis approaches that can be applied in a community health environment.

As has been previously mentioned, the ability to ask the right questions and to search for answers is vital for a community health organization. This chapter assumes little or no previous background in spatial analysis and leads the reader through the concepts of designing a spatial question, how and where to collect sample data, and the pitfalls of working with spatial data. One such data issue comes with having an adequate sample, and a background population against which to compare or test the sample, what is known in the game as numerators and denominators. For example, if one house has a birth, and that baby dies, the house and possibly the road will have an infant mortality rate of 1,000/1,000. Similarly, if a car crashes with twins on board, a hot spot of mortality will likely show in the neighborhood. However, if we demand a minimum number of births for any area to counteract this "law of small numbers," what happens when two babies die in the same apartment building to two different mothers when there were only ten births in total? Is something going on there?

The reader is also presented with a simple analysis approach, a version of the difference of proportions t-test that can easily be created in a spreadsheet such as Excel. Simply put, the GIS can reduce a dataset by both queries that are spatial (for one zip code or half a mile around a clinic), and aspatial (by all teen births). These data are exported to Excel, where a comparison is made to the larger population, such as the whole city. The question in this case is, is there a significant difference to teen mothers in this area? Certainly, more powerful techniques are available, and some of these will be presented in Chapter VI, but there is an argument for keeping it simple. The novice GIS user, along with any introductory GIS student, as long as he or she has had at least one basic statistics class, can perform the analysis example provided in this chapter.

Chapter VI

The goal of this chapter is to expose the reader to a series of techniques on the "cutting edge" of spatial analysis. These techniques can be applied to find spatial intensities or "hot spots" as well as causative associations.

The inexperienced GIS user might want to skip this chapter, at least during an initial learning phase. For those who are more used to GIS analysis, or with a

reasonable background in statistics, this chapter is designed to discuss approaches to find clusters (including nearest neighbor hierarchical clustering and the spatial filter), fill-in space (kernel density analysis), and spatial associations (spatial regression). Again detailed examples are provided of these approaches.

Several excellent books have been written on GIS and spatial analysis, and this chapter only scratches the surface. Nevertheless, Chapters V and VI should provide a reasonable armory for a GIS user irrespective of where he or she falls on the expertise continuum.

Chapter VII

The goal of this chapter is to present an example of how a data condition can affect GIS analysis. This condition is how mobile are pregnant women? It is hoped the reader will begin to understand how important data are — and not fall for garbage-in, garbage-out.

Spatial variation will always occur between analyses. It is very unlikely that an analysis will reveal a hot spot that can be used to precisely identify subsequent hot spots — unless there is an environmental association, such as proximity to a landfill. In order to counter this problem, the areas of analysis have to be large enough to reduce the effect of this "spatial error." This leads to another problem, in that too large an area does not really help outreach.

This is the problem: What is the point of performing a GIS analysis unless we can extract some good out of it, such as targeting resources? In order to effectively target resources, relatively small areas are needed to be identified — and yet what is the point in sending outreach teams into a neighborhood if the problem by that time has cycled out? This chapter presents different research approaches that have been used to identify how spatially stable results are.

A further issue is raised — although spatial variation will always exist, what happens if the women themselves are mobile? If we create a hot spot map of infant death — what is the important residence? That on the birth certificate, or on the death certificate, or at a certain stage of the pregnancy? Unfortunately, the chapter offers no gold standard in terms of a solution, because it is not really discussed in the geographic analysis literature.

The purpose of this chapter is therefore twofold: to initiate this discussion, and to make the reader aware that most data have limitations.

Chapter VIII

The goal of this chapter is to provide some guidelines as to how to map health while preserving patient confidentiality.

In this chapter, it is argued that national standards (e.g., HIPAA) are good at preserving somebody's statistical information, usually recorded as text or in a spreadsheet format, but lack appropriate rules for visualizing this information on maps. Privacy rules for visualizing confidential data on maps (i.e., spatial confidentiality) become more and more important, as governmental agencies increasingly use GIS as a standard tool for collecting, storing, analyzing, and disseminating spatial information. One possibility to preserve somebody's spatial confidentiality is by manipulating the location of health records just enough to protect the confidentiality of individuals, but not too much to change essential visual characteristics of the true, original spatial distribution of those records. This method is referred to as "geographic masking" and this chapter evaluates the usefulness of different geographic masking methods in preserving spatial confidentiality.

Chapter IX

The goal of this chapter is to show the reader how a GIS might be woven into a community health initiative, from inception (grant getting), through outreach, basic research and the day-to-day running of the program.

This chapter may be the most important for many readers, as it can be seen as a quick reference as how to involve GIS within a grant application. The chapter then goes on to show how GIS has been implemented in all aspects of the program over the following four years, from outreach to a data collection tool. Although many of the actual approaches described in this chapter have previously been mentioned, it is here where they are set in a context.

If one message comes from this chapter, is should be that a stronger working relationship should exist between community health organizations and academia. There are many able geography graduate students who would love to bring their skills to a topic as important as improving infant health. These collaborations could easily end up being theses and dissertations, resulting in even more applicable research for the health unit. So go seek out your local geography department…

Chapter X

The goal of this chapter is to open everyone's eyes to the way pregnant women are missed in disaster response and recovery planning.

Although this chapter might seem somewhat of a departure from the rest of the book, it raises an important point. Pregnant populations, especially those in neighborhoods with high pregnancy risks, are often the most vulnerable whenever disaster strikes. These women often have little family support, live in densely packed inner-city areas, and have no transport. In addition, these women are pregnant. Although vulnerability mapping is not a new concept, very little has been written about the specific risks faced by pregnant women. During a disaster, and often after the event has occurred, the victim suffers from an elevated stress load. Stress has been found to link directly, or indirectly through coping measures such as smoking, to negative birth outcomes.

This chapter describes measures that can be used to reduce this stress load, and by using the type of GIS incorporated at the Baton Rouge Healthy Start — (near) real-time help can be directed to those most at risk. The chapter ends with a suggestion as to how Homeland Security funding could be channeled to develop community health GIS systems in the name of syndromic surveillance.

Chapter XI — Guest Chapter by Jackie Mills

The goal of this chapter is to explain some of the differences a community health unit will face if they come from a more rural location.

The authors are fully aware that many community health centers are located in rural areas. Indeed, the Monroe Healthy Start in Louisiana works with rural counties. Although most GIS books tend to focus on either urban or rural (usually the former), many of the points raised in this text are still relevant to both, though with a little tweaking. We therefore asked Jackie Mills, who has worked on GIS and poverty in the Mississippi Delta for the last six years, to give us her insight into the particular problems likely to be faced by a rural health unit implementing GIS.

This chapter highlights the importance of understanding the geography of an area of analysis in regard to its urban or rural nature. This distinction has ramifications not only for types of health care issues specific to rural places, but in particular to the availability and applicability of data in a GIS environment. This chapter discusses the complexity of rurality, that a firm or best definition for

being rural does not exist. Several agencies of the U.S. government provide definitions based on county or census designations, but the scale varies and these classifications must be researched in respect to your particular place and health care interests.

Some of the "risks" faced in rural areas will be distinctly different. For example, exposure to agricultural chemical applications has been noted as a health risk to pregnant and lactating women in rural areas. The agricultural structure (agristructure) of a rural community can also impact the availability of jobs, services, and overall well-being of rural places which impacts the age of pregnancy, health of pregnancy, and infant health in general. In addition, the existence of rural ghettos in the southern Cotton Belt of the U.S. has its own risks, much like those faced in urban ghettos, yet these places are not often considered when examining health issues in the rural south.

Finally, this chapter delves into the types of data that may be important in a GIS analysis of rural health issues. Concerns of confidentiality, availability, and scale are particular areas of note, as are some helpful ideas for making the best use of data.

Summary

I hope you enjoy this book. Working with the Baton Rouge Healthy Start has been a fulfilling experience. This experience has ranged from presenting at a congressional breakfast to working late nights with nurses and social workers as we plan the next step of our program. I have met some wonderful people along the way, and, hopefully, employed GIS to its maximum potential. But there is so much more that can be done. I would love to see a national approach by organizations such as Healthy Start providing the resources to encourage all community health groups to incorporate the technology. But for now, this book will have to do. If we have made any mistakes in the writing, as I am sure we have, I apologize. This was our best effort at covering a large subject in such a way to make it palatable and hopefully, a somewhat enjoyable read.

All writers royalties, as such they are, will be fed directly back into the research described in this project.

Andrew Curtis
Michael Leitner