LSU Doctoral Dissertations   Graduate School

2007

Spatial and temporal structure of typhoid fever in Washington, D.C., 1895-1909: a geographic information systems exploration of urban health concerns

Sarah Elizabeth Hinman
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

 Part of the Social and Behavioral Sciences Commons

SPATIAL AND TEMPORAL STRUCTURE OF
TYPHOID FEVER IN WASHINGTON, D.C., 1895-1909:
A GEOGRAPHIC INFORMATION SYSTEMS EXPLORATION
OF URBAN HEALTH CONCERNS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Geography and Anthropology

by
Sarah Elizabeth Hinman
B.A. Mary Washington College, 2000
M.A. Ohio University, 2002
May 2007

To Karl Crawford...
Thank you for showing me this path

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The period between 1880 and 1920 was one of the most dynamic in the history of medicine. Morbidity and mortality rates for infectious diseases dropped quickly. Concurrently, miasmatic theory gave way to germ theory. Many of these dynamic changes occurred in the urban centers of North America, which were also entering into a period of dramatic growth and change. Following the 1905 completion of infrastructure improvements intended to improve public health in Washington, D.C., typhoid fever rates unexpectedly increased. Previously, for mitigation purposes, Dr. George Kober investigated a typhoid epidemic in 1895, and as a result of the 1906 increase in typhoid morbidity the United States Public Health and Marine-Hospital Service conducted investigations between 1906 and 1909 to better understand the origins of typhoid in the city. These studies include dot maps of typhoid case locations for the entire 1895 epidemic, at two-week intervals for 1906, 1907, and 1908, and monthly intervals for 1909. These point locations are used to construct a geographic information system (GIS) displaying the spatial distribution of individual typhoid cases. The creation of this GIS allows for the investigation of urban typhoid at a localized geographic scale. The temporal resolution of the data and supplementary data included in the reports provides an opportunity to explore urban typhoid within years, between years, to compare morbidity to mortality, and to compare the spatial pattern of multiple diseases. This dissertation describes the creation of this GIS and the results of the spatial analyses using Ripley's K-function and the G_i* statistic to evaluate spatial clustering patterns. The G_i* statistic identified localized hotspots that refute the conclusions of the original reports. Typhoid clusters varied in size and location, and lacked temporal stability. The findings

of this dissertation indicate that typhoid in early twentieth century Washington, D.C. originated from multiple sources whose impact decreased over time. Studies of this type make use of geospatial approaches unavailable when the original data were collected, in order to investigate potential patterns of typhoid fever invisible a century ago. This research helps to provide a better understanding of the historical geography of urban health in general.

CHAPTER 1
INTRODUCTION: HEALTH IN THE PAST

Introduction

The period between 1880 and 1920 was one of the most dynamic in the history of medicine. Morbidity and mortality rates for infectious diseases dropped quickly during this time. Concurrently, miasmatic theory was giving way to germ theory, bacteriological testing, and anti-toxin development. Many of these dynamic changes occurred in the urban centers of Europe and particularly North America, which were also entering into a period of dramatic change and growth. The urban centers of the United States grew in terms of physical size, population, and their built environment. The phrase "built environment" refers not only to the structures above ground, but also the below ground built environment of drinking water distribution pipes, storm sewers, and sanitary sewers (Boone 2003; Melosi 2000). It is this intersection of health improvement, urban change, and infrastructure development that is the backdrop for this dissertation and its focus upon the spatial aspects of typhoid fever at the turn of the twentieth century.

Cliff et al. (1998) briefly mention the continued role of infectious diseases in the era following W. H. Stewart's 1969 declaration that the war on pestilence was over, and infectious diseases were in the past. These authors discuss the need to not only analyze new infectious diseases, but also disease patterns from the past in order to establish a baseline of contagion (Cliff et al. 1981, 1998; Trevelyan et al. 2005b). In so doing, Cliff et al. (1998) display the error of Stewart's declaration and highlight the number of infectious diseases that currently impact the globe. In the most recent of this team's publications, Trevelyan et al. (2005b) focus on the 1916 poliomyelitis epidemic in the

northeastern United States with the purpose of better understanding the spatial structure of the epidemic though the utilization of currently available computing and analysis techniques. This approach not only provides insight into the historical significance of the disease, but also epidemiological characteristics of an epidemic.

Trevelyan et al. (2005a, 2005b) succinctly summarized the relevance of studying historical disease surfaces when they wrote that this type of research allows for the re-examination of the spatial structure of a disease. Another way to think about this subject is that while many infectious diseases; tuberculosis, typhoid fever, cholera, diphtheria, poliomyelitis; might be better understood biologically today, their spatial structures have not been carefully considered through the lens of late-twentieth century developments in spatial analysis. Geographic information systems (GIS) and the associated forms of spatial analysis provide fresh tools to consider epidemic datasets from the nineteenth and early twentieth centuries. One approach to investigating such epidemic structures is to control the geography of the investigation, and concentrate on a single disease. Most studies investigate historical disease(s) across large areas using aggregated data (Trevelyan et al. 2005a, 2005b; Smallman-Raynor and Cliff 1999; Patterson 1994; Pyle 1969; Cliff et al. 1998, 1981, 1986). These studies have been insightful in explaining temporal changes in diffusion processes, for example, but a GIS allows for more sophisticated individual level, or city block level analyses. The benefit of investigating data at this geographic scale is obvious, actual patterns and processes of spread from person to person, or neighborhood to neighborhood can be identified.

Typhoid fever is a suitable disease for an epidemic-structure investigation as 1) its symptoms allow for confidence in data quality and 2) as with other water-borne diseases

such as cholera, contemporary epidemiological investigations were often rich with spatial

context. A number of reports produced for Washington, D.C. in the nineteenth and early

twentieth centuries contain maps of typhoid mortality and/or morbidity locations. In this

study, a GIS is built using spatial data from these reports, and the hypotheses and

conclusions of the original authors are tested. In addition to re-testing the original

hypotheses, exploratory data analysis is employed to investigate questions not described

or possible during the time-frame of the original report. Before presenting the exploration

of typhoid data from Washington, D.C. this chapter will first outline the three intersecting

backdrops needed for this study: an understanding of typhoid fever, a review of key

literature as to place this study in context, and an overview of Washington, D.C. during

the time in question.

Urban health and its geography is a construct of multiple inputs that converge to

form identifiable disease surfaces. These inputs can include political decisions, the local

environment, the nature of medical understanding, economic and social constraints,

individual behavior, and the presence or absence of disease, or the conditions supporting

disease. These categorical parameters on health distribution hold for both health

inequalities in the present as well as the past (Gatrell 2002; Boone 2003; Olson 1997;

Haines 1995; Williams 1992; Thornton and Olson 2001). In the late nineteenth century,

good health largely depended on a person's socio-economic status. A higher income

allowed greater residential choice within a city, but the local environment provided by

that choice also played a role in health risk (Williams 1992, Haines 1995; Thornton and

Olson 2001; Olson 1997). While this dissertation focuses upon the spatial aspects of

typhoid fever, it is undeniable that social geography, along with the political decisions

that facilitated public health infrastructure, and the developments in medical knowledge that urged political actions all contributed to the health of the city (Condran and Crimmins-Gardner 1978; Boone 2003; Boone and Modarres 2006; Williams 1992; Melosi 2000; Curson 1985; Thornton and Olson 2001). As such, each will be briefly discussed before the spatial structure of the epidemics under investigation are analyzed.

This dissertation focuses primarily on a single disease, typhoid fever, in a single city, Washington, D.C., for the period 1895 to 1909. Although other water-borne diseases, such as cholera, have been investigated in urban environments, most notably by John Snow in 1854 (Brody et al. 2000; Snow 2002), similar works on typhoid are relatively lacking. Detailed urban spatial analyses of epidemic typhoid is even more scarce, partly due to the lack of historical studies adopting GIS as a tool of inquiry, although this trend is slowly changing (Knowles 2000, 2002; Gregory 2005; Yu and Christakos 2006).

This dissertation adopts the journal-style approach of each chapter being a publishable unit. Five chapters investigate different facets of the selected data and/or approaches to spatial analysis. The remainder of this introductory chapter provides the necessary background for each of these analysis chapters, explaining typhoid fever itself and placing the disease in the context of the period 1895 to 1909. A review of key literature related to historical health studies and medical geography is provided, as is a brief history of key events in Washington, D.C. that relate to typhoid fever and health in general. Chapter 2 outlines the data entry process and the methodology used in four of the five analysis chapters. Chapters 3 through 7 describe each of the data explorations embarked upon in this dissertation. In each of these chapters the premise and problem

statement is explained at the outset of the chapter followed by a methodology, results, and discussion section. Chapter 8 is a concluding chapter that ties the findings of chapters 3 through 7 together, places the overall findings within the larger context, and provides suggestions for future research.

The Geographic Setting of Typhoid Fever

Typhoid fever is one of many gastrointestinal diseases that continue to affect certain populations around the globe. The disease is media-borne and most often transmitted through contaminated water or food, especially in parts of the world where water treatment and sewerage infrastructure are limited (Parker and Parker 2002). Typhoid fever is caused by *Salmonella typhi,* a bacteria endemic to humans. The bacteria are carried in the bloodstream and gastro-intestinal tract of infected persons and shed through feces. Typically, through lax sanitary practices either at the individual level or through a lack of sanitary infrastructure, typhoid bacteria are transferred from feces to drinking water, milk, or other foods intended for human consumption (Parker and Parker 2002; CDC 2005a).

The bacteria itself was identified early in the development of the germ theory of disease causation. Discoveries related to *S. typhi* occurred first in 1880 and then again in 1884, when Robert Koch and members of his institute first cultivated the typhoid bacteria. Although medical professionals could identify the bacteria causing typhoid and knew that the bacteria survived in water and milk, a more complete understanding of the means of typhoid transmission did not develop until 1902 (Whipple 1908). It was at this time that scientists grasped the nature of the asymptomatic carrier of the disease (Brock 1988). These carriers shed *S. typhi* in their feces, but do not present any signs of illness

(Mortimer 1999). The most famous typhoid carrier was Mary Mallon, better known as Typhoid Mary (Leavitt 1996; Bourdain 2001). Approximately three percent of any population were typhoid carriers, thus the scope of their potential threat to public health was limited compared to contaminated water supplying an entire city (Leavitt 1992, 1996).

Typhoid continued to be common in both North America and Europe during the first quarter of the twentieth century. As with other diseases, typhoid rates dropped during the last quarter of the nineteenth century as a result of, but not limited to, improved epidemiological understanding and general improvements in the urban built environment. The specific reasons for the decline in typhoid rates are usually attributed to alterations made to the urban built environment that included changes in water supply sources, water treatment, the adoption of comprehensive sewerage systems, and the sanitization of milk production (Casner 2001; Parascandola 1997; Condran and Crimmins-Gardner 1978; Atkins 1992). These developments began during the second half of the nineteenth century and continued into the early part of the twentieth century as communities began to implement major public works projects for a variety of reasons including, but not limited to public health (Condran and Crimmins-Gardner 1978; Casner 2001).

There are approximately 400 cases of typhoid annually in the United States today; 70 percent of those are contracted in other countries (Parker and Parker 2002). However, even with an understanding of the causes and transmission mechanisms of typhoid, outbreaks continue to occur in other countries. For example, the World Health Organization (WHO) posted updates in June 2003 about a typhoid outbreak in Haiti and

in 2004-2005 for the Democratic Republic of the Congo. In Haiti, at least 200 cases and 40 deaths were reported to the WHO and a team was sent by the Pan-American Health Organization and Haiti's Ministry of Health to control and investigate the epidemic. Eighteen months later in the Democratic Republic of the Congo, between September and January 42,554 cases and 214 deaths from typhoid were reported to the WHO. A lack of clean drinking water and poor sanitary conditions were cited as the cause of this particular outbreak (WHO 2003, 2004, 2005a).

Typhoid fever is not the only infectious disease that has been largely eliminated in North America and Europe but continues to exist elsewhere. Cholera is another example of a gastro-intestinal ailment that no longer exists in places with modernized sanitary infrastructure, but often severely impacts other populations (CDC 2005b). The Indian subcontinent is particularly vulnerable to outbreaks of cholera (Emch 1999; WHO 2001), and a cholera outbreak began in July 2005 affecting eight countries in West Africa, and in Sudan beginning in April 2006 (WHO 2005b, 2006). Cholera has received attention in research, both as a nineteenth century disease, and as one that persists in pandemic form today (Patterson 1994; Patz et al. 1996; Pyle 1969; Emch 1999). Typhoid, on the other hand, has received limited attention despite a similar means of transmission and continued presence in some parts of the world. This limited attention is particularly visible in the historical literature, although there are exceptions (Smallman-Raynor and Cliff 1999, 2001; Cliff et al. 1998). Typhoid was a more common disease than cholera during the nineteenth century, and was considered endemic in most large cities, with cases and deaths occurring annually. Therefore, typhoid can provide a consistent record

for a geographic investigation of a water-borne disease within a single community (Board of Commissioners 1880, 1898; Rosenau et al. 1907; Sedgwick 1893).

In his comprehensive volume on typhoid fever and its continued presence in North America, George Whipple identified in the following passage two key elements that defined the continuation of problems with this disease in the early twentieth century.

> It is easier to keep the pig from getting out of the pen than it is to catch him when he is out. It is easier to keep the sparks from scattering from the fireplace than it is to put out the conflagration that the sparks have kindled. So, also, it is easier to prevent the germs of typhoid fever from leaving the sick-room than it is to avoid them, or to discover and destroy them after they are out of bounds. But, in spite of all precautions, typhoid germs will escape through all the barriers (Whipple 1908, 41).

In other words, of course prevention is the best "cure" for infectious diseases, but at the time, despite a fairly complete knowledge of what caused typhoid fever, how it was spread, and how to prevent the disease from affecting a community, this was not enough to eliminate the risk. Of particular interest is how even the epidemiological investigations described by Whipple (1908) often included a geographic assessment of typhoid cases or deaths as part of the analysis process. Contemporary analytical capabilities limited these geographic elements of epidemiological investigations to visually interpreting patterns of disease. Advances in computer technology and spatial analysis techniques developed in the latter part of the twentieth century now allow for these historical disease patterns to be further investigated. In addition to re-examining the questions of the period, more insight can be gained into the spatial structure of a disease in order to gain a fresh perspective of its spatial epidemiology. At a broader level, this research can add insight

into the specific role urban structure plays in an epidemic, and more generally, how the urban environment can affect the geography of a disease.

Supporting Literature

Studies that explore historical spatial health issues tend to adopt one of two perspectives. There are studies that investigate the different social and geographic variables related to mortality decline in the late nineteenth and early twentieth centuries. While a number of other studies attempt to better understand and model spatial patterns and the diffusion of disease. Among medical geographers, research tends to focus on more contemporary diseases (Emch 1999; Gatrell 2002; Cromley and McLafferty 2002; Meade and Earickson 2000; Getis et al. 2003). Historical medical geography studies tend to be limited with the exception of the diffusion and mortality studies mentioned above, and a few other works primarily concerned with the social aspects of disease (Craddock 1998; Kenny 1995; Frenkel and Western 1988; Whitmore 1991; Lovell 1992). A final area that provides context for this study comes from public health, infrastructure, and urban environmental history.

Mortality Studies

Quality of life and differences in health vary spatially at all geographic scales. One city will bear a greater health load than another, and specific neighborhoods within that city will also be more prone to illness than others. This is as true for modern communities as it is for the past (Lee 1991; Williams 1992; Smith and Easterlow 2005). While undoubtedly socio-economic conditions and problems play a major role in explaining this spatial variation in health, so too does the environment of the individual. This environmental risk can also be found at all geographic scales, with health problems

in subtropical locations being different to those of more temperate climates. Urban proximity to wetlands or marshes (in the past), or heavy industry (currently), for example, can create localized elevated disease presence.

Beginning in the 1870s American mortality rates began to decline, signaling the start of what is often called the epidemiological transition (Elman and Myers 1999). During this transition, urban mortality rates, previously much higher than rural mortality rates, began to decline faster than rural rates and continued to decline into the twentieth century (Preston and Haines 1991; Elman and Myers 1999). Numerous studies investigate differences between urban and rural mortality rates, causes of mortality decline, and often attempt to determine the key variables associated with patterns of high mortality in urban settings (Preston and Haines 1991; Williams 1992; Woods et al. 1988; Lee 1991; Meeker 1972; Condran and Crimmins-Gardner 1978; McKowen 1976). Typically, high urban mortality is associated with infectious diseases that spread more readily through dense populations and places with poor environmental conditions, such as overburdened privies and unclean drinking water (Condran and Crimmins 1980; Williams 1992; Preston and van de Walle 1978; Thornton and Olson 2001).

Often authors attribute the increased urban mortality rate to an undefined factor called the "urban effect" (Preston and Haines 1991). This effect, also termed the "urban-sanitary-diarrhoea effect," essentially refers to the inability of research to identify the exact explanatory factors that caused higher mortality rates among urban residents, and specifically infants, than among rural residents (Preston and Haines 1991; Williams and Galley 1995; Woods et al. 1988; Mooney 1994). Typhoid's highly infectious nature and

easy means of spreading through water or milk fits well in the urban-sanitary effect equation (Condran et al. 1984; Preston and van de Walle 1978; Higgs and Booth 1979).

One suggested cause of urban mortality decline is the sanitization of the urban environment (Meeker 1972; Preston and van de Walle 1978; Watterson 1986). Simply put, particularly by historians, once cities built piped water systems and sanitary sewer systems, mortality rates declined and public health concerns shifted from infectious, and essentially preventable diseases such as typhoid and malaria, to chronic and lifestyle related diseases (Rosenkrantz 1972; Duffy 1968, 1974; Elman and Myers 1997). Research that uses typhoid as a proxy for sanitary conditions is usually part of the larger body of literature focused on changes in mortality patterns during the late nineteenth and early twentieth centuries (Condran and Crimmis-Gardner 1978; Higgs and Booth 1979; Preston and van de Walle 1978; Condran et al. 1984; Watterson 1986; Woods et al. 1988; Thornton and Olson 2001).

Studies focused on turn of the twentieth century health or mortality and that use typhoid fever as a disease of investigation can usually be subdivided into two types. First, typhoid is used as a proxy for either sanitary conditions (higher rates of the disease in a city ward indicating unsanitary conditions) or as an indicator of the impact of sanitary improvements (typhoid rates changed after the introduction of water filtration) (Condran and Crimmins-Gardner 1978; Condran et al. 1984; Higgs and Booth 1979). Second, the diffusion process of typhoid spread is analyzed, usually at a regional or national level, and often in combination with a number of other parallel diseases (Cliff et al. 1998; Smallman-Raynor and Cliff 1999, 2001). In both cases, typhoid's relationship with urban sanitation is clear, but in neither group has the spatial pattern of the disease been studied

within a single city. This is noteworthy since many historical reports did attempt such large scale investigation (Sedgwick 1893; Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911; Kober 1895; Provincial Board of Health Ontario 1912).

Typhoid fever often plays a prominent role in mortality studies as an indicator of health improvements, as mentioned previously. Condran and Crimmins-Gardner (1978) compared the effect of environmental improvements to the disease-specific mortality rates of known water-borne diseases, such as typhoid and diarrheal diseases. To achieve this, the authors measured environmental improvements through the variables of miles of water and sewer pipe, and the construction and maintenance expenditure figures for these features. Condran and Crimmins-Gardner (1978) concluded that water and sewer reforms were not primary factors in urban mortality decline as evidenced by a limited association with water-borne disease decline. Instead of supporting the hypothesis of health benefits from infrastructure improvements, they suggested that income might be a more important explanatory factor of urban mortality decline (Condran and Crimmins-Gardner 1978).

In a later study by Condran et al. (1984), the authors discussed the importance of income to the decline in mortality in Philadelphia. They argued that income influenced all of the other suggested changes that caused the decline in urban mortality rates. "Increased per capita income leads to increases in the purchase of health-enhancing goods–especially better food and better housing" (Condran et al. 1984, 163). Additionally, increased income led to improvements in public services by increasing the tax money available for civic improvements. With this idea of environmental improvements stemming from increases in income, the authors investigated the effect of

water filtration on typhoid fever deaths and other water-borne disease deaths. They

argued that while water filtration put an end to typhoid epidemics, it did not have the

same effect on other diseases. Condran et al. (1984) concluded that the most important

factor in Philadelphia's mortality decline was the general increase in income and living

standards, but concurrent improvements to public health infrastructure also contributed to

the decline (Condran et al. 1984). One problem with using miles of pipe laid as a proxy in

these two studies is that the existence of a sewer pipe does not necessarily indicate that

households were hooked up to the system. It only means that residents could chose to

make use of the sewer system initially (Green 1963; Colten 2002).

Higgs and Booth (1979) used city-wide typhoid rates as proxies for urban sanitary

conditions in their study of mortality differentials between the wards of 17 American

cities. While it was not the ideal variable to use, the authors suggest that typhoid fever

rates per 1000 population in each city are the best available indicator of urban sanitation

given its known association with clean water and sanitary sewage removal. Typhoid was

a significant variable explaining mortality differences in this study, particularly among

children. In the summary of typhoid's significance, Higgs and Booth (1979) critique

Condran and Crimmins-Gardner's (1978) use of expenditure on piped water and sewer

lines as a way to understand the impacts of sanitary reform on mortality rates for the

same reasons Meeker (1972) advised against the use of this variable. By using deaths

from typhoid fever per 1000 population as a proxy, Higgs and Booth (1979, 365)

attempted to simply address the "general sanitary condition of [each] city, and of the

water supply in particular." Even this proxy is problematic, as addressed by the authors,

in that it is not a direct measure of the sanitary conditions in each city, but rather a

"stopgap measure" chosen only after being unable to find data pertaining to actual water quality and sewerage (Higgs and Booth 1979, 365). In their opinion, though, typhoid mortality rates provided a more empirical proxy than sewer connections, miles of water mains, or public expenditure, since those choices vary in accuracy and bias from city to city.

Contrary to other studies, Preston and van de Walle (1978) found that environmental improvements did, in fact, play a more important role in French urban mortality decline than increased wealth. Comparing the three economically similar urban areas (defined as a major city and its surrounding area) of Paris, Lyon, and Marseille, Preston and van de Walle found significant differences in the mortality rates between these cities. If Condran et al.'s (1984) hypothesis that sanitary improvements only augmented increases in income were true, then Preston and van de Walle's (1978) urban mortality rates would have tended toward uniformity, since the economic situations of the French cities were similar, the sanitary conditions were not. Preston and van de Walle examined the implementation dates of sanitary improvements compared to the dates when water-borne disease deaths, particularly typhoid, declined. They concluded that the timing of water-borne disease mortality decline and sanitation improvements coincided. Thus, sanitary improvements were the primary factors associated with urban mortality decline (Preston and van de Walle 1978).

The preceding summary of some mortality pattern literature illustrates how this type of research uses typhoid as an indicator of sanitation, but that there is little agreement on the importance of sanitation in mortality decline. These studies do not usually consider their research questions based on a single city and the patterns within

that city. When the study does focus on an individual city, the analysis is either of the city as a whole (Condran et al. 1984; Preston and van de Walle 1978) or of large areas to provide comparisons (Williams 1992).

None of the mortality studies mentioned above consider the *geography of typhoid* within a single city. An alternative group of historical investigations, what could be called diffusion studies, do consider the geography of typhoid and other diseases, but again these studies usually do not confine themselves to typhoid patterns within an individual city.

<u>Diffusion Studies</u>

Much of the historical medical geography research is quantitative in nature and tends to focus on diffusion patterns, ranging from studies of measles in Iceland to cholera pandemics in Russia (Cliff et al. 1981; Patterson 1994). Recent disease diffusion studies incorporate typhoid at a number of different geographic scales including the military camp, multi-city, or country levels (Smallman-Raynor and Cliff 1999, 2001; Cliff et al. 1998). In addition to the variation in disease and locale of study, so too do diffusion studies differ in approach from descriptive to model development. The following summarizes some of the key diffusion studies. One feature common among these diffusion studies is the use of data over large areas, ranging from a county, through a number of provinces, to an entire country (Cliff et al. 1981; Patterson 1994; Pyle 1969; Wilson 1993).

Diffusion studies focus on how diseases move through a population, often re-defining approaches based on the famous diffusion models introduced by Hägerstrand (1953) (Cliff et al. 1981). Typical diffusion related questions include velocity and

directionality of disease movement, cyclicity, and diffusion structure. Disease diffusion can be categorized in one of three ways: re-location, contagious (contact), and hierarchical. Re-location diffusion occurs when a disease arrives in a place that is not part of the geographic area from where it originated. For example, cholera would not naturally spread from Europe to North America without an infected individual "re-locating" from Europe to a port in North America. From a point of origin a disease can then spread in a contagious manner from person to person. So that person A has a conversation with person B, coughs on person B, and infects him/her with the flu. Then person B has dinner with person C and infects person C. The idea behind this type of diffusion is the farther a person or place is from the point of origin the longer it will take for the disease to reach that location, but in the meantime, places that are closer to the origin will become infected sooner. At the urban level an epidemic will spread between neighboring cities through basic trade routes (frequently rivers) and continue to diffuse out in a linear fashion. Hierarchical diffusion differs from contact diffusion in that rather than moving from nearest to farthest location, a disease would first spread between the largest and most important cities, which would be the most connected by transportation routes, and then diffuse to the second-tier cities and on down through the urban hierarchy. The result would be that diffusion to other large cities would happen before reaching a smaller city in between the two large ones.

Andrew Cliff and Matthew Smallman-Raynor lead the field in studying historical disease diffusion, for a variety of different diseases (Cliff et al. 1981, 1986, 1998; Smallman-Raynor and Cliff 1998a, 1998b, 1999, 2001; Trevelyan et al. 2005a, 2005b). The first of these diffusion studies by Cliff et al. (1981) investigated waves of measles

epidemics in Iceland in order to develop a model of outbreaks and diffusion structure for that particular disease. Subsequent investigations include the influenza pandemics (Cliff et al. 1986), typhoid fever (Cliff et al. 1998; Smallman-Raynor and Cliff 1999, 2001), and most recently poliomyelitis. This last study also utilizes a variation of the Getis and Ord statistic described in chapter 2 (Trevelyan et al. 2005a, 2005b).

Yellow fever, smallpox, and enteric fever (another name for typhoid fever) were studied during the Cuban insurrection against Spain, the diffusion processes being compared between both war-time and peace-time (Smallman-Raynor and Cliff 1999). A more recent study by these authors investigates the diffusion of typhoid between United States Army camps during the Spanish-American War (Smallman-Raynor and Cliff 2001). Here a new type of diffusion in disease transmission is proposed, transfer diffusion. Each of these diffusion studies focus on the movement of disease from city to city. It is likely that this level of resolution is the most detailed available for the studies in order to maintain consistency and incorporate a temporal component. Nonetheless, a geographic scale missing from these kinds of studies is the large scale process, the diffusion within a node of the epidemic. The authors do note the importance of investigating at this finer resolution even if their data do not allow for such an analysis (Trevelyan et al. 2005; Smallman-Raynor and Cliff 1999).

Cliff, Smallman-Raynor, and their colleagues are not the only ones to write about the movements of diseases in America and Europe in the past. Wilson (1993) studied the diffusion of smallpox through towns in Finland, while Patterson (1994) focused on cholera diffusion in Russia. Wilson (1993) incorporated two additional elements to his diffusion study that set it apart from much of the rest of the historical medical research

17

discussed to this point. He focused on the interactions between villages and the diffusion of smallpox, and in so doing addressed the issues of geographic scale. In addition, he also approached the question of choosing between morbidity and mortality data. He explained that one of the key problems with using mortality data is that not all cases become deaths. This can be problematic since mortality patterns of certain diseases may not provide a good sample of the actual disease distribution. Unfortunately, as in Wilson's case, sometimes mortality data are all that is available and the researcher must make do. With this in mind, Wilson developed a simulation model of morbidity to help explain why mortality data is not always a good representation of disease diffusion or distribution patterns. While Wilson had access to household level data, he conducted his study at the village level. Wilson concluded that a micro-scale study such as this one should not use mortality data since individual villages might have different mortality rates even from similar morbidity rates. At the same time, he did suggest that mortality data will work as a proxy for morbidity data in more macro-scale studies, as the micro-scale difference in morbidity will be smoothed out through the scale of the study (Wilson 1993).

Patterson (1994)'s analysis of cholera in Russia at the province level focuses on the types of diffusion patterns that occurred during each of the series of nineteenth century epidemics. Patterson found that cholera tended to follow the major transportation routes, including rivers such as the Volga, rather than following a hierarchical diffusion between urban areas as discussed by Pyle (1969). The available data limited Patterson's (1994) study to only consider the equal or unequal cholera distribution across each province.

A diffusion study by Pyle (1969) considered the impact of the cholera epidemics that swept through North American urban areas during the nineteenth century. Here, Pyle focused on changes in the type of diffusion pattern exhibited by cholera as the United States urban hierarchy developed in the nineteenth century. While cities relied upon rivers and oceans as their primary transportation routes for goods and people to and from their shores, diseases tended to diffuse along these pathways following a contagious diffusion pattern. Wilson (1993) also employed this idea in his study. As transportation improved and a hierarchical urban system developed in North America during the nineteenth century, cholera epidemics began to follow a similar hierarchical pattern (Pyle 1969). The disease first appeared in the largest coastal cities and diffused down the urban hierarchy following the railway system rather than simply in a radial pattern from an urban center to its nearest neighbors (Pyle 1969).

A final diffusion study in need of a mention here is Curson's (1985) exploration of multiple epidemics in Sydney, Australia during the nineteenth century. This book provides an exception to the focus on diffusion over large areas by focusing instead on a single city, although still using areal-based analysis. His primary foci are on how social geography impacted the diffusion of epidemics in the city and how reactions to these epidemics brought about social, policy, and geographic change to the city (Curson 1985). By concentrating this investigation upon one city and considering how six different diseases moved in both space and time Curson (1985) was able to identify how much impact epidemics had on sections of the city housing residents of lower socio-economic status. In his conclusion, Curson (1985) describes how disease after disease, and year after year, the same socially marginalized neighborhoods bore the heaviest disease loads

regardless of where or when the disease in question was introduced into the urban system. While by focusing on a single city Curson (1985) takes a different approach to diffusion than many of the other studies mentioned here, the nature of his data, which vary from epidemic to epidemic, force him to consider disease diffusion at a neighborhood and areal scale. Therefore, while this historical medical geography of disease diffusion in Sydney gets closer to investigating highly localized disease patterns, the analysis does not consider the spread of disease and its patterns on an individual basis (Curson 1985).

Medical Geography

Medical geography as a geographic sub-field provides one of the two final pieces of context for this study. This sub-field includes a number of approaches to medical and, more generally, health related topics ranging from the distribution of a particular infectious disease to the accessibility of health care facilities. In terms of methodological contributions, the field of quantitative medical geography relies heavily on the visualization of disease patterns through maps and the statistical analysis of disease distributions to identify places where these disease distributions do not follow expected patterns (Gatrell 2002; Cromley and McLafferty 2002; Meade and Earickson 2000). To achieve these goals, there are a number of statistical techniques available, and the ones relevant to this study will be discussed in chapter 2. A typical approach in quantitative medical geography is the search for hotspots (clusters) of a disease within a dataset as an indication of an above normal (or expected) concentration of the disease. Such an occurrence would thus be an indication of something out of the ordinary happening in that place (Gatrell 2002; Cromley and McLafferty 2002; Meade and Earickson 2000).

Often the identification of a disease hotspot is followed by an analysis of possible factors contributing to the increased rate in that location: focusing on the disease's ecology (Gatrell 2002; Paul 1985; Meade and Earickson 2000). In other words, the search for spatial process or association follows an initial identification of spatial patterning. Some medical geography studies have focused on urban patterns and even a single disease within a city, but in these cases a historical perspective is not usually used (Pyle 1973; Pyle and Rees 1971). Craddock (1998) does deal with specific patterns of tuberculosis in nineteenth century San Francisco, but only as a means of understanding socio-political aspects of public health. Understanding the medical geography of the past can provide insights into the spatial structure of epidemics today, and historical datasets are invaluable for testing new technologies and methodologies that can later be used to answer modern disease questions (Cliff et al. 1981, 1998; Trevelyan et al. 2005a, 2005b).

<u>Public Health, Infrastructure, and Urban Environmental History</u>

Some environmental change supporters within mortality research indicate that the building of infrastructure altered the local environment in such a way that health improved (Preston and van de Walle 1978; Thornton and Olson 2001). Public health histories tend to agree with this environmental hypothesis concerning late nineteenth century health improvement, especially in American cities. The focus of the history of the public health movement, though, is not on mortality patterns or changes, but on the social and political forces that led to what we now take for granted as public health infrastructure and public health departments (Melosi 1994, 2000; Tarr 1979; Duffy 1968, 1974). Studies focused on late nineteenth century public health history and infrastructure history tend to end their discussion with the introduction of scientific and technological

solutions to health problems, such as the building of a sewerage system, operating under the assumption that following a project's completion health improved in the city in question (Melosi 1994; Tarr 1979). Unfortunately, public health histories of a single city tend to paint this story with a rather broad brush and gloss over the fact that infrastructure systems are built over a long period of time and that usually there was uneven service after the project was completed (Colten 2002; Crane 2000). Much of the uneven pattern of service provision resulted from political decisions (Colten 2002; Lessoff 1994; Green 1963).

The first phase of the public health movement was closely tied to miasmatic theory and the pollution problems created by rapid industrialization and urban population growth. As city populations grew, so did the amount of waste generated. In this context, waste not only includes domestic refuse, but also human and animal wastes, most of which were deposited into open gutters, privies, and cesspools until the construction of comprehensive sewerage systems in the late nineteenth century and the establishment of organized municipal garbage collection (Boone 2003; Melosi 1994; Boone and Modarres 2006). The repulsive, noisome collections of wastes in urban centers were identified as the source of disease under the auspices of miasmatic theory (the belief that vapors rising from decaying vegetative matter made a person ill) in the mid-nineteenth century. In reaction to the identification of this source of disease causation cities began to sanitize their environments through improved drainage in the form of sewerage systems that remove both storm water and unsanitary wastes from the city and thus remove the source of miasmas/bacteria residing on the wastes (Melosi 2000; Boone and Modarres 2006).

Cities as they develop, quickly pollute local waterways through the deposition of wastes into water systems, and usually forced urban dwellers to seek alternative water sources such as wells or the development of large engineering projects to get water from distant, unpolluted waterways (Boone and Modarres 2006; Gandy 2002; Gumprecht 1999; Melosi 2000). Wells, whether supplied by a water company drawing water from a river and piping it to the well location or tapping ground water sources, were susceptible to contamination from any number of sources, particularly if they were located near a privy. Without proper attention, poorly constructed privies, in particular, would leak human wastes into the surrounding soils. If those privies were in close proximity to a well the leaking wastes could contaminate drinking water (Kober 1895; Wills 1996). While urban populations in the mid-nineteenth century did not know that bacteria in drinking water could cause illness, stinking, visibly dirty water was recognized as undrinkable and likely a cause of sickness (Wills 1996). By the 1880s, many city governments in the United States recognized, through the development of comprehensive waterworks, that they could control the quality of drinking water with more equality and integrity than either private companies or the inconsistent system of wells, thus helping to maintain their city's health (Melosi 2000). These waterworks projects simply supplied drinking water to cities and did not treat the water in any way until after the bacteriological revolution.

By the 1890s, the bacteriological revolution and the transition from a focus on sanitation to one on laboratory testing was firmly entrenched as the second stage of the public health movement. The results of this shift manifested themselves in a number of ways. Diseases were identified through laboratory testing for bacteria presence, rather

than simply through visual assessments of a patient's symptoms (Hammonds 1993; Rosenau et al 1909). Wells and water distribution systems were tested for disease-causing bacteria (Rosenau et al. 1907). The developments in scientific testing and understanding of disease prevention also led, in this new phase of public health history, to city governments creating programs aimed at eliminating particular diseases from their communities, one example being New York City's diphtheria campaign (Leavitt 1996; Hammonds 1993). Programs of this kind relied upon testing large portions of a population for disease presence, and for those without immunity to the disease, administering a toxoid (inoculation) treatment (Hammonds 1993). As a part of this scientifically-based form of public health, water treatment, through filtration, became another means of actively seeking to prevent disease. Bacteriological testing had identified that simply drawing water from a water source that appeared to be clean did not mean that invisible bacteria were not lurking in and waiting to cause an epidemic (Sedgwick 1893; Wills 1996; Kober 1895; Parker et al. 1907). Water treatment usually began in the form of sand filtration, which involves the forcing of water through a bed of sand that would collect contaminants, such as invisible bits of fecal matter to which typhoid bacteria clung, as the water percolated through the filters and then on through the distribution pipes to consumers (Rosenau et al. 1907).

During this period from the middle of the nineteenth century until the early decades of the twentieth century, numerous investigations of the origins of epidemics were conducted. Like the investigations used in this dissertation, many of these studies mapped the distribution of cases and/or deaths in order to visualize the distribution of a disease in the hopes of identifying a source of the epidemic (Sedgwick 1893; Whipple

1908; Jordan 1909). Two epidemiological reports of particular interest for demonstrating the ways in which disease understanding developed, outline water-borne disease epidemics in the 1890s. The first, by the William Sedgwick (1893), one of the prominent bacteriological scientists of his time, details the epidemiological investigation of a typhoid fever epidemic in Lowell and Lawrence, Massachusetts. In his report to the Massachusetts State Board of Health, Sedgwick described how he collected information for every identified case of typhoid in Lowell, including information about when the person first showed signs of symptoms, where the person primarily got their drinking water from, where they worked, and which company delivered their milk. Using the location of the person's home, each case was mapped as a point (Figure 1.1). The final report included three maps, one of all cases of typhoid fever, one of those cases not associated with canal water, and one of those cases that were possibly associated with drinking canal water. Through investigating the possible source for each case of typhoid in Lowell, Sedgwick was ultimately able to determine that most of the people affected by the disease drank water from the city's waterworks which drew, but did not filter, water from the Merrimac River. This water had been contaminated upstream from Lowell by typhoid infected feces deposited into the waterway (via a privy overhanging a tributary to the Merrimac). Following the epidemic in Lowell and the related epidemic in the downstream city of Lawrence, Lowell ceased to draw its public drinking water from the Merrimac River, and Lawrence mitigated its typhoid problem by building a filtration system (Sedgwick 1893; Whipple 1908).

The second example of this type of "shoe-leather epidemiology" comes from a secondary account of an 1892 cholera epidemic in Hamburg, Germany. In this example,

Figure 1.1 Image of map from the Lowell, Massachusetts typhoid epidemic report, 1892 (Sedgwick 1893).

Jordan (1909) summarizes an investigation of the source of this cholera epidemic. Once again, the investigators of this epidemic, following the model of John Snow's famous cholera investigation, mapped cases of cholera in Hamburg and the neighboring city of

Altona. These neighboring cities were separated not by distance, but by a simple political boundary – a city line. The environmental characteristics of the two cities were the same, both cities drew their water from the Elbe River, but only Altona filtered that water before distributing it for consumption. A cholera epidemic broke out on 20 August, 1892 in Hamburg and as a part of the epidemiological investigation cases of cholera were mapped in both cities. The resulting map displayed a striking pattern: nearly all of the cases of cholera occurred in Hamburg, and the few that did develop in Altona could be traced back to Hamburg. In some instances, cholera struck the side of a street that was

administered by Hamburg and the opposite side of the same street was cholera-free in Altona. Clearly, in these two cities water filtration played a large role in keeping Altona safe from cholera, and this example is important in illustrating the impact that the water filtration systems developed following the bacteriological revolution had on public health (Jordan 1909).

The two examples above provide good illustrations of the impact of urban sanitation on the distribution of health. Every city has its own specific challenges in meeting the sanitary needs of their residents, of some of which are based upon a city's site and some upon the social/cultural/racial geography that developed over time in a place. The following section outlines some of the urban geography elements for the time period of this dissertation in order to place the subsequent analyses in the appropriate context of Washington, D.C. at the turn of the twentieth century.

Washington, D.C. in the Nineteenth and Twentieth Centuries

The city of Washington developed on the shores of the Potomac River and the Eastern Branch of the Potomac River (now called the Anacostia River). Rock Creek and James Creek flowed through the city feeding these two branches of the river that ultimately drains into the Chesapeake Bay and from there into the Atlantic Ocean. This location gave the city a good site for developing commerce and industry in need of shipping routes. Yet, these initial intensions of the city never developed. Instead, the primary raison d'etre for Washington became government, and the city developed around this focus.

Designed to be an impressive capital city of a youthful and progressive nation, the plans for Washington, D.C. drawn by Pierre Charles L'Enfant deviated from a simple

urban grid pattern. L'Enfant's plan included wide sweeping boulevards, streets that cut

diagonally across the regular grid, and in so doing created city blocks of varying sizes

and odd shapes. The design of the city divides Washington into four sections of unequal

size radiating from the Capitol Building, the city's focal point (Figure 1.2). While not

built to the letter of the original plan, the design evolved over the course of the nineteenth

century and remains today the canvass upon which Washington, D.C. continues to evolve

as an urban center (Lessoff 1994).
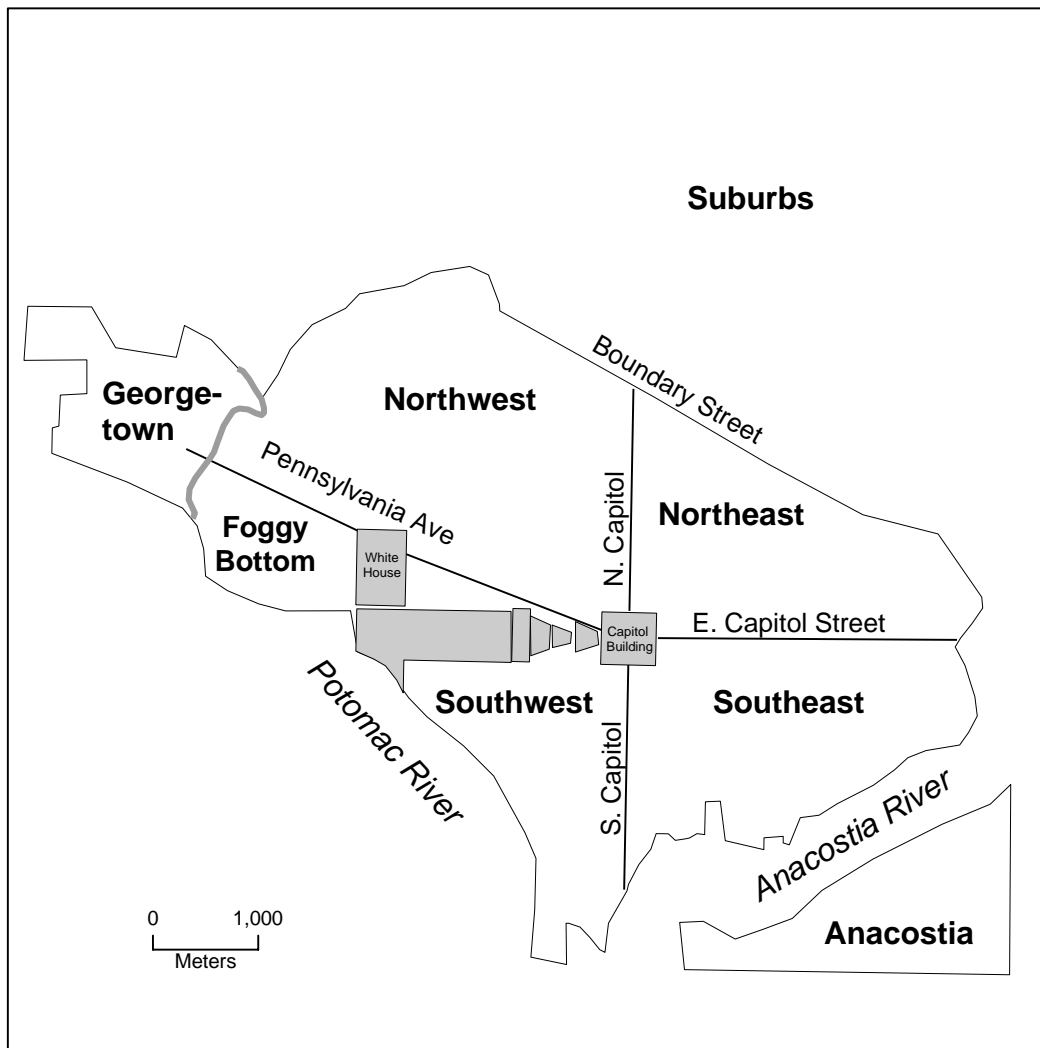
Figure 1.2 Regions of Washington, D.C., as defined in this study. This map will help
to simplify the discussion of the different parts of the city and is referred to in all of
the following chapters.

One element of L'Enfant's design greatly impacted the residential geography of the city. By designing large city blocks often including space for backyards, people needed to access the interior of those blocks and their backyards via alleyways. Over time the interior spaces on many blocks developed into residential spaces of their own. Owners of homes that faced the street would build a small structure on the back part of their lot and either rent out the structure or sell the portion of their land facing the alleyway. Entire communities developed in the homes along these interior alleys (Borchert 1980). For a full development history of Washington's alleys, their population, and culture see James Borchert's *Alley Life in Washington* (1980), the following description section describes this feature of Washington's urban geography to place the analysis results from this dissertation in context.

Alleys developed in many nineteenth century North American cities. The combination of narrow alleyways and wider main and side streets created a hierarchy of street widths and a social segregation mechanism that organized city residences by class, race, and ethnicity (Olson 1997; Borchert 1980). The alleys were typically narrow mazes of streets allowing access to the interior of blocks and acted as residential spaces for the poor. Houses along these alleys tended to be small, overcrowded, poorly ventilated, and poorly constructed. Inadequate drainage along improperly graded alleys, especially those in low-lying areas, allowed water to collect along them, resulting in a micro-environment of disease risk in both the summer and the winter. Despite this, these spaces were highly sought after by the newly arriving immigrants and blacks who could afford no other housing, needed to live near jobs available in the industrial districts, and/or through racial discrimination were relegated to these less desirable living conditions (Borchert 1980;

Groves 1974, 1973-74; Olson 1997). In Washington, with only a small immigrant population, alley housing served primarily the residential needs of blacks and some poor whites. The design of Washington, D.C., with its extensive network of alleys on blocks throughout the city, meant that alley residents lived in most parts of the city regardless of the social class of the street fronts. Some of the city's poorest resided behind the homes of the wealthy (Borchert 1980), creating a complex spatio-social geography for the city.

Washington's population grew rapidly during the last quarter of the nineteenth century and into the early twentieth century (Table 1.1). Between 1880 and 1910 the city's population nearly doubled in size. This rapid growth placed stress upon the existing housing stock, despite an ongoing construction boom. For the disadvantaged portions of the population, poorly maintained and crowded houses and tenements along alleys were often the only viable residential choices (Groves 1974). As the population grew and new homes were built, a socially-based residential geography developed. Broad descriptions of this late nineteenth century residential geography are outlined below.

Table 1.1
Population of Washington, D.C.

Year	Population
1880	177,624
1890	230,302
1900	278,718
1910	331,069

A construction boom began in the 1870s and continued until about 1900. During this time Northwest Washington, particularly around Dupont Circle, became quite fashionable leading to further entrenchment of affluent residences in this section of the city (Figure 1.3) (Myers 1973-74; Green 1963). While many reasons contribute to the

30
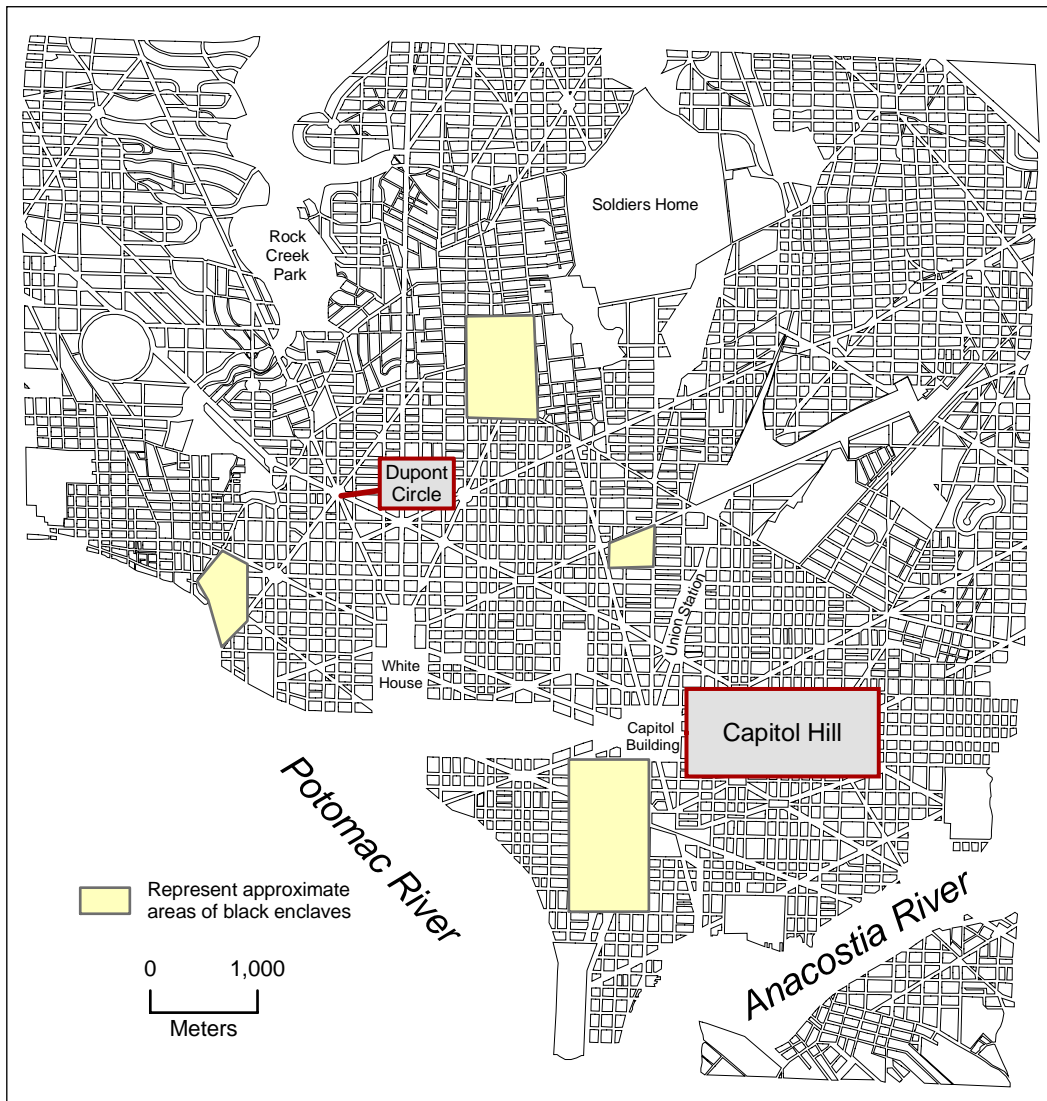
Figure 1.3 Neighborhoods and locations of black residential enclaves

explanation of why the affluent members of Washington's population chose to live in this part of the city, some universal tendencies among the wealthy in nineteenth century cities probably also hold true here. Olson (1997) described the reasons explaining why the wealthy in Baltimore shifted their homes from near the waterfront to new neighborhoods to the north and west of the central business district. The reasons provided by Olson were two-fold, first by building new houses the residents were able to incorporate modern

conveniences such as water closets, gas lighting, and hot water systems into their homes.

Yet more importantly in Baltimore, these new homes were positioned on higher ground

where the residents could take advantage of natural drainage in an effort to escape the

damp, filth, and standing water of the lower-lying areas of the city (Olson 1997).

Although these drainage-based decisions were made before miasmatic theory fully gave

way to germ theory, the precedent of those who could afford to do so living in better

drained areas was well established (Olson 1997). While this example comes from

Baltimore, the Dupont Circle area and much of Washington's Northwest quadrant, in

general, was some of the highest ground in the city and occupied some of the best drained

land (Figure 1.3).

Like the elite neighborhoods of Northwest, Capitol Hill, as the name implies, is

on a hill within the city and therefore on higher ground. Farther to the east of the Capitol

Hill neighborhood was the Eastern Branch of the Potomac River and the swampy, poorly

drained land along its banks that were perceived as a breeding ground of disease. This

was until public works project belatedly brought drainage improvements to that area in

the late 1890s. Myers (1973-74) indicates that public works projects in the late nineteenth

century tended to favor Northwest and ignored the eastern part of the city. Washington's

middle classes, which were mostly composed of government clerks, eventually settled

into the area east of the Capitol Building between C Streets, Northeast and Southeast

referred to as Capitol Hill. This neighborhood extends to the north and south of C Street

into both the Northeast and Southeast quadrants, though the actual boundaries are ill

defined (Figure 1.3) (Myers 1973-74). By the 1880s, Capitol Hill was one of the few

parts of the city that had not been "claimed" and developed by a particular social class.

This essentially became home to much of the city's middle class for a number of decades. Proximity to government offices and other city conveniences made this neighborhood attractive to families employed as civil servants (Myers 1973-74).

The working poor, regardless of race, mostly lived in the Southwest and Southeast quadrants of the city (Figure 1.2) (Myers 1973-74; Groves 1973-74). However, the proliferation of alleys and alley housing in the city distributed some of these marginalized populations throughout Washington, particularly the poor black population (Groves 1974, 1973-74; Borchert 1980). The large concentration of working class Washington residents in Southwest was divided between a predominantly black enclave between 4 ½ Street and South Capitol Street, and a predominately poor white population between 4 ½ Street and the Potomac River (Figure 1.3) (Groves 1973-74). One distinctive feature of Washington's social fabric was the black population and its residential and occupational patterns. By 1830, free blacks outnumbered black slaves in Washington, unlike other Southern cities where slaves continued to be in the majority until as late as the 1860s. It was not until the "Great Migration" of the 1920s that Northern cities became a major destination for migrating blacks, but during the nineteenth century, both before and after the Civil War, Washington was an important destination for blacks migrating north. As a destination city, Washington absorbed a growing, if impoverished, black population into the housing stock and labor force during the nineteenth century (Groves and Muller 1975). Residential patterns of blacks in Washington reflected attitudes of segregation relegating blacks mostly to the alleys and some side streets. The exceptions to this trend were the black enclaves, particularly the one in Southwest (Figure 1.3), where some successful black families were elevated to a middle or even wealthy status. As a result,

33

their residences shifted to street front housing though still remaining within the black residential enclave (Groves and Muller 1975).

Three clusters of black residents were identified by Groves (1973-74) as having developed by 1880 in Northwest. These were centered on 19th and R Streets, Connecticut Avenue and L Street, and New York Avenue and M Street, and together housed approximately 10 percent of the city's black population. Another 10 percent of the black population resided in Southwest in that year. The enclaves in Southwest and at New York Avenue and M Street remained stable into the twentieth century, while other concentrations of black homes shifted between 1880 and 1920. By 1920, two new and fairly stable enclaves had appeared. One of these new enclaves was located in the very north of the city straddling Boundary Street (Figure 1.2, 1.3). The other was in the section of Northwest known as Foggy Bottom proximate to Rock Creek (Figure 1.3) (Groves and Muller 1975).

These broad and over generalized descriptions of the residential geography, for the purposes of this study, provide an overview of where different social classes resided within Washington, D.C. However, using Northwest as an example, the reader should be mindful that while much of this area was home to the city's elite, it was not a racially or socially homogeneous quadrant. Therefore, caution must be employed in making sweeping judgments based upon the location of clusters of poor health, as these will oversimplify what was a socially complex milieu. Since health is tied to socio-economic status, residential location, and urban environmental conditions of a city, the final section of this chapter describes the status of sanitary infrastructure in Washington, D.C up through the early years of the twentieth century.

Outline of Washington's Infrastructure Status until 1905

Washington, D.C.'s site is tied closely to aspects of public works developments and health in the city. Until public works projects focused upon draining and improving sanitation in the city in the 1870s began, Washington, D.C. was little more than the seat of the federal government and a place where few with other residential choices lived permanently (Myers 1973-74; Green 1963; Lessoff 1994). The improvements to the city in the 1870s included paving streets, adding street lighting, and developing a sewerage system (Myers 1973-74; Green 1963). Once these urban features were in place, people began to *choose* to live permanently in the Capitol City (Green 1963).

Addressing the drainage needs of Washington, D.C. was a continuous concern in the city from its founding until the completion of the sewerage system in the early twentieth century. As a low-lying site along the shores of two rivers, Washington flooded easily, particularly after development in the city began to alter the local environment, this included street paving and covering over existing waterways, which in the process inhibited storm water and street gutter drainage. Other challenges faced by the District of Columbia were the tidal marshes, where the National Mall now sits, and a tendency for sewage to be washed back up the Rock Creek and James Creek by the tides. These streams were originally intended to remove waste from the city. Sewage problems were further intensified as both Rock Creek, and to a lesser extent James Creek, often did not have enough water flowing through the channel to transport the sewage deposited into them from the city's early sewerage system (Green 1963).

The Board of Public Works in the early 1870s began building a comprehensive sewerage system of the city, but due to financial problems (over 5,000,000 dollars being

spent on the system), and poor city management, the project was not completed (Schultz and McShane 1978; Green 1963; Lessoff 1994). The governance of Washington, D.C. was brought under the full supervision of Congress in 1878 under the Organic Act. While this shift made the city financially solvent again, there were limited resources available to invest in repairing the faulty portions of the partially constructed sewer system and to continue its expansion to the remainder of the city (Green 1963). Further problems with the sewerage system, beyond emptying into waterways with too little water flow to remove wastes from the city, were that some lateral sewers were laid in such a way that water would have had to flow uphill (Schultz and McShane 1978; Green 1963). The sewerage problem could no longer be ignored by the early 1890s, and despite financial constraints, Congress approved a plan developed by the country's top sanitary engineer, Rudolph Henning, and a special board of sanitary engineers for the city, to address the sewerage and sanitation problems within Washington. This plan, drawn up in 1890, for a combined storm and sanitary sewerage system had been designed with an expectation of urban expansion and once completed served the sewerage needs of Washington until the late 1950s, when sanitary experts finally pushed for a separate storm and sanitary sewer system. By 1901, the sewer project was about half completed, and was completed fully by 1907 (Green 1963).

Washington's residential drinking water was drawn from the Potomac River 15 miles upstream at Great Falls (Parker et al. 1907). The city developed a piped water supply with the building of the Washington Aqueduct in 1853, however the city was not efficiently served until the early twentieth century. Throughout the later years of the nineteenth century many Washington residents continued to draw their water from deep

and shallow wells located throughout the city (Kober 1895). Green (1963, 42) identified key problems with the city's water distribution system as "the insufficient head of water at the reservoir above Georgetown, leakages in the distributing system, and the inadequacy of a single three-foot distributing main to meet the needs of the rapidly growing population." By 1896, the city was using two reservoirs, Georgetown and Daleclaria, to try and meet the needs of the city in terms of a clean drinking water supply. This addition may have been in response to an outbreak of typhoid fever in 1895 linked to contaminated well water (chapter 3). The city added the Washington Reservoir, later known as the McMillan Reservoir, in 1902, in yet another attempt to significantly improve the city's water supply and reduce disease (Green 1963; Lumsden and Anderson 1911).

However, in the 1890s, Washington continued to have a typhoid rate that was four times higher than in comparable European cities, despite the sanitary measures in place (Lessoff 1994; Whipple 1908). The use of three reservoirs as settling basins for Washington's drinking water supply did not completely eliminate typhoid and other water-borne diseases from the urban system and this prompted the city's officials to push for the building of a sand filtration plant (Rosenau et al.1907; Green 1963). Water filtration was already a proven means of removing bacteria from drinking water in European cities (Jordan 1909). The German government mandated and enforced water filtration for its cities, resulting in a typhoid mortality rate a quarter the size of the typhoid fever mortality rate in the United States (Whipple 1908).

The sand filtration plant in Washington, D.C., located at the site of the McMillan Reservoir, opened in October 1905 (Rosenau et al. 1907). Then in June 1906,

Washington's Health Officer noted an increase in typhoid fever cases across the city. Instead of experiencing the expected decrease in typhoid fever rates in 1906 due to the incorporation of the presumed key factor in supplying clean water to consumers, filtration, there was an increase in typhoid morbidity. The majority of the analyses conducted in this dissertation focus on the causation and patterning of disease experienced in Washington, D.C. as recorded in the reports of the epidemiological investigations into the unexpected 1906 increase in typhoid fever morbidity.

CHAPTER 2
DATA AND METHODS:
CREATING A GIS FOR SPATIAL AND TEMPORAL ANALYSIS OF THE PAST

Introduction

Many of the materials and much of the methodology used in this dissertation

applies to all of the analyses performed in the following five chapters. Therefore, rather

then repeat the same information, the materials and methods shared by the different

chapters will be explained once in this chapter. Each analysis chapter asks a different

question or questions about the data being used. These questions at times led to the

addition of other methodologies appropriate to that particular chapter, and therefore are

described in that context. Below are detailed descriptions of the available data utilized in

this study and the form of the local spatial auto-correlation approach used to explore

these data.

Available Data

Detailed investigations of typhoid epidemics occurred in Washington, D.C. in

1895, 1906, 1907, 1908, and 1909. Doctor George Kober conducted the first of these

investigations as a special medical sanitary inspector at the request of Washington's

Health Officer for an outbreak of summer typhoid between July and October 1895

(Kober 1895). The United States Public Health and Marine-Hospital Service (PHS)

conducted the four typhoid investigations between 1906 and 1909 (inclusive), again at

the request of the city's Health Officer (Rosenau, Lumsden, and Kastle 1907, 1908, 1909;

Lumsden and Anderson 1911). All five reports included tables, charts, text, and maps to

explain the likely source(s) of typhoid fever in each year. The investigations met with

mixed results in identifying a particular cause of typhoid fever in Washington, D.C. The

spatial data in these reports can now be re-examined to either confirm the geographic findings of the original authors, or identify patterns previously not considered.

In December 1895, the Commissioners of the District of Columbia published their annual reports to the House of Representatives for the year ending 30 June, 1895. The Health Officer's report contains a summary of a special investigation of a typhoid fever outbreak during the summer of that year. This special report includes not only a written summary of the outbreak and descriptions of the origins of some cases, but also a map of typhoid cases and deaths, and a map of the locations of public wells that tested clean or contaminated. Every year the Health Officer's report included maps of deaths at the residence level from specific causes including typhoid, but the 1895 special typhoid report is unique to that one year. The typhoid morbidity map and the special report it came from provide a rare glimpse of typhoid morbidity for a city that typically at the time only mapped disease specific mortality data (Health Officer 1898, 1881, 1895). Non-fatal and fatal cases of local and non-local origin were mapped using two different symbols and two colors to identify the residential location of each case that developed between 1 July and 31 October, 1895. This means that a red cross represents non-fatal cases of non-local origin, a blue cross represents a fatal case of non-local origin, a blue dot represents a fatal case of local origin, and a red dot represents a case of local origin.

The four reports published by the United States Public Health and Marine-Hospital Service between 1907 and 1911 contain the remainder of the spatially recorded typhoid data used in this dissertation (Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911). Typhoid cases were mapped as points at the block level in the reports at either two-week or monthly intervals (Figure 2.1). The reports investigating typhoid in
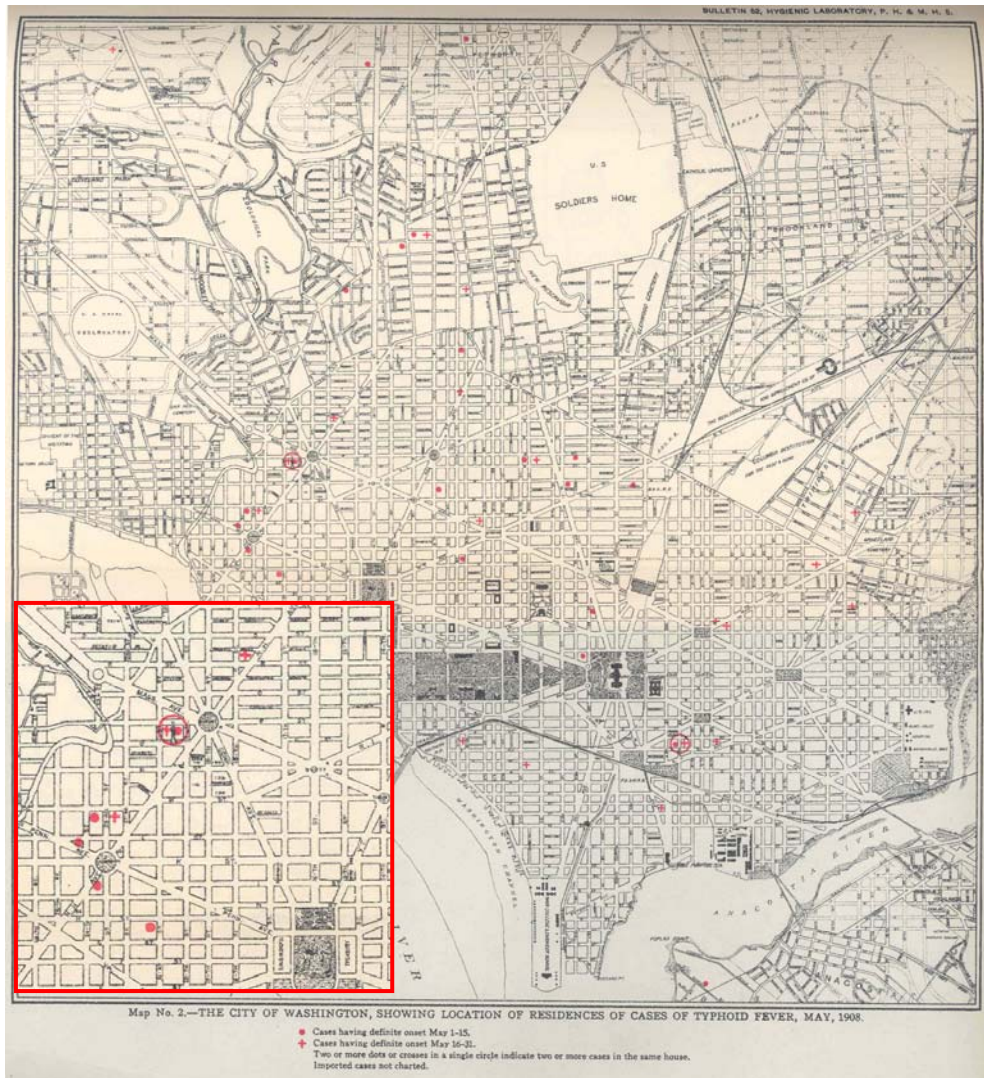
Figure 2.1 Image of a 1908 typhoid morbidity map, includes a zoomed in section of the map to better illustrate how typhoid cases were marked.

1906, 1907, and 1908 only studied typhoid cases for the six summer months. This resulted in six maps, one for each month between May and October (inclusive). On each map the PHS noted all cases for the first half of the month in one symbol and the second half of the month in a different symbol. All cases that showed signs of onset between 1 May and 15 May were identified with a dot over their location and all cases between 16 May and 31 May were notated with a cross. The 1909 report included typhoid case

locations for the entire year beginning in January and ending in December. Unlike the previous reports, the 1909 report displayed all cases occurring in a single month in one symbol and not by two-week intervals. This means that all of the cases that developed in January 1909 were recorded as a dot on the January map.

Each typhoid case was heads-up digitized over an 1898 map of Washington, D.C. georeferenced in ArcGIS 9.0 (Appendix A).[1] Since the original reports mapped the typhoid cases at different temporal scales, the data were digitized at the finest temporal resolution, allowing for the possibility of subsequent temporal aggregation by merging shapefiles. The digitizing process resulted in a total of 49 shapefiles of typhoid cases.[2]

Global and Local Spatial Auto-correlation

One of the key elements of geography is the search for spatial patterns. The previous chapter mentioned that medical geographers often examine spatial disease data for hotspots or other patterns that can provide clues to underlying causative factors in the social or natural environment. Spatial auto-correlation provides one such measure of spatial cluster identification within data (Gatrell 2002). The basic premise of spatial auto-correlation is that phenomena in close proximity are more likely to be similar than phenomena that are farther apart (Goodchild 1986; Cromley and McLafferty 2002). This is an appropriate concept when considering contagious disease, not only in terms of disease spread within an outbreak (the neighbor of person A is more likely to display symptoms than person B located in a separate neighborhood), but also across similarly

---

[1] See Hinman (2002) for a detailed explanation of both the georeferencing and digitizing processes for historical data in a GIS setting.

[2] The total number of 49 shapefiles is reached from: 1 shapefile for the entire 1895 epidemic. Then for 1906 through 1908 two shapefiles were created for each of the six months investigated. One file was for the 1st through the 15th of a month and another for the 16th through the 30th or 31st. Finally, for 1909, one shapefile was created for each individual month of the calendar year.

diseased environments. For example, an outbreak of influenza in Foggy Bottom

(Washington, D.C., figure 1.2) is more likely to be related in terms of its origin to an

influenza outbreak in Northwest than to an outbreak in Anacostia. Therefore, if all of

Washington, D.C. were tested for spatial auto-correlation, Foggy Bottom and Northwest

could be considered a "cluster" or having positive spatial auto-correlation, while

Anacostia would not be a part of that particular cluster. Spatial auto-correlation is heavily

dependent upon the scale of study and two county sized areas might constitute a cluster

of disease incidence, but when those counties are broken into smaller geographic

aggregations, further hotspots and even coldspots (where no disease is present) would

appear (Goodchild 1986).

The issue of scale in hotspot detection is one of the key aspects of this

dissertation. The original PHS reports essentially claimed a lack of global spatial auto-

correlation among the mapped typhoid cases, but if the scale of the analysis were

changed, would clusters of typhoid be visible? Global spatial auto-correlation is the

search for hotspots across an entire dataset, while local spatial auto-correlation searches

for clusters of events within a smaller area and compares the patterns found to other small

areas, or to a global mean, to determine which localities can be considered as statistically

significant clusters (Cromley and McLafferty 2002). In a GIS context, Getis and Ord

(1996) identify one of the key problems with measures of global spatial auto-correlation

and large datasets.

> The global statistics summarize an enormous number of possible
> disparate spatial relationships for a given set of data. In addition,
> since each datum represents a very small portion of the study area,
> the likelihood that near-neighbour data point are similar is so high
> that any global statistic measure at a large scale of analysis
> provides little useful information (Getis and Ord 1996, 262).

This quotation supports Goodchild's (1986) obervation about the scale of analysis changing spatial auto-correlation patterns that may or may not be present within a dataset. At the same time, Getis and Ord (1996) present a strong justification for the use of local measures of spatial auto-correlation in a GIS setting when presented with large datasets, such as the typhoid data used in this dissertation.

Measures of both global and local spatial auto-correlation are used in this dissertation. The former is used to confirm or deny the original PHS findings of a general distribution of typhoid, and in the event of a lack of global clustering, as a justification for the interest in exploring the data with local spatial auto-correlation techniques. Based upon the conclusions of the Washington, D.C. Health Officer and the PHS it seemed to be a logical to test whether or not typhoid clustered locally.

The two statistical tests selected for the analysis in this dissertation are Ripley's K-function as a test of global spatial auto-correlation and the Getis and Ord (1992; Ord and Getis 1995), G_i^*, as a test of local spatial autocorrelation.

Ripley's K-function

The Ripley's K-function was used to measure global spatial auto-correlation with the original point data for each year (summer 1906, 1907, 1908, 1909; whole year 1909). The Ripley's K-function was conducted using ClusterSeer2 (Terraseer, Inc, Crystal Lake, Illinois). The K-function is written as (following Durbeck et al. 2000):

$$\hat{K}(h) = \frac{R}{n^2} \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} \frac{I_h(d_{ij})}{w_{ij}}$$

Where R is the area of the region of interest (in this case the defined study boundary-metropolitan Washington, D.C. and the surrounding "suburban" areas; Figure 1.2), n is the total number of typhoid cases within the study area (R), d_{ij} is the distance between

cases *i* and *j*, and $I_h(d_{ij})$ is an indicator function that equals 1 if d_{ij} is less than *h*, and equal

to 0 otherwise. Clusterseer employs w_{ij} as an edge correction factor (defined from 0.5 to

1) that ensures that cases near the study boundary are evaluated equally (Durbeck et al.

2000). Clusterseer employs a second formula to evaluate the K-function in comparison to

a homogeneous Poisson distribution, described as L(h). That formula is expressed as

(Durbeck et al.2000):

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}}$$

Ten distance steps from 100m to 1000m were used[3] and 1000 Monte Carlo simulations to

evaluate L(h) in comparison to Complete Spatial Randomness.

G_i* Statistic

In order to test for statistically significant local typhoid clusters within

Washington, D.C., and to determine the spatial extent of these clusters, the Getis-Ord G_i*

statistic was used (Getis and Ord 1992; Ord and Getis 1995). The G_i* statistic is useful

for identifying individual members of local clusters by determining the spatial

dependence and relative magnitude between an observation and neighboring observations

(Getis et al. 2003). This particular statistical technique was chosen as an approach to

study typhoid fever patterns in Washington, D.C. because it a common and established

technique in epidemiological studies. Examples of its use range from infant mortality to

vector-borne diseases like dengue fever. One dengue fever study uses data collected from

---

[3] Ripley's K-function uses circles of increasing size, referred to as distance steps, as a part of the test of
spatial randomness and as a means of potentially identifying the spatial extent of global clusters. The
process works in the following way: if a circle drawn around a point is one tenth the size of the entire study
area, then under complete spatial randomness approximately one tenth of the points would fall within that
circle on average. This measure is repeated for every point under study. If the points are in fact clustered
then on average significantly more points than a tenth will occur within a circle one tenth the size of the
study area and if the points are widely dispersed then fewer than one tenth would appear within the radius
of that circle (Levine 2004).

individual houses, a scale even more detailed than the one used in the current study,

while Trevelyan et al. employ a variant of this statistical in their historical study of

poliomyelitis (Getis et al. 2003; Wu et al. 2004; Trevelyan et al. 2005). The G_i* statistic

is written as (following Getis and Ord 1992; Wu et al. 2004):

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

Where x is equal to the number of typhoid cases within a given grid cell, S is the standard

variance of typhoid cases, when the distance from grid cell j to grid cell i is within

distance d. A weights matrix is derived where $w_{ij}(d) = 1$; otherwise $w_{ij}(d) = 0$ to

determine whether cases at j are within distance d of case i (Wu et al. 2004). The G_i*

statistic includes the value at i in the calculation of G_i*. G_i* is calculated and then output

as the standard normal variant with an associated probability from the z-score distribution

(Wu et al. 2004).

Grid Development and Methods

The G_i* is a group-level statistic, where point data must first be aggregated to

areas. Two grids were developed as the areas for aggregation for this project. The G_i*

results from both grids were compared to determine the better grid choice for the

remaining analysis. First, a 115m x 90m vector grid surface was developed to aggregate

the spatial distribution of the point data set using the National Park Service Grid Tool

extension for ArcGIS 9.0 (http://www.nature.nps.gov/im/units/mwr/gis/grid_tools.htm).

As the point data available for this study were digitized from block-level maps

contemporary to the original outbreaks, the city block was the finest resolution available

for spatial statistics. The city block shape in Washington, D.C. was, and still is, complicated and not comprised of only simple geometric squares or rectangles. Several blocks were triangular or irregular in shape. The NPS grid tool only develops symmetrical grid cells. To develop a grid surface that best represented this irregular network of block shapes, a systematic measurement survey was performed within the GIS to calculate the average block size. First, the city was divided into its four primary sections (Northwest, Northeast, Southwest, Southeast) to insure that all parts of the city were included in the grid cell size determination (Figures 1.2 and 2.2). Measurements of length and width were collected for 25 square or rectangular blocks in each of the four parts that were the most typical for that section of the city. The mean length and width were calculated for the 100 measured blocks and used as the grid cell size.

The second grid consisted of the digitizing of actual city blocks. Each block was heads up digitized from the georeferenced image of a typhoid report map from 1908 (Figure 2.3). By creating a grid of the actual city blocks the intention was that any resulting hotspot surfaces would better reflect the geography of the city. Additionally, a more geographically accurate grid could facilitate future research by identifying the specific blocks in need of more investigation.

The decision to compare an arbitrary grid instead of the block-based grid surface occurred for two reasons. First, an arbitrary grid can be created more quickly than the digitized grid, thus, if it effectively produces results, in other test environments time would be saved on unnecessary digitizing. Second, the uniformity of the grid cells in the arbitrary grid means that in the analysis there will be a more consistent potential for equal
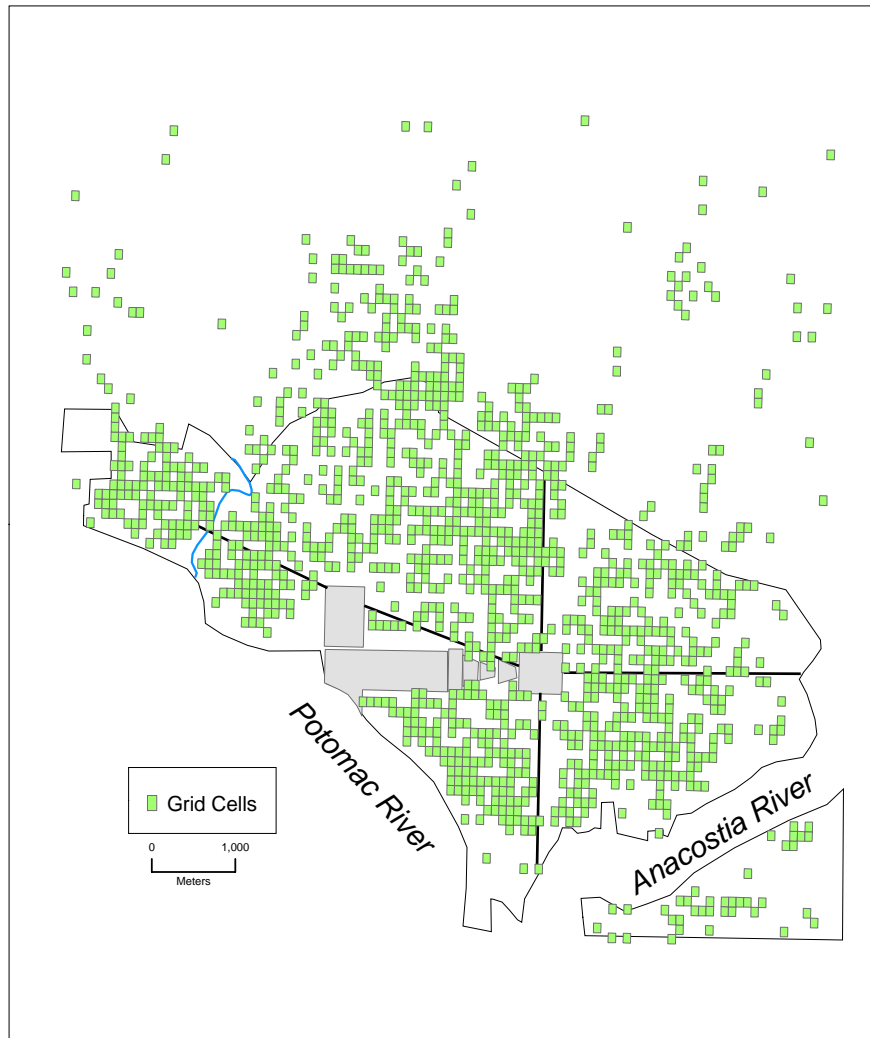
Figure 2.2 Typhoid positive cells in the uniform grid

numbers of cells to be included in the G_i* equation, whereas many of the larger city

blocks, even at 1000m may be the only polygon processed in the equation.

The number of cases occurring in each grid cell, regardless of which grid was

used, were summated using the Count Points in Polygon Tool available in the Hawth's

Anaylsis Tools extension for ArcGIS 9.0 (www.spatialecology.com). G_i* was calculated

using the spatial statistics tool in ArcGIS 9.0 Arc Tool Box.
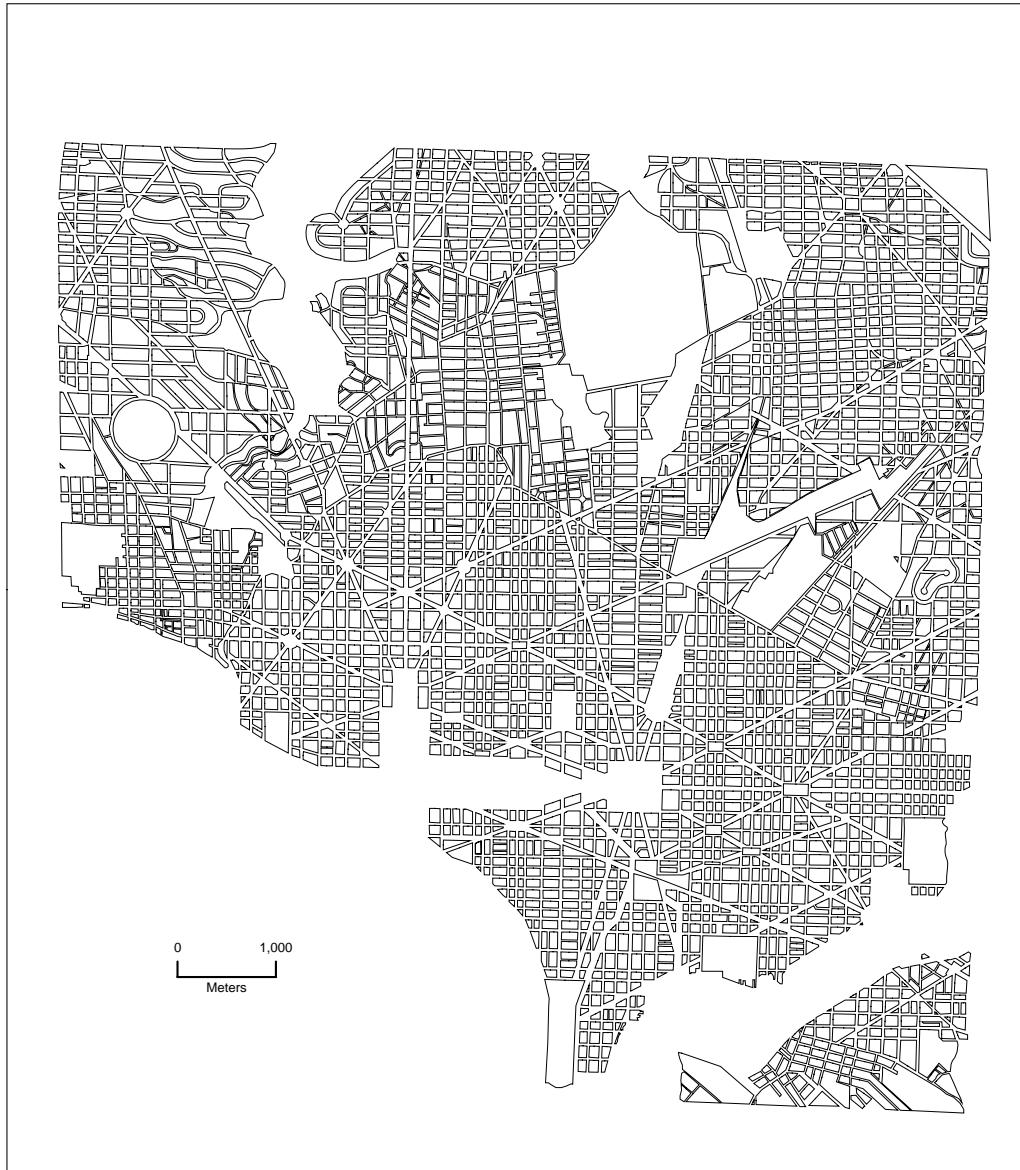
Figure 2.3 Digitized city blocks for G_i* analysis

To determine at what scale typhoid clusters appeared during any of the typhoid outbreaks, multiple distance values are used in this study. The distances *(d)* were set to 150, 250, 500, 750 and 1000 meters. The smallest distance, 150m, was selected to capture localized infections, such as residences clustered around a single shared water source. The largest distance, 1000m, was selected to capture larger outbreaks more representative of a global infection source, such as the city-wide water supply. As the G_i* values are

normal variants of the z-distribution, only those G_i^* values greater than 2.0 were considered significant, in order to be more conservative than $\alpha = 0.05$. This choice results in a 4.6 percent chance of a Type I Error. In other words, there is slightly less than a five percent possibility that statistically significant clusters occurred by chance alone and a 95.4 percent possibility that significant clusters are non-random occurrences. Following Getis et al. (2003) the highest G_i^* value for every grid cell was considered the peak of the typhoid cluster. In this way, although a cluster at 500m may have a G_i^* value exceeding 2.0 at the 500m distance, if the G_i^* value at 150m exceeded that G_i^* at 500m, the 500m distance was not counted as being significant. In other words, for a grid cell to remain a member of a statistically significant cluster from one distance to another, the G_i^* value must increase from test distance size to test distance size. If the G_i^* value did not increase with distances, though values may have been greater than 2.0, they were not considered members of clusters. This is defined as the critical distance, d_c in Getis and Aldstadt (2004). All clusters presented in this study are defined at d_c.

Maps of the significant clusters were produced for each year or part of a year tested. As is shown in Figures 2.4 and 2.5 the G_i^* results from both grid surfaces resemble one another. Still, the block-based grid displays more geographically specific results and is easier for a wide variety of readers to interpret. Therefore, the block-based grid was selected for the G_i^* analysis in this project. In other types of G_i^* studies, an "arbitrarily" created grid should still provide accurate results.

**Arbitrary Grid Results for 1908
Using the Gi* Statistic**

Cluster Size (m)

- 150
- 250
- 500
- 750
- 1000

0   1,000

Meters
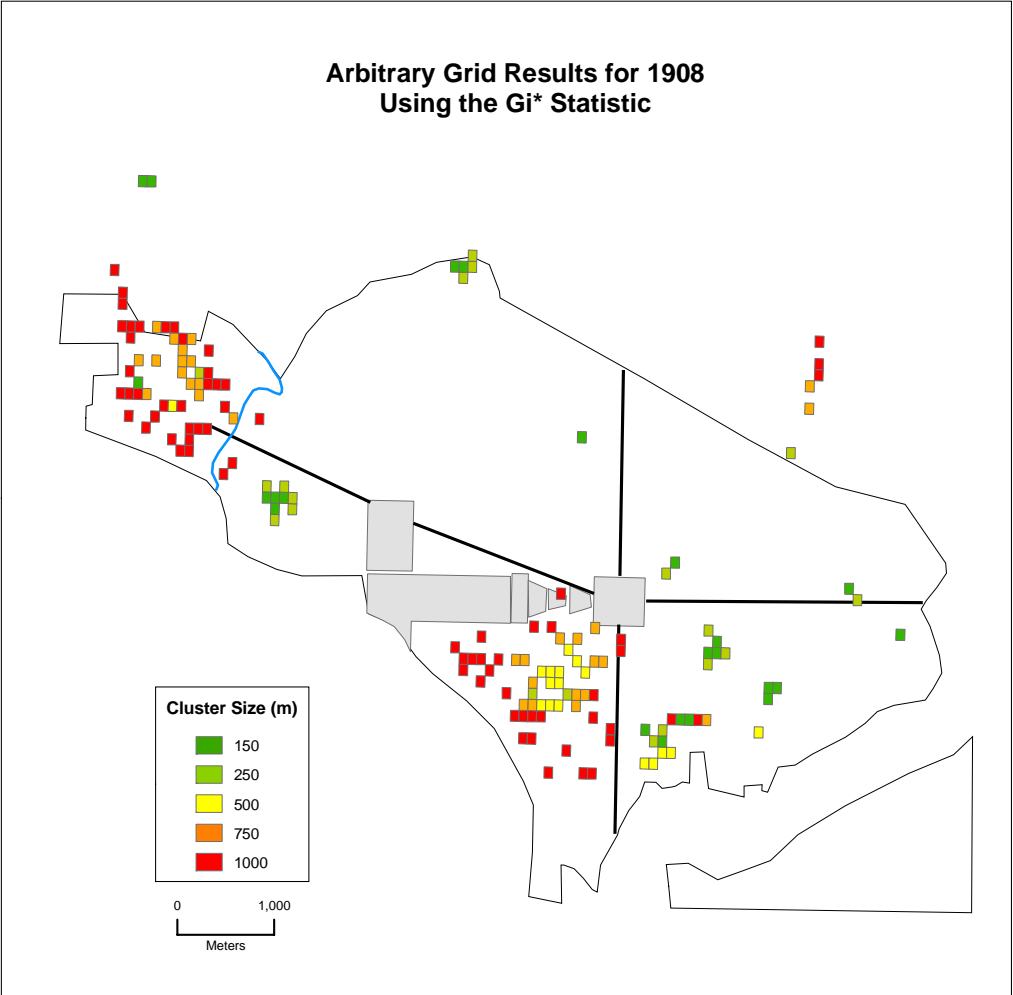
Figure 2.4 G_i* results for 1908 using the uniform grid

**City Block Grid Results for 1908
Using the Gi* Statistic**

Cluster Size (m)
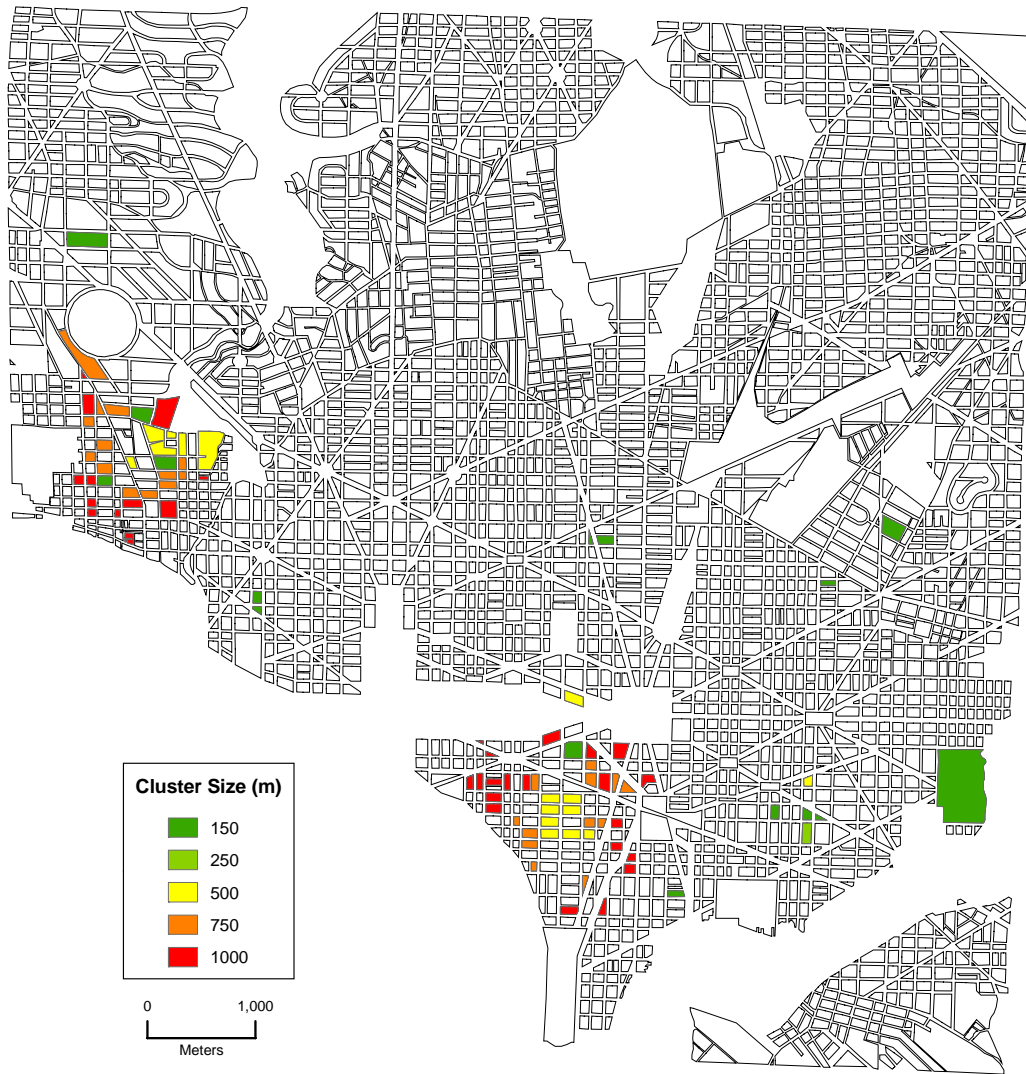
150
250
500
750
1000

0   1,000

Meters

Figure 2.5 G_i^* results for 1908 using the block-based grid

CHAPTER 3
COMPARING MORBIDITY AND MORTALITY SURFACES:
TYPHOID IN 1895

Introduction

In part due to data availability and in part data accuracy, little historical health

research has been conducted using morbidity data from the nineteenth century (Elman

and Myers 1997, 1999; Wilson 1993). The majority of late nineteenth century health

studies use mortality data (Mercier and Boone 2002; Thornton and Olson 1991, 1997,

2001; Woods et al. 1988, 1989; Condran and Crimmins-Gardner 1978, 1980; Smallman-

Raynor and Cliff 1999; Cliff et al. 1998; Pyle 1969; Haines 1995; Preston and Haines

1991; Williams 1992). The general assumption in the use of mortality data is that it is a

reasonable proxy for morbidity data when these are not available (Alter and Riley 1989).

Some studies do use morbidity data, but at generalized scales (Elman and Myers 1997,

1999; Wilson 1993). Not only are morbidity data difficult to come by for use in historical

studies, but until approximately the 1890s these morbidity data were often inaccurate and

under representative of the disease in question. The rise of bacteriological testing in the

1890s, more systematic approaches to public health surveillance, and mandatory

reporting of certain diseases began to change the amount and quality of health data

collected (Hammonds 1993).

As medical care improved in the nineteenth and early twentieth centuries a shift

occurred in the ratio between morbidity and mortality events (Alter and Riley 1989).

Prior to bacteriological testing and an understanding of germ theory medical diagnoses

were derived from a doctor's observations and interpretation of the patient's symptoms,

such as fever, sore throat, and the manifestation of a particular rash. As many diseases

present with similar early symptoms misdiagnosis is only to be expected. For example, early symptoms of diphtheria and scarlet fever include a fever and a sore throat, with only some cases of diphtheria presenting the tell-tale membrane across the patient's throat that traditionally characterizes it (Hammonds 1993; Brainthwaite et al. 1996). Additionally, the simple understanding of the importance of bedside cleanliness and household sanitation meant that mortalities became increasingly skewed to the more vulnerable populations, such as the very young and elderly. In relying upon mortality data, resulting spatial patterns might reflect the distribution of the susceptible cohort rather than the underlying disease surface. Therefore, morbidity data, even given the limitations of symptom similarity, can provide a more holistic spatial impression of an epidemic when the data are available (Elman and Myers 1999; Wilson 1993).

In order to better understand the spatial differences between morbidity and mortality data in the late nineteenth century, the two types of data must be analyzed together and compared. Elman and Myers (1997, 1999), like many others, use data for the whole country and/or entire cities, and therefore tend to consider numerical trends in morbidity and morbidity reporting. Alter and Riley (1989) used days of missed work as a proxy for aspatial non-disease specific morbidity. By considering the geography of both morbidity and mortality for a single disease it is possible to visualize whether or not mortality data accurately represent the overall spatial distribution of cases. The special investigation contained within the Health Officer's report for the year ending 30 June, 1895 provides a rare combination of both morbidity and mortality data in map form. This enables the analysis of the two types of data together in order to determine if typhoid morbidity data from the 1890s might prove more useful than mortality data.

The 1895 Typhoid Epidemic

At the request of Doctor William Woodward, the Health Officer in Washington,

D.C., Doctor George Kober began a special investigation of the large number of typhoid

fever related deaths beginning in July 1895. The city did not require the submission of

information to the Health Officer for cases of typhoid fever at the time. Therefore, Dr.

Kober began his investigation using death certificates held by the Health Officer since it

was compulsory to submit the certificates to this office. Working backwards from the

death certificates listing typhoid fever as the cause of death, Dr. Kober contacted

hospitals and all local physicians for information about all of the known typhoid

mortalities and to report all cases of typhoid fever that they might have treated. In this

manner, Dr. Kober received information about 149 deaths from the death certificates

submitted to the Health Officer and reports of 428 cases of typhoid fever that occurred

between 1 July and 31 October, 1895 from doctors and hospitals. Given that the

morbidity data were collected based upon the memories (and varied degrees of record

keeping) of local physicians, Dr. Kober believed that he had collected a reasonable

sample of typhoid fever cases for his investigation and used the mortality information to

calculate an estimate of the actual number of cases, which were approximately 795.[4]

Ultimately, Dr. Kober investigated 500 reported cases of typhoid fever from both local

---

[4] Dr. Kober described how he estimated the approximate number of typhoid cases in the following way, using the 149 reported deaths as a guide. First, six deaths were removed from the calculation as being of non-local origin. Of the remaining 143 deaths, Dr. Kober calculated that if each death represented 10 cases, then there would have been 1430 cases of typhoid between 1 July and 31 October of that year. If instead, 20 percent of the cases were fatalities, then there would have been 715 cases that originated in Washington, D.C. in 1895. According to Dr. Kober one of his contemporaries calculated typhoid's fatality rate at approximately 17.5 percent and the U.S. Army's doctors reported a fatality rate of 11 percent. Based upon this information, according to Dr. Kober, and the fact of the epidemic occurred in the summer when typhoid fatality rates (and case rates) tend to increase led him to estimate the typhoid fatality rate around 18 percent. Using this fatality rate and the 143 reported deaths of local origin would indicate approximately 795 cases occurring during the four summer months of 1895 (Kober 1895, 253-254).

and non-local origins, which to his mind would provide a good sample of the total cases. He indicated that through investigating the 428 reported typhoid cases that he anticipated discovering more morbidities. However, there is no specific mention as to how Dr. Kober identified the source(s) of the additional 72 cases not included in the originally reported 428 cases.

The investigation of each reported typhoid case included completing an information card with the person's: name, age, sex, race, address, date of attack, travel outside of the city prior to the attack, consumer of Potomac water, consumer of well water, location of the well, and name of the milkman. From the information sought about each typhoid case and Kober's own summary of the causes of typhoid fever, it is clear that germ theory was embraced as the primary theory of disease causation in this special investigation (Kober 1895).

The results of this investigation were written as a report including a map of 500 cases of typhoid, including fatalities. The typhoid morbidity and mortality data included cases of both local and non-local origin. These point locations were mapped to the residence level, but due to the scale of the map, confidence in spatial accuracy is reduced to their appropriate city block. Kober's report encompasses only cases that occurred between 1 July and 31 October, 1895.

These data pre-date the building of the water filtration system in Washington, D.C., as well as many of the capital improvements made to the city's sewerage system (Green 1963; Rosenau, Lumsden, and Kastle 1907). Additionally, Dr. Kober's investigation determined most cases of typhoid fever to be related to drinking contaminated well water and that a large percentage of the impacted households

continued to use box privies. The specific linkages to local environmental factors causing the 1895 epidemic and the identification of two parts of the city having greater prevalence of typhoid, Northeast and the Suburbs, provide a contrasting situation to that described in the later 1906 to 1909 PHS typhoid reports (Figure 1.2). In other words, in the later reports most cases were not linked to a particular cause, whereas in 1895 most cases were linked to contaminated wells. Additionally, Dr. Kober (1895) specifically references areas of the city where typhoid appeared to cluster, which again did not happen in the later reports that claimed typhoid to be generally distributed through out the city (Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911).

By exploring the 1895 typhoid fever epidemic two separate research questions are possible. First, the 1895 data provide a glimpse into the geography of typhoid morbidity compared to its mortality, something which is not often readily available for this time period and even more rarely available at this scale. A typical hypothesis to be tested would be that typhoid morbidity and mortality data exhibit the same spatial pattern. In effect, no clusters should be found of one without the other. Second, these data are an example of a typical typhoid pattern before the improvements in infrastructure between the 1890s and early 1900s. Therefore, these data provide a possible test dataset for current cities in developing countries where water and sewage systems more closely resemble pre-improved Washington, D.C.

The questions mentioned above can be explored using the G_i^* statistic to determine if typhoid fever clustered in particular parts of the city and at what scale it clustered. Given that Dr. Kober identified well contamination as the primary source of typhoid fever the resulting clustering surface should include a skewing towards small

cluster sizes of 150m or 250m. Also, if the typhoid cases did cluster locally, did these locations coincide with those identified by Dr. Kober? Further, was there similarity between typhoid morbidity and mortality clustering?

Methods

The basic methodology needed to answer the questions posed in this chapter was described in detail in chapter 2. The specific data used in the analysis here are the non-fatal cases of local origin and fatal cases of local origin that were extracted from the map of all 500 cases of typhoid that occurred between 1 July and 31 October, 1895 included in George Kober's special report on typhoid in the District of Columbia. Cases and mortalities of non-local origin were excluded since these cases were not caused by environmental conditions in Washington, D.C. All of the data included on the map (cases and deaths from both local and non-local origin) were heads up digitized and given an attribute according to the type of typhoid case or death. Using the grid of city blocks described in chapter 2 a count of points in each block polygon was run using Hawth's Analysis Tools (www.spatialecology.com).

The existing 1895 typhoid GIS can be used to further explore the typhoid morbidity clusters identified using the G_i* statistic, in particular fatality rates can be calculated for each cluster in order to assess the differences between the mortality and morbidity surfaces. In order to visualize the fatality rates of each statistically significant morbidity cluster, a buffer is extended to the size of the cluster around the centroid of each significant block. In other words, if block A represents a significant cluster at 750m, then the circular buffer is extended to a radius of 750m from that centroid. The Polygon in Polygon tool included with Hawth's Analysis Tools is used to summarize the number

of cases and fatalities on each block contained within each of the buffers created for the significant blocks. Finally, using the Field Calculator in the attribute table for the newly created buffer shapefile the percentage of typhoid cases that turned into fatalities was calculated as: deaths/cases * 100.

Results

The results of the mortality and morbidity digitizing are contained in Figures 3.1 and 3.2. Cases and deaths appear in all parts of the city and the suburbs. Please refer to Figure 1.2 for the location of the named regions under discussion in the remainder of this chapter. Based upon a purely visual assessment, the cases in the suburbs appear to be more clustered than those in the city proper. Additionally, there appear to be two "gaps" in the distribution of death locations, in Northeast and Foggy Bottom, both are highlighted in Figure 3.2 to show that there were cases in these areas, but they did not become deaths.

G_i* - Mortality

Significant clusters of typhoid mortality occur on 17 blocks in the city. Fourteen of these clusters are located to the east and west of Union Station (Figure 3.3). The strongest cluster is a block with four deaths on it and a G_i* value of 6.02. The remaining three blocks that were part of significant clusters of typhoid mortality are located in Anacostia and are associated with hospital deaths. The inclusion of the three blocks was an artifact of how the data were originally reported from an orphanage. There were no mortality clusters larger than 750m, although 11 of the 17 total clusters were either 500m or 750m in size, indicating a slight skewing towards larger cluster sizes (Table 3.1).

**Deaths from Cases of Local Origin
Between 1 July and 31 October, 1895**

Each dot represents one death

0   1,000

Meters

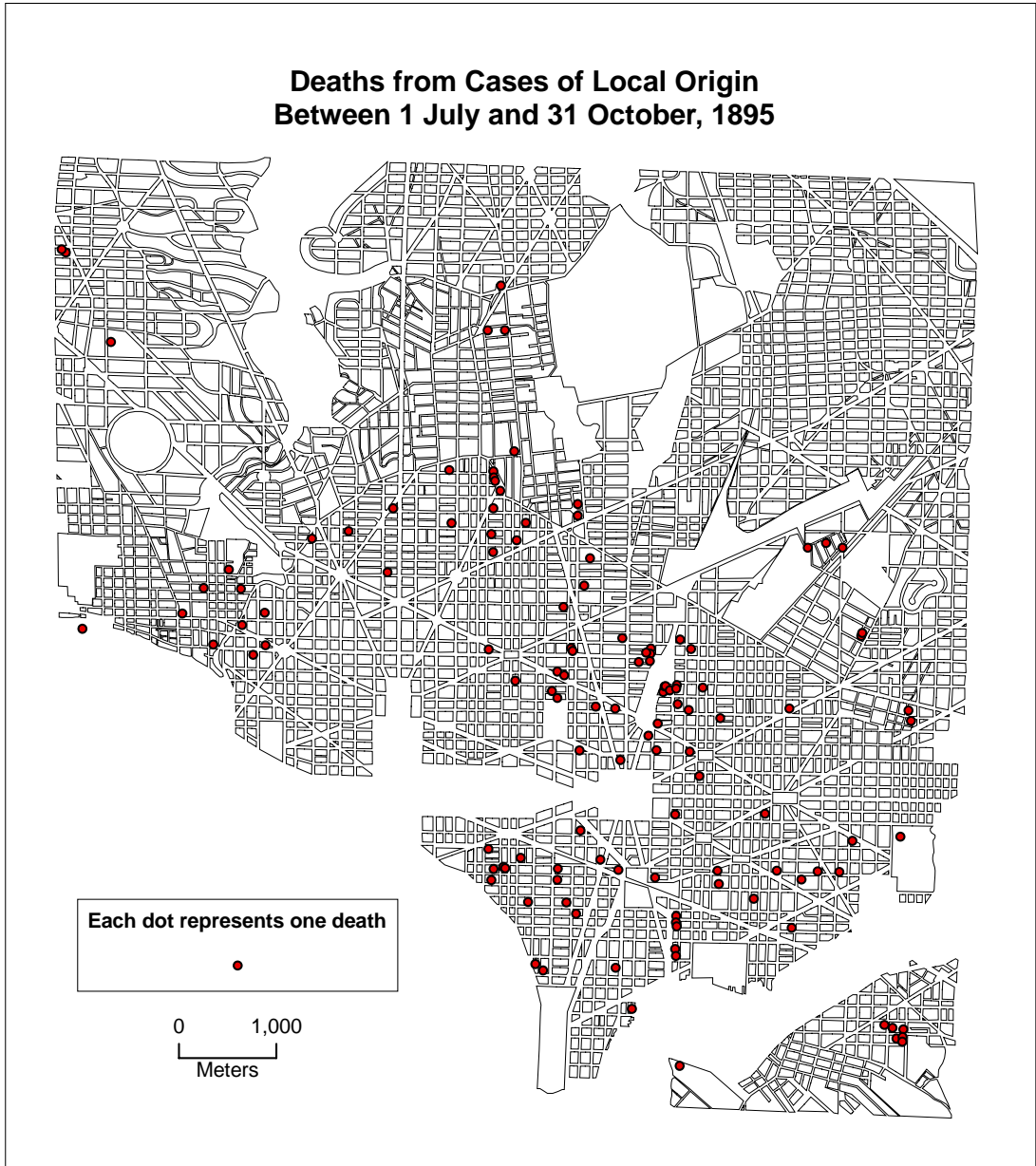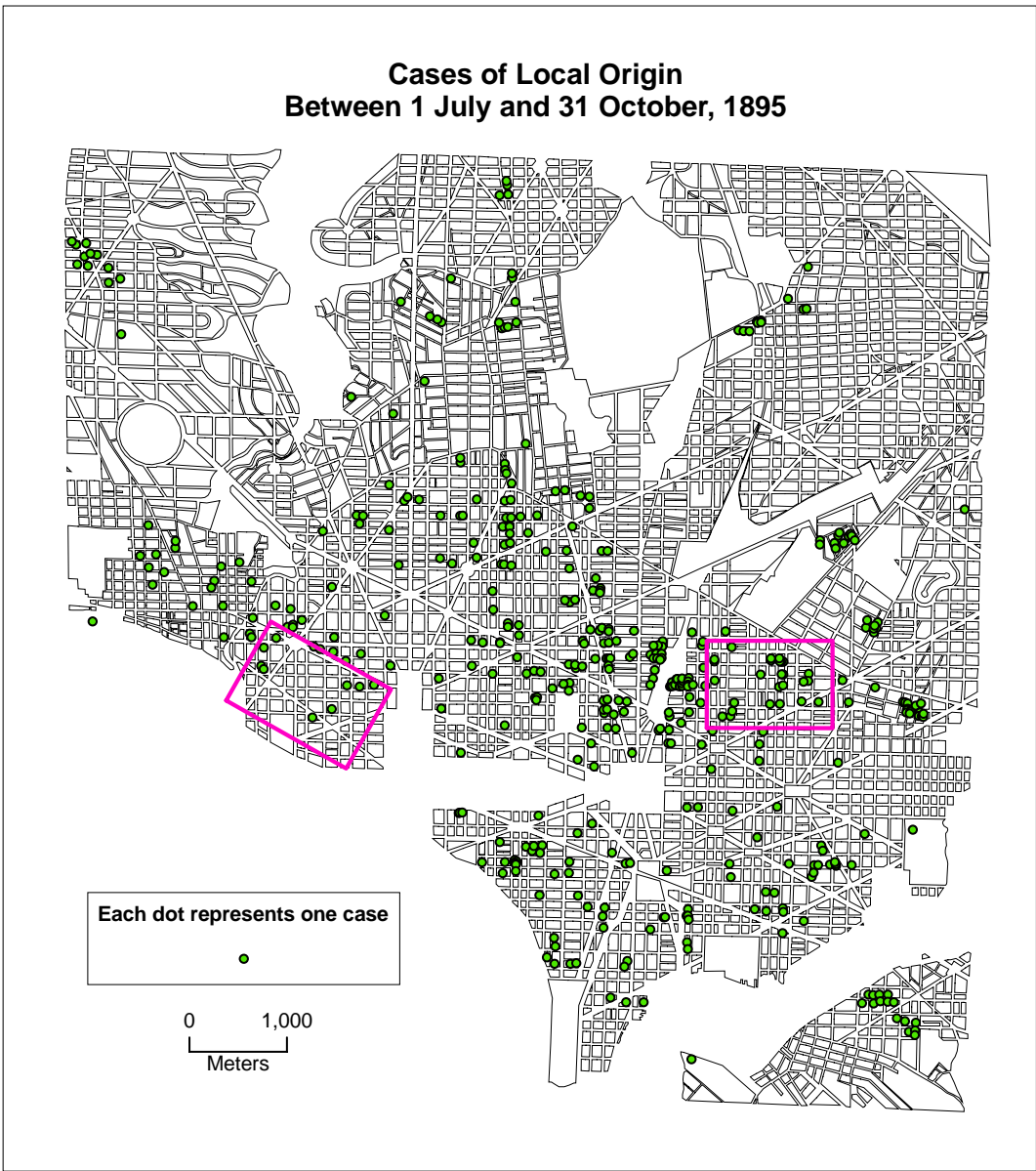**Figure 3.1** Typhoid mortality of local origin

Figure 3.2 Typhoid morbidity of local origin. Pink boxes highlight two areas where cases occurred, but no deaths.

**Clusters of Typhoid Fever Deaths of Local Origin,
1 July through 31 October, 1895**

Cluster Size (m)

150
250
500
750
1000

0   1,000

Meters
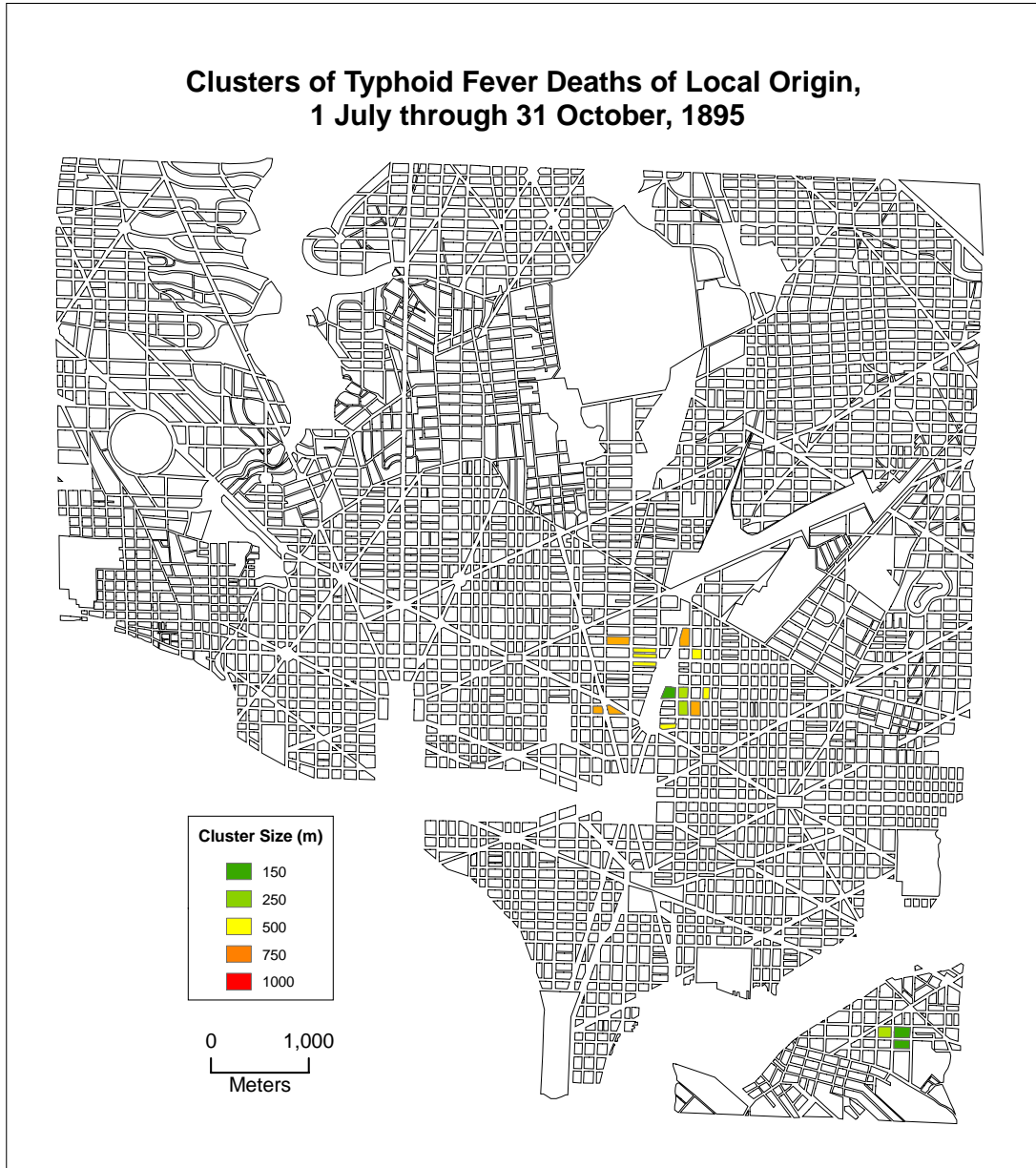
Figure 3.3 Results of G_i* statistic for fatal typhoid cases of local origin

Table 3.1 Summary of the *Gi** results for morbidity and mortality in 1895

	Distance (m)	Cells	min G_i^*	max G_i^*
Mortality				
	150	3	4.02	6.03
	250	3	3.05	3.08
	500	6	2.04	3.99
	750	5	2.18	3.21
	1000	0	0.00	0.00
Morbidity				
	150	8	2.04	10.06
	250	5	2.77	4.22
	500	14	2.02	4.99
	750	14	2.08	4.61
	1000	15	2.06	3.94

G_i^* - Morbidity

The cases mapped and reported by Dr. Kober clustered locally on a total of 56 blocks in Washington, D.C. Thirty-nine of those significant clusters are grouped on the east and west sides of Union Station, just north of the Capitol Building in Northwest and Northeast (Figure 3.4). The group of six significant clusters in the far eastern part of the city constitute another area of interest, although the G_i^* values for these clusters are low, being less than three. Much of the city is devoid of statistically significant typhoid

63

clusters, but the point data indicate that a number of other groups of typhoid cases throughout the city existed (Figure 3.2). Dr. Kober's report describes numerous groups of cases that were significant in the context of the epidemic that did not meet the statistical criteria used for the G_i^* analysis. Clusters occurred at all of the distances tested, but tended toward larger cluster sizes, 500m or larger (Table 3.1).

<u>Fatality Rates</u>

Figure 3.5 displays the percentage of typhoid cases of local origin that became deaths between 1 July and 31 October 1895. Each block shown in a color on this map represents the center block of a significant cluster. Fatality rates by cluster range from 11.76 percent to as high as 41.67 percent.

Discussion

The use of the G_i^* statistic answers the key question posed here concerning whether or not typhoid fever deaths and/or cases cluster in any particular parts of Washington, D.C. As would be expected there are many more clusters of cases than of deaths, and the hotspots of typhoid mortality coincided with clusters of cases. The latter part of the previous statement helps to identify that deaths from typhoid tend to cluster in the same locations as cases. Still, comparing the results of the two clustering surfaces does identify that focusing solely on mortality data can limit the scope of understanding the geography of an outbreak.

The map of the fatality rates is particularly interesting as there appears to be a geographic pattern to the rates. For example, the nine of the ten clusters with the highest fatality rates greater than 32.69 percent, are all located directly to the east of Union Station, while nearly all the significant clusters with the lowest fatality rates are on the
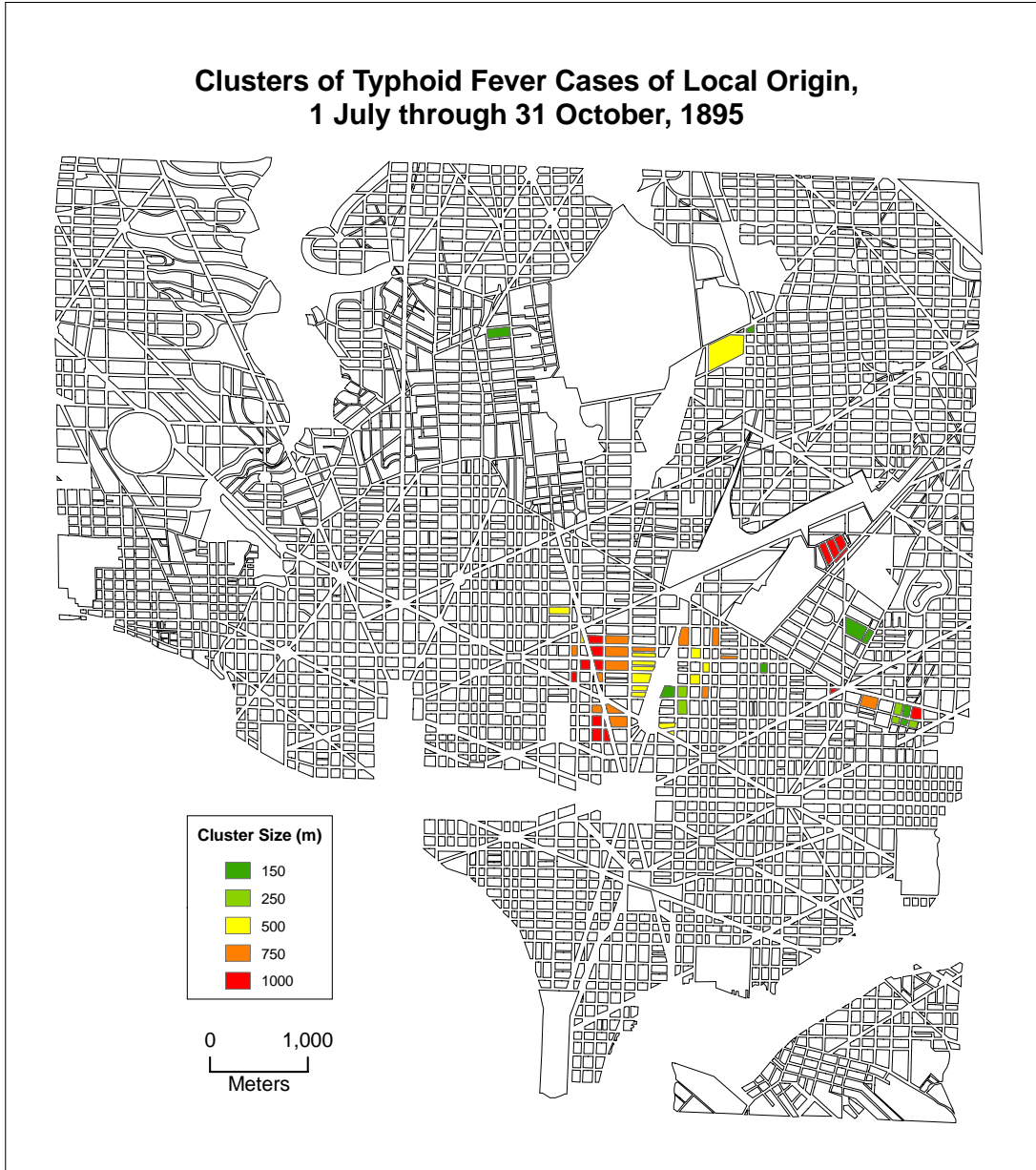
Figure 3.4 G_i* results for the all non-fatal cases of typhoid in 1895
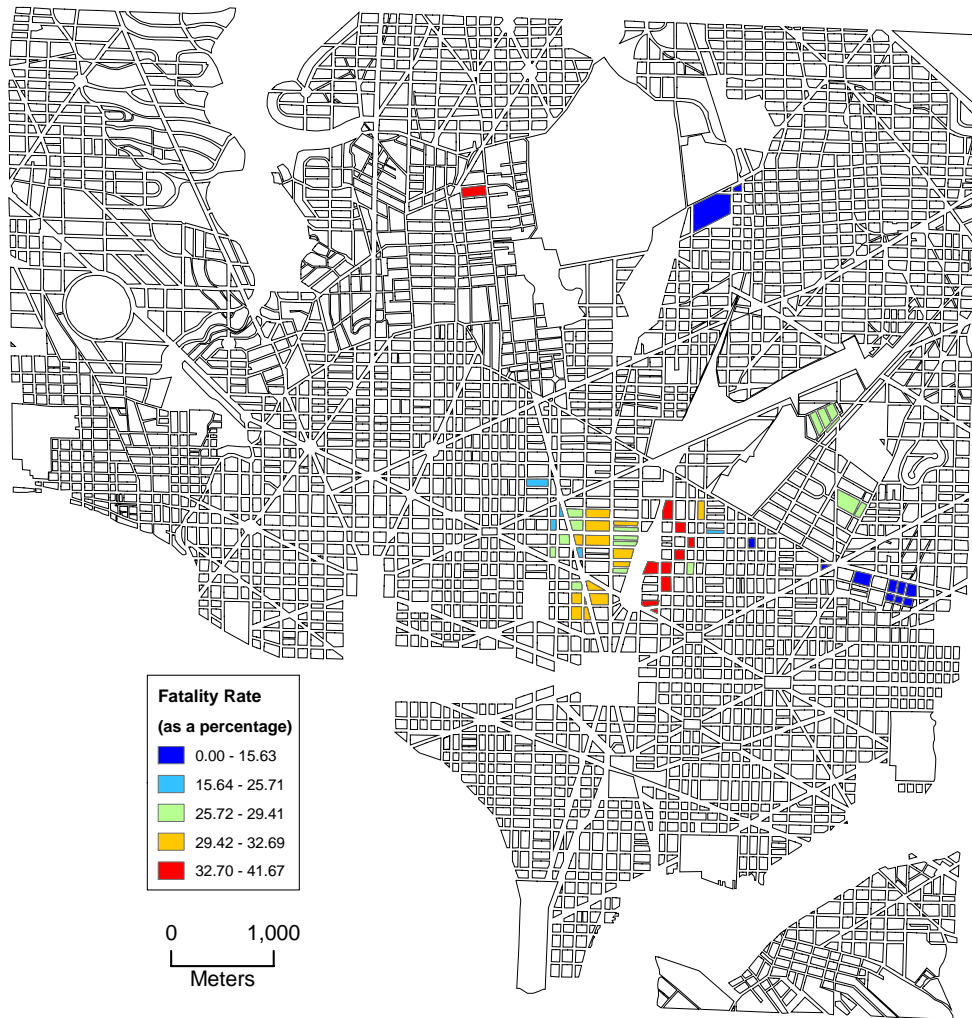
Figure 3.5 Fatality rates for all significant morbidity clusters

eastern edge of the study area. This patterning seems to indicate that the location of a hotspot might impact an individual's chance of recovery from typhoid.

The 1895 typhoid GIS also aids in the exploration of whether public wells were a potential source of typhoid in Washington, D.C. Forty-four of the 56 significant clusters of typhoid cases included a contaminated well within the spatial extent of that cluster. Many of these were 1000m clusters, raising a question about the relationship between contaminated wells, since if one extends the size of a cluster far enough, that cluster is quite likely to eventually contain a polluted well. However, only eight of the 21 wells confirmed to be contaminated were located within the significant critical distance of a cluster.

These results do indicate a few things. First, proximity to a contaminated well does not necessarily mean that a block is a member of a significant typhoid cluster. Two contaminated wells did not have any cases within 500 meters of their locations, just as not all clusters were within a test distance of a contaminated well. When fatality rates are calculated for both, 250m and 500m around all 21 contaminated wells the rates do vary across the city. For example, the highest fatality rate within 500m of a contaminated well is 55.55 percent of the nine cases that developed in that area in Southwest, whereas, of the eight cases within 500m of a contaminated well in Georgetown there were no fatalities. As will be seen in a number of instances throughout this dissertation, the small number of cases or fatalities requires careful interpretation of the results. More cases and fatalities occurred near Union Station than in Southwest, therefore, while often times the fatality rates in Southwest appear to be higher (one of two cases turned into a death), the greater number of cases in other parts of the city tend to provide more accurate fatality

rates. Dr. Kober determined the typhoid fatality rate to be around 20 percent of all cases, and using the data entered into this GIS, the fatality rate is approximately 30 percent of all cases. Therefore, the results presented here must be interpreted with caution and consideration given to the number of cases and deaths being used to calculate the fatality rates. Fatality rates most likely did vary across the city. A good example of this is the comparison of the fatality rates for the blocks immediately to the east of Union Station to the group of six blocks in the extreme eastern part of the city. The number of cases varied between these two areas, the six significant clusters in the east all had fatality rates of less than 20 percent, while all of the blocks to the east of Union Station had greater than 30 percent fatalities from cases of typhoid (Figure 3.5).

The special investigation conducted by George Kober also considered specific sources of typhoid cases, and in the process of doing so not only mapped all of the investigated cases, but discussed some spatial aspects of cases in neighborhoods of interest. Some of these cases discussed by Dr. Kober did not fall into statistically significant clusters using the G_i* statistic. This poses the question of whether Dr. Kober's observations are insightful or erroneous. Overall, Dr. Kober concluded that contaminated public wells were the culprits behind many of the typhoid cases. While not all cases could be directly linked to a well that tested positive for fecal contamination, those cases that could be directly linked to contaminated well water usually were in close proximity to the contaminated well and/or to a place for human waste disposal indicating a potential means for soil contamination or for flies to transfer the bacteria to food. In addition to the usual means of typhoid infection, polluted water and milk, Dr. Kober described other possible avenues of contracting typhoid, including fomites (infected clothing), infected

hands, or infected sewer air. In the results of his study no cases were directly linked to these other possible sources, but they are still relevant to understanding the context within which the report was written (Kober 1895).

The list and companion discussion of possible origins of the typhoid outbreak not only provides insight into how typhoid entered Washington's urban system in 1895, but also the state of medical knowledge at the time. The status of medical knowledge would, of course, influence which potential origins of typhoid were considered, which were ignored, and the evaluation of the information collected. One fascinating piece of information that appeared in the text was a reference to miasmatic theory.

> All scientific physicians agree, however, upon one point, viz, that typhoid fever is caused by an organized germ capable of reproducing itself within and without the body, instead of such hypothetical matter as miasms or contagia, whose nature has never been demonstrated to our senses. On no other theory except the germ theory can we explain the occurrence of typhoid-fever epidemics, spread through the water and milk supply (Kober 1895, 257).

From the above excerpt of the 1895 typhoid report it is apparent that although germ theory was the prevalent theory of disease causation at this time, threads of miasmatic theory still lingered on. Dr. Kober's statement, clearly addresses the futility of any remaining belief in miasma.

What is more interesting, is how in subsequent statements references are made to the possibility of contracting typhoid through infected sewer air. Examples are used from other cities that supposedly linked the inhalation of typhoid-infected sewer air rising from defective plumbing in homes. Dr. Kober firmly states that based upon the amount of evidence for such happenings that it is likely that "when [typhoid fever is] present in stagnant sewers or in the soil, may be liberated and infect the air… [and that there is

evidence] that the germs of malaria are carried from the soil by the ascending currents of air" (Kober 1895, 260). So while miasmas could not be the cause of disease, typhoid bacteria and even malaria bacteria could became airborne, much like miasmas that one could then breathe in and become ill. Conceptually, the idea is not too far fetched, though still fundamentally wrong.[5] Certainly some diseases spread through the air, such as tuberculosis, which can spread by the inhalation of an infected respiratory droplet exhaled by a person with the disease. Neither typhoid nor malaria can spread in this way.

This report was written prior to a few key discoveries concerning typhoid, including that of healthy carriers (1902) and how the disease could spread through contact (not washing one's hands after using the water-closet and then coming in contact with food or drink). Reference is made by Dr. Kober to infected hands being possible modes of typhoid transfer, but not in the same context as our modern knowledge of hand washing. Whipple (1908) shed light on the issue of transfer through infected clothing by mentioning that infected fecal matter had been found on clothing, thus supporting the concept of this means of transfer, but Whipple emphasized that the transfer of typhoid through clothing was due to the presence of typhoid-contaminated fecal matter on that clothing rather than the clothing acting as a particular type of typhoid reservoir on its own.

Milk as a common vehicle for the spread of typhoid is also discussed by Dr. Kober. Here he distinguishes the means by which milk becomes infected, since cows cannot contract the disease or pass typhoid directly into their milk. Instead, milk infections tended to develop because of insanitary practices on farms and in dairies, such

---

[5] The mosquito was identified as the vector for malaria in 1897, thus this report pre-dates that piece of disease understanding.

as washing milk containers in contaminated well water or cows wading through polluted water and bacteria ending up on the udders. Drawing from 130 typhoid epidemics he investigated world-wide, Dr. Kober mentioned milk-related epidemics developing from utensils being soaked in contaminated well water, diluting milk (which was considered illegal) with contaminated well water, infected persons with mild cases continuing to work, which "those who are familiar with the personal habits of the average dairy boy will have no difficulty in surmising the manner of direct digital infection" (Kober 1895, 259).

Dr. Kober included a number of detailed descriptions in his report of specific groups of cases, their local-level geography, and possible sources. One of these local areas he described was the group of six significant clusters in the far eastern part of Northeast (Figure 3.6). Here, according to Dr. Kober, was an unusual prevalence of typhoid fever that could be traced to both a polluted well and to contaminated milk, although the specific contributions of each were unknown. The polluted well was located within 20 feet of an overflowing barrel used to collect human waste. When tested the well was confirmed to be positive for fecal matter, and thus was presumed to be a source of typhoid fever. Additionally, milk cows, owned by the residents on these blocks, wandered about the common land near these homes. These cows grazed along the banks of a stream polluted by sewers. The wastes in this stream could contain typhoid and Kober himself observed these cows wading in the stream. It is feasible, although unlikely, that their teats and udders could have come into contact with the bacteria infected water, which in turn found its way into the milk during the milking process (Kober 1895, 266).

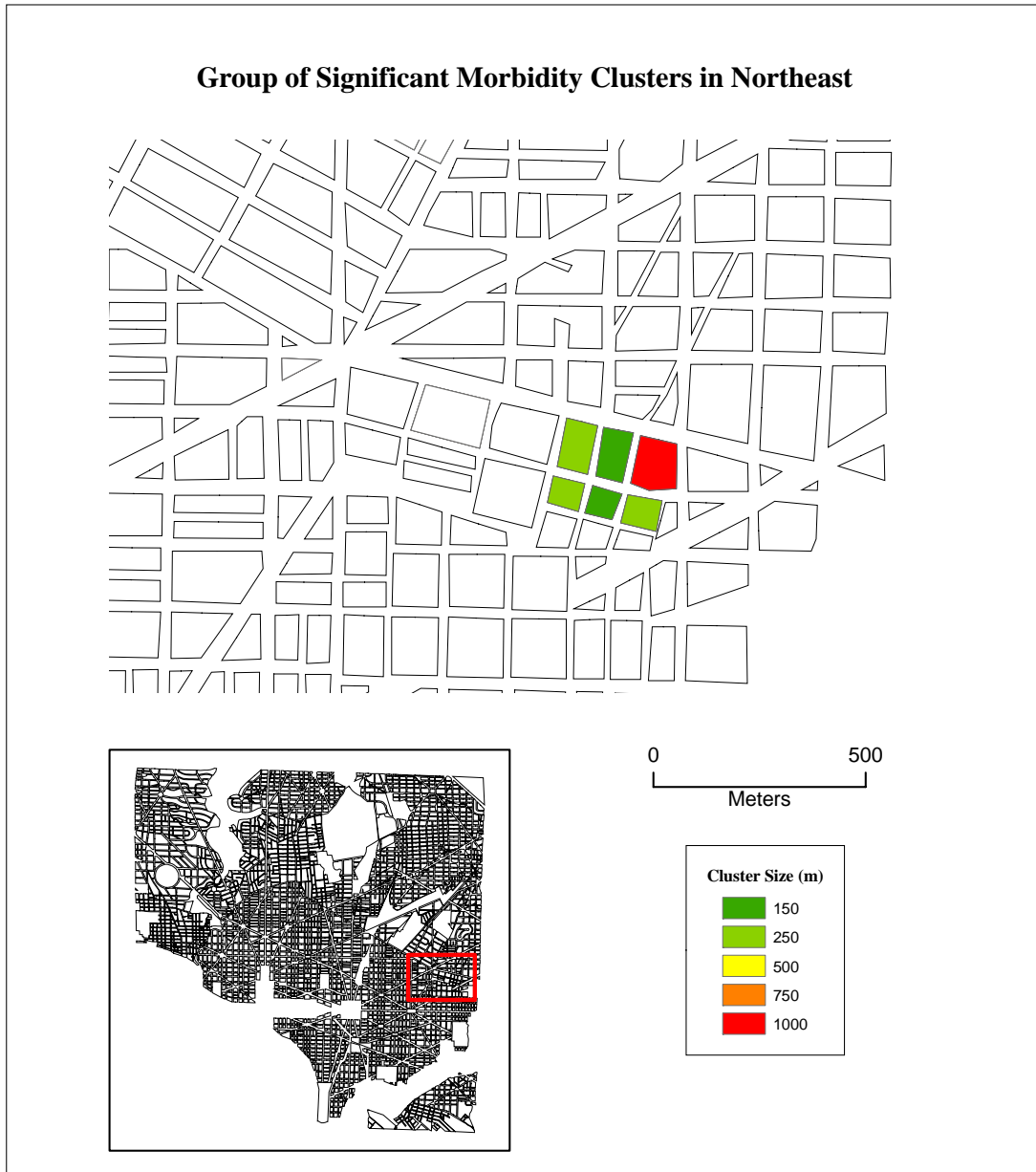**Figure 3.6** Detailed map of six significant typhoid morbidity clusters in Northeast

The typhoid mortality data from this focused study did provide enough information to produce some hotspot results, but as can be clearly seen by comparing the morbidity and mortality G_i* results, the mortality data returned much narrower results (Figures 3.3 and 3.4). Had these been the only data available, analyses would have been

severely limited due to a lack of fatal cases. For many other years during the late

nineteenth century there were too few typhoid fever deaths reported to the city's health

officer to permit the use of the G_i* statistic. For example, while the health officer's report

for the year ending June 30, 1898, included a map of typhoid fever deaths, not enough of

these 90 deaths occurred on common blocks to run a valid G_i* statistic test. In 1880, there

were even fewer typhoid deaths, 66, again without a G_i* compatible geography for

analysis.

There was an increase in typhoid fever morbidity and mortality rates in the 1890s

for reasons that are unclear (Kober 1895; Rosenau et al. 1907). A more virulent strain of

typhoid fever may have entered into Washington, D.C. in those years, leading to an

increase in prevalence. Alternatively, changes in the local urban environment may have

been behind the increases in typhoid. Regardless, the typhoid epidemic of 1895 provides

a snapshot of typhoid fever patterns before the expected solution to the city's typhoid

problem, the water filtration plant, was opened in 1905. Due to the paucity of typhoid

mortality data by the turn of the twentieth century, when available, morbidity data

increase the chances of being able to statistically test for local clusters by providing more

events with which to work. Based upon the analysis here mortality data can display

clusters, and these do tend to coincide with clusters of cases, but clearly morbidity data

provide a more complete picture of typhoid's spatial patterning. Still, from George

Kober's report, even identifying statistically significant clusters does not account for all

related groups of cases. The remaining chapters of this dissertation focus upon typhoid

morbidity data collected by the United State Public Health Service (PHS) between 1906

and 1909. The analyses in those chapters build upon the ones used here and incorporate

additional techniques beyond statistical cluster analysis, even if that is where the

exploration begins.

CHAPTER 4
SPATIAL AND TEMPORAL STRUCTURE OF SUMMER
TYPHOID FEVER, 1906-1909:
EVALUATING LOCAL CLUSTERING WITH THE G_i* STATISTIC[†]

Introduction

The city of Washington, D.C. added a sand filtration plant to its water supply

system in October 1905, intending to reduce water-borne disease outbreaks and

particularly the annual typhoid fever rate. However, during June 1906 a typhoid

epidemic erupted in the city resulting in a higher morbidity rate than before the 1905

infrastructural improvements (Table 4.1). In reaction to this situation, the city's Health

Officer commissioned the PHS to investigate the origins of the epidemic, and

simultaneously conduct a bacteriological study of the water from all public wells in the

city. The PHS chose to continue annual investigations of typhoid until 1909, when the

typhoid rate in Washington dropped to acceptable levels for a city given its sanitary

infrastructure (Lumsden and Anderson 1911). Interestingly, in the original request for

help to the PHS, the Health Officer made direct reference to the general spatial

distribution of the cases during the typhoid outbreak, meaning that typhoid appeared in

all parts of the city rather than a single sector (Rosenau et al. 1907). This suggested a

mass consumption source of the disease rather than localized origins of infection, yet

each of the four reports failed to identify a single originating source of the disease. In

meeting the city's request for help, PHS doctors investigated the geographic distribution

of typhoid, resulting in a series of maps in each report displaying the location of each

[†] Note: The majority of this chapter has been previously published in the *International Journal of Health Geographics*. The citation is as follows:
Hinman, S. E., J. K. Blackburn, and A. Curtis. 2006. Spatial and temporal structure of
typhoid outbreaks in Washington, D.C., 1906-1909: evaluating local clustering
with the G_i* statistic. *International Journal of Health Geographics* 5 (13).

typhoid case studied (Lumsden and Anderson 1911; Rosenau, Lumsden, and Kastle 1907, 1908, 1909). These maps, supported by the accompanying descriptions in the reports, have now been transferred into a GIS allowing for more sophisticated spatial analyses - in effect finally meeting the mission charged of the PHS.

Table 4.1 Mortality and morbidity rates are listed per 100,000 population. Italics indicate that the numbers represent estimates of typhoid morbidity for July through October only.

Year	Mortality Rate	Morbidity Rate	Number of Cases
1880	47.3	-	-
1895	73.8	*289.87*	*795*
1900	74.1	-	-
1902	74.0	-	-
1905	43.9	-	-
1906	49.3	345	1126
1907	34.5	287	945
1908	36.5	275	936
1909	33.8	227	716
1910	23.2	205	780

In addition to considering the geography of typhoid cases between 1906 and 1909, the PHS investigators collected much of the same information about typhoid patients as George Kober described in his 1895 typhoid investigation and as William Sedgwick explained with regard to a typhoid investigation in Lowell, Massachusetts 1891-1892 (Kober 1895; Sedgwick 1893). Each case of typhoid investigated by the PHS was recorded on a case card, along with the patient's name, age, race, sex, address, sanitary condition of their home, their primary water supply, their milk supplier, observation of flies and mosquitoes in the house, water closet versus privy. Each of the

variables mentioned on the case card were tallied for all of the investigated cases and numerical calculations made, such as the number of people who contracted typhoid who drank only well water compared to those who only drank Potomac River water. The doctors, for example, considered whether or not their observations of abundant flies in a residence seemed to be connected to the majority of cases across the city, or the majority of cases where the victim also used a privy rather than a water-closet. The reports were essentially inconclusive concerning concrete spatial associations of origin and infection. Two milk related outbreaks were identified, one in 1906 and the other in 1908 (Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911). In 1909, a small number of cases were traced to particular sources. A well, that likely should have been closed sooner, accounted for four cases, and a handful of cases were caused by the consumption of raw oysters distributed through a less than reputable vendor (Lumsden and Anderson 1911). Beyond these few instances of finding the source(s) of typhoid fever in Washington, D.C. the cause of the sudden increase and subsequent decrease in typhoid morbidity remains unknown.

The spatial analytical capabilities of a GIS now allow for these four years' of typhoid cases to be further investigated in the form of three general research questions. The reports evaluated typhoid as being generally distributed throughout the city, GIS can be used to determine if the spatial distribution of typhoid in Washington, D.C. was truly uniform, or if clusters of disease did exist within the city. Second, if clusters were present, at what spatial scale did they exist – highly localized, or generalized across large parts of the city? City-wide epidemics would generate large spatial clusters with relatively homogenous morbidity surfaces. Although concentrated disease clusters would

likely occur as secondary sources of infection, the ratio of these to large clusters would diminish during city-wide epidemics. Finally, did these clusters remain temporally stable, with individual hotspots existing across multiple outbreak years? As the socioeconomic characteristics of neighborhoods were unlikely to change dramatically between years, any general source of infection, especially without specific intervention, would likely impact similar areas of the city annually. Additionally, if a small disease cluster were to appear for numerous years, this might reveal an underlying and consistent disease threat, such as an unsanitary well or a small manufacturer with poor hygiene facilities.

Methods

Using the methodology outlined in chapter 2, the two questions mentioned above can be answered. Here the digitized typhoid cases were temporally aggregated to whole "year" units. In this way all of the typhoid cases that developed in 1906 were analyzed together and compared to those in 1907. Since the typhoid data collected for 1906, 1907, and 1908 only included the months of May through October, the term "year" here refers to those six summer months and not all 12 months of a calendar year. For 1909, only the data from May through October were used in order to keep the analysis symmetrical with the other three years tested. First, the annual typhoid morbidity events were tested for global clustering using Ripley's K-function. Next using the block-based grid and the five test distances (150, 250, 500, 750, and 1000 meters) described in chapter 2, the G_i^* statistic was run to identify the presence of any local clustering.

To determine whether typhoid outbreak regions were temporally stable across the study period, a summary program was developed in SAS v9.0 to tabulate the total number of years (1 − 4) that a grid cell (city block) was a member of a significant cluster

at each given distance value. Additionally, the total number of years any given grid cell

was significant regardless of d_c was evaluated to identify stable cluster locations that may

have resulted from stable causes.

Results

      Figure 4.1 illustrates the original distribution of point data for each of the four

years on which the K-function analysis was applied. Figure 2.3 illustrates the block-based

grid used in the G_i^* analysis and represents the grid cells referenced in the individual
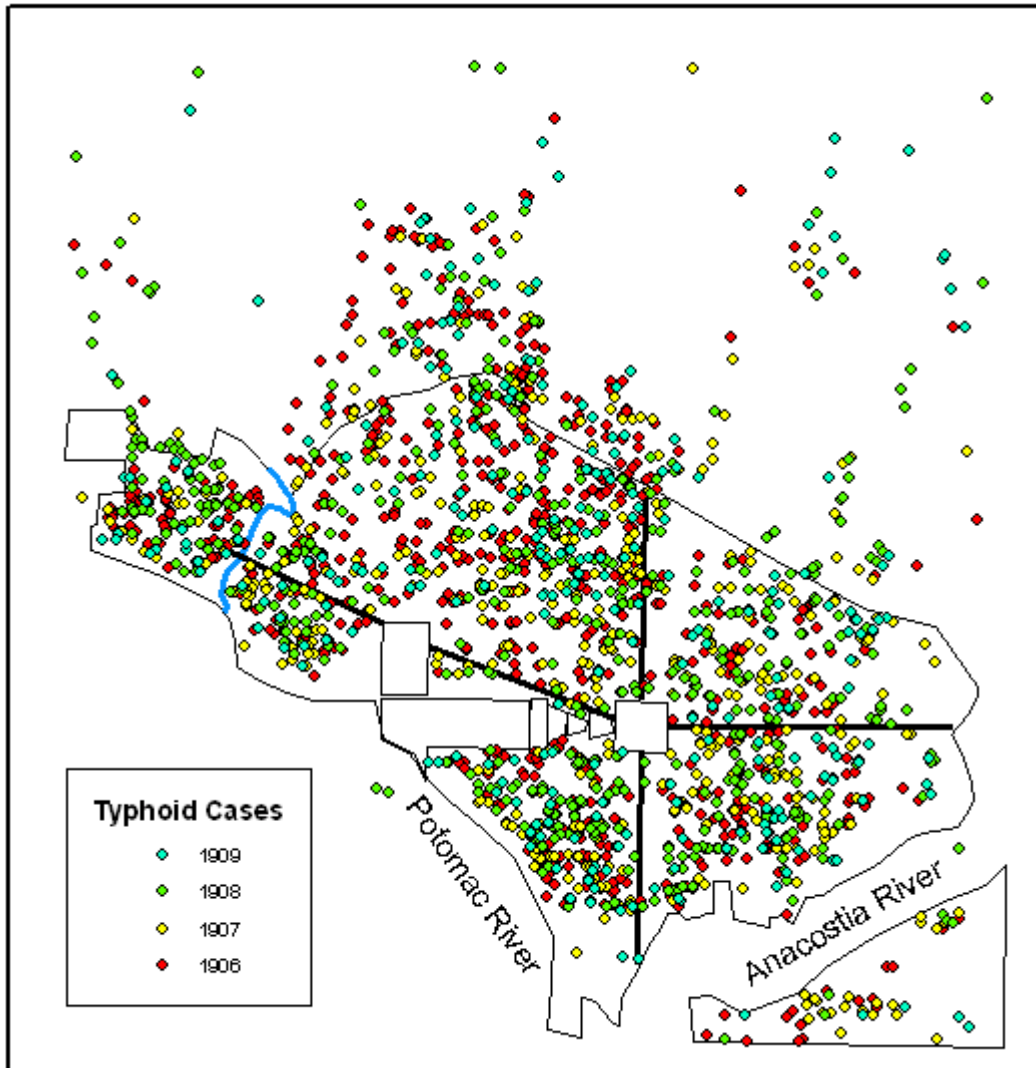
results sections.

Figure 4.1 All typhoid fever cases tested, 1906-1909

Ripley's K-function

     The global spatial auto-correlation statistic indicates global typhoid clustering in all four years. Table 4.2 and Figures 4.2 though 4.5 summarize the results of this analysis. The values for L(h) were fairly similar at all distances in all years studied, but the when plotted on a graph the L(h) values fell above the simulated envelope from the Monte Carlo simulations indicating global spatial auto-correlation within the case data for each of the four years under study.

Table 4.2 Ripley's K-function results, 1906-1909

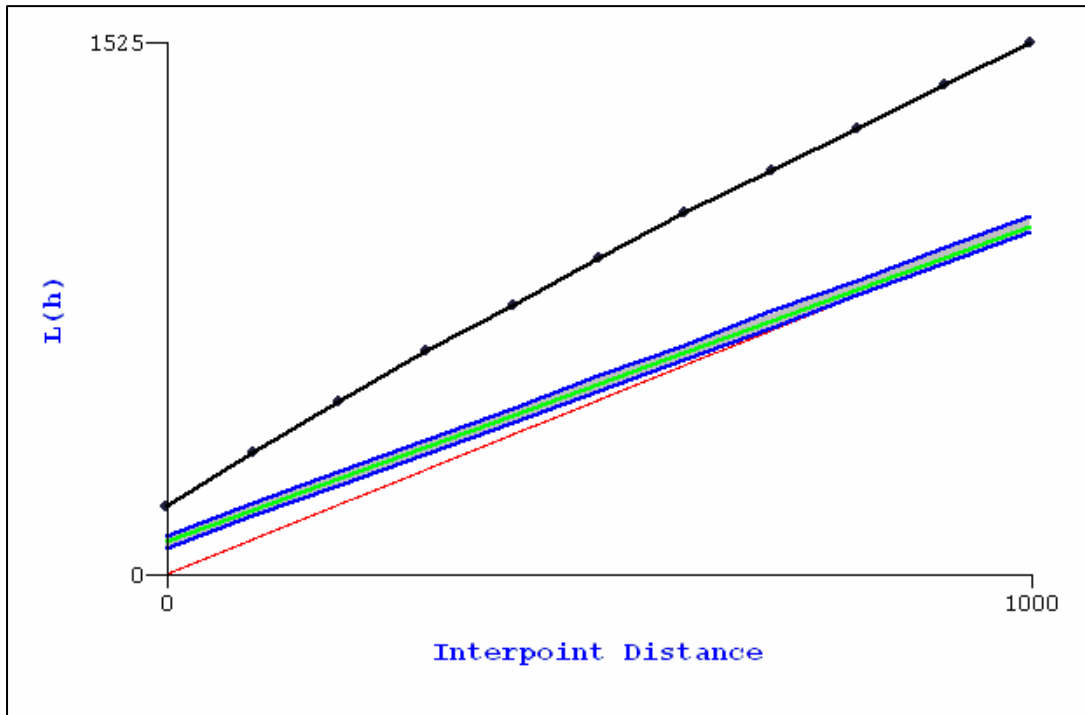| Distance (m)	1906	1907	1908	1909
0	192.88	197.80	216.99	170.77
100	350.24	368.19	381.27	312.82
200	495.41	511.65	534.79	457.20
300	639.88	644.29	681.38	603.48
400	772.84	789.88	833.06	729.71
500	905.25	914.98	974.53	858.60
600	1036.84	1047.79	1098.68	973.26
700	1160.23	1177.82	1226.26	1090.83
800	1279.21	1304.02	1350.98	1207.75
900	1404.37	1423.07	1468.40	1330.13
1000	1525.93	1535.74	1580.81	1438.93

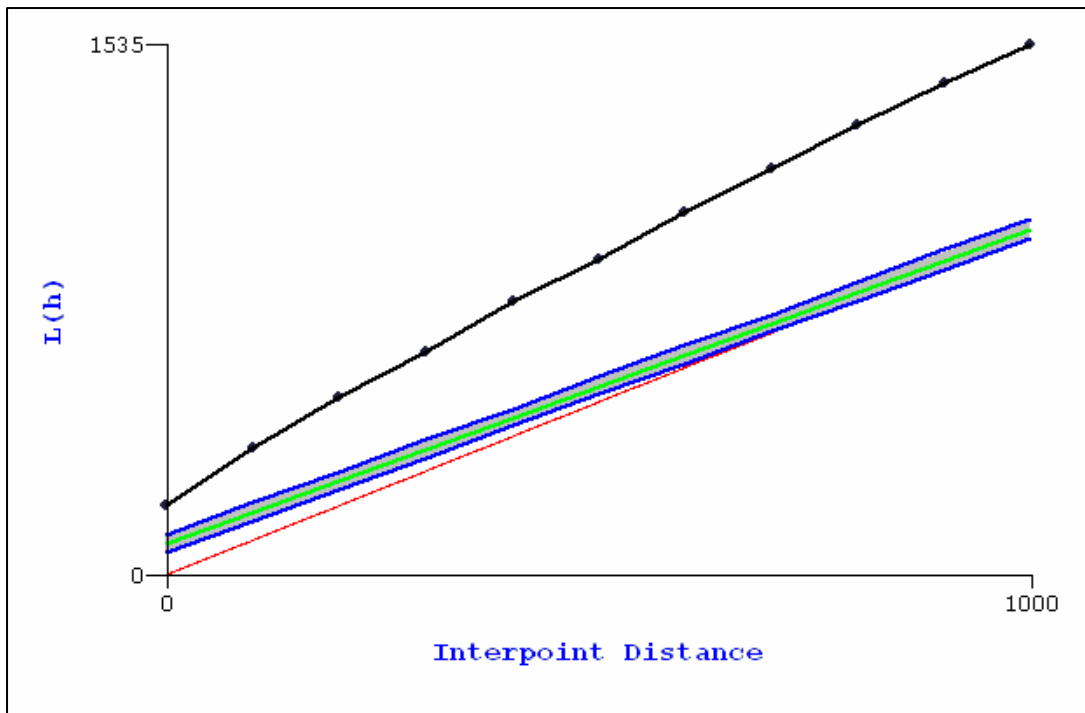**Figure 4.2** Graph of Ripley's K-function results for 1906

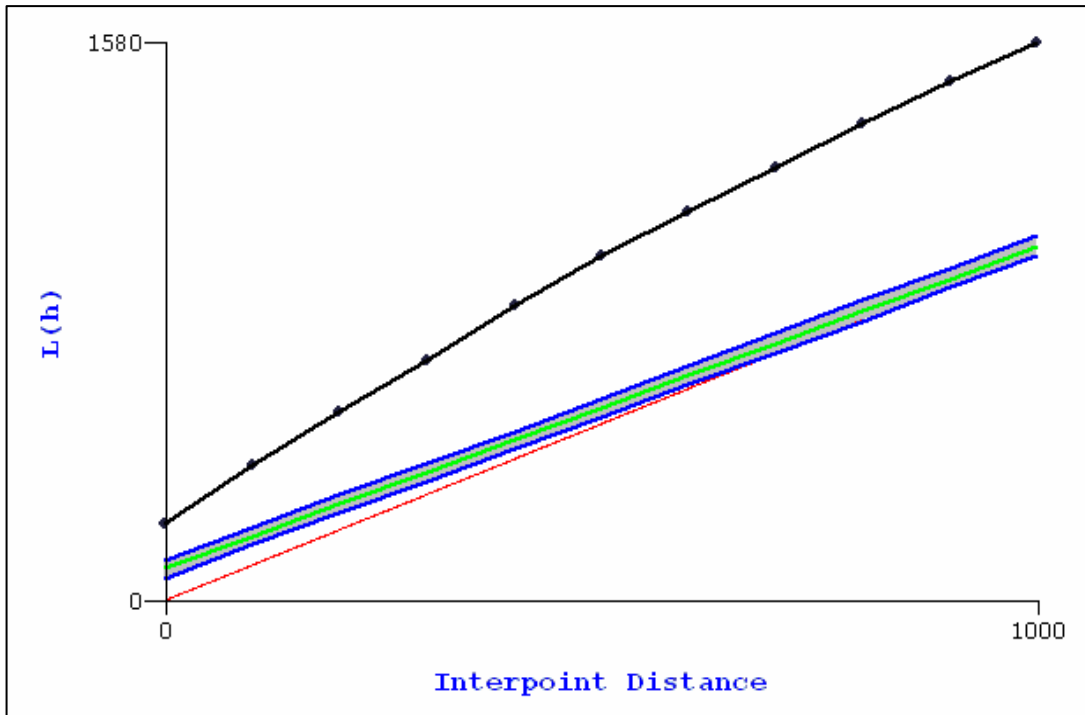**Figure 4.3** Graph of Ripley's K-function results for 1907

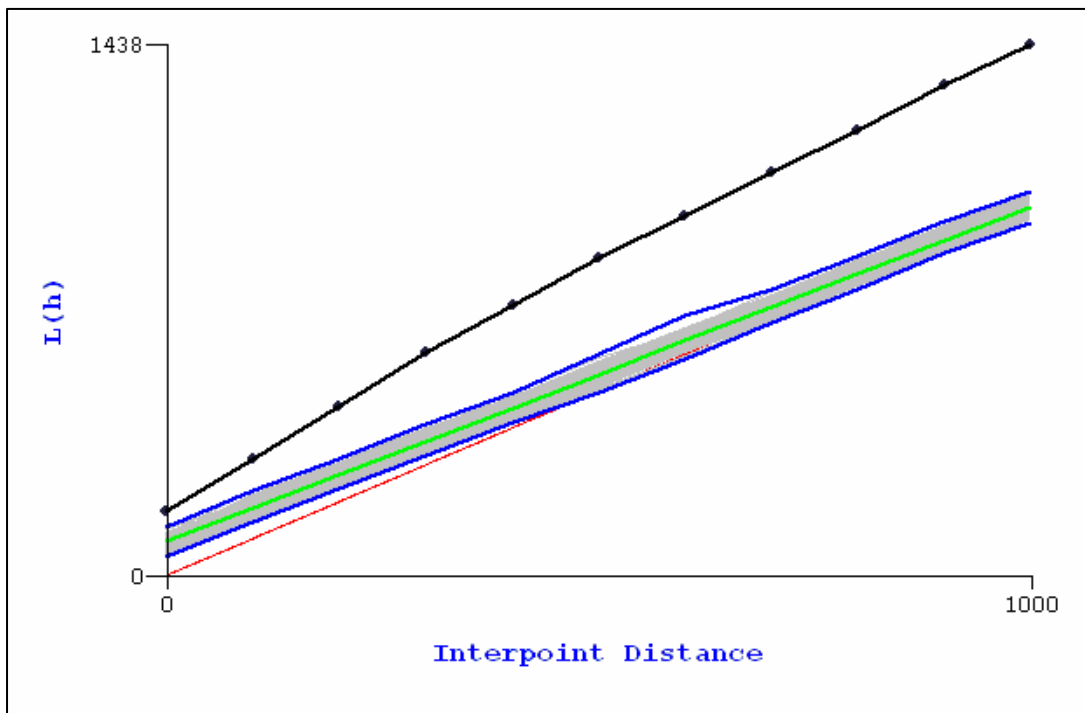**Figure 4.4** Graph of Ripley's K-function results for 1908

Figure 4.5 Graph of Ripley's K-function results for 1909

G_i* - 1906

The number of blocks that were significant at each distance value for 1906 are summarized in Table 4.3. The skewing towards the smallest cluster size suggests a localized geographic distribution of infection. Figure 4.6 displays significant blocks by cluster size for this year. The four 1000m clusters appeared in the suburbs near the northern boundary (2 clusters) and in Georgetown (2 clusters) (Figure 1.2). All of the clusters in Southwest were more localized, 500m or smaller.

G_i* - 1907

The summary of the number of blocks that were significant at each distance value are reported in Table 4.3. Figure 4.7 displays the significant blocks by cluster size for 1907. Like with the 1906 results, in 1907 the significant clusters were skewed towards the 150m distance. No clusters were reported for the 1000m distance.

G_i* - 1908

Table 4.3 summarizes the number of blocks that were significant at each distance value in 1908. Figure 4.8 displays the significant blocks by cluster size for this year. It is evident that Georgetown and Southwest were the areas primarily impacted during 1908, and these areas tended to display clusters at larger distance values. A few scattered small clusters appeared in Southeast.

G_i* - 1909

The number of blocks that were significant at each distance value for 1909 are summarized in Table 4.3. Figure 4.9 displays the significant blocks by cluster size for 1909. This year displays approximately the same number of blocks to have significant

clusters at each distance. The significant blocks were grouped mostly in Northwest and a

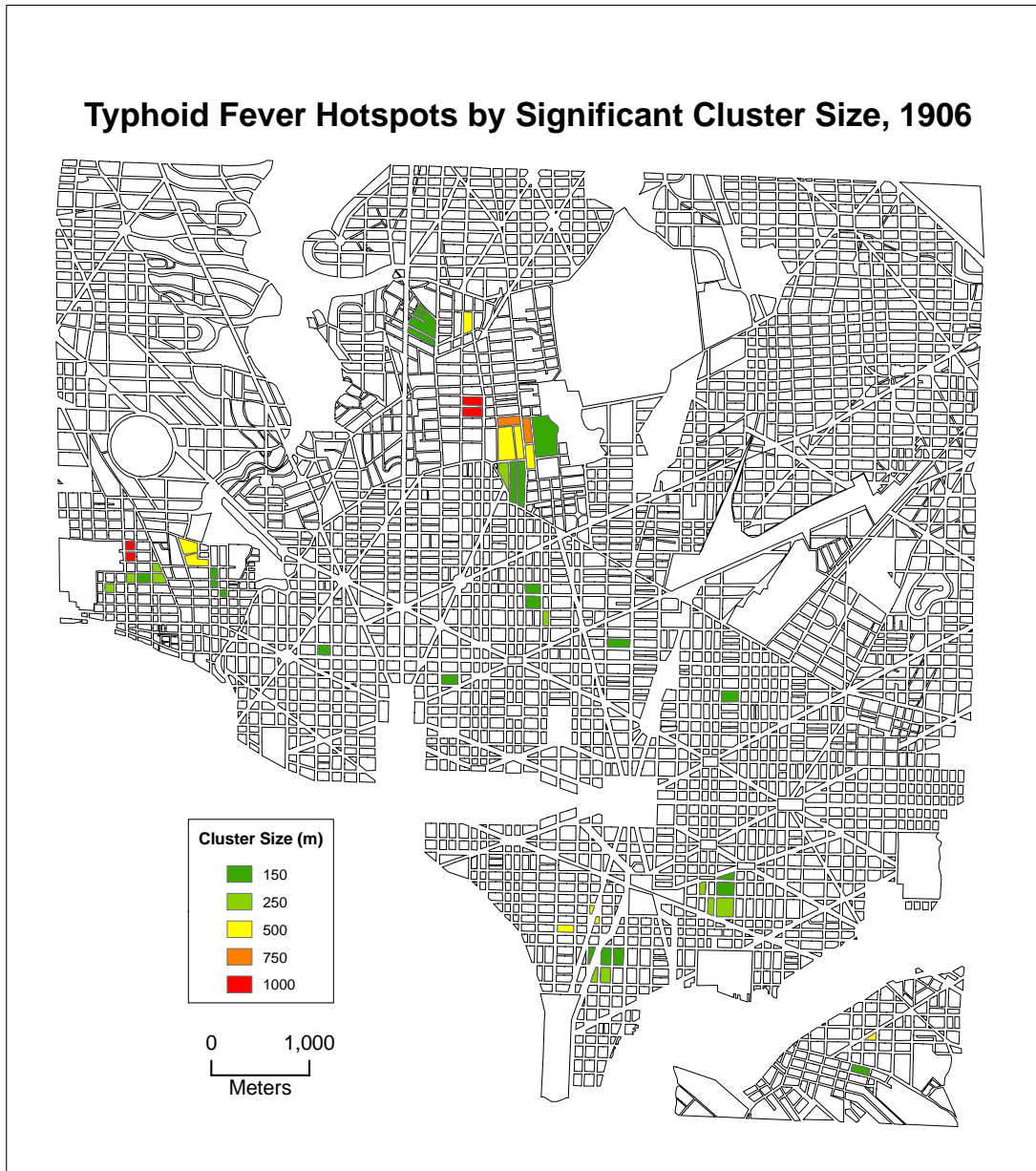few other clusters were scattered in Southwest and Southeast.

Figure 4.6 G_i* results for 1906

Table 4.3 Summary of G_i* results, 1906-1909

Year	Distance (m)	Cells	min G_i*	max G_i*
1906				
	150	21	2.03	4.43
	250	11	2.02	2.68
	500	9	2.03	2.72
	750	2	2.19	2.35
	1000	4	2.01	2.63
1907				
	150	20	2.41	5.48
	250	9	2.06	4.87
	500	6	2.15	3.60
	750	3	2.04	2.41
	1000	0	0.00	0.00
1908				
	150	16	2.22	5.04
	250	1	2.44	2.44
	500	15	2.07	6.12
	750	25	2.32	4.69
	1000	33	2.06	4.56
1909				
	150	9	3.65	5.17
	250	6	2.40	3.70
	500	6	2.05	3.78
	750	7	2.06	4.35
	1000	8	2.04	2.49

Typhoid Fever Hotspots by Significant Cluster Size, 1907
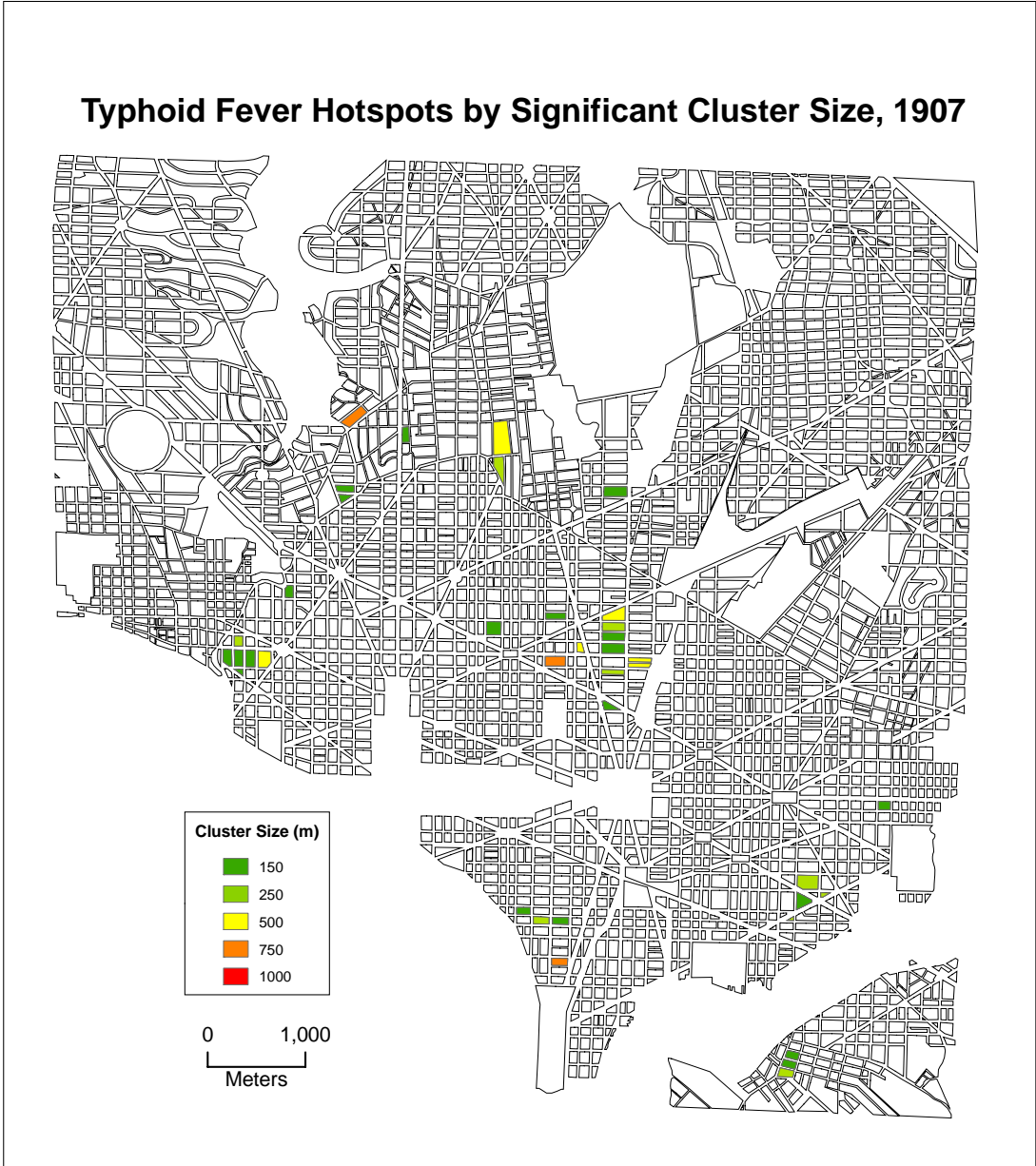
Figure 4.7 G_i* results for 1907
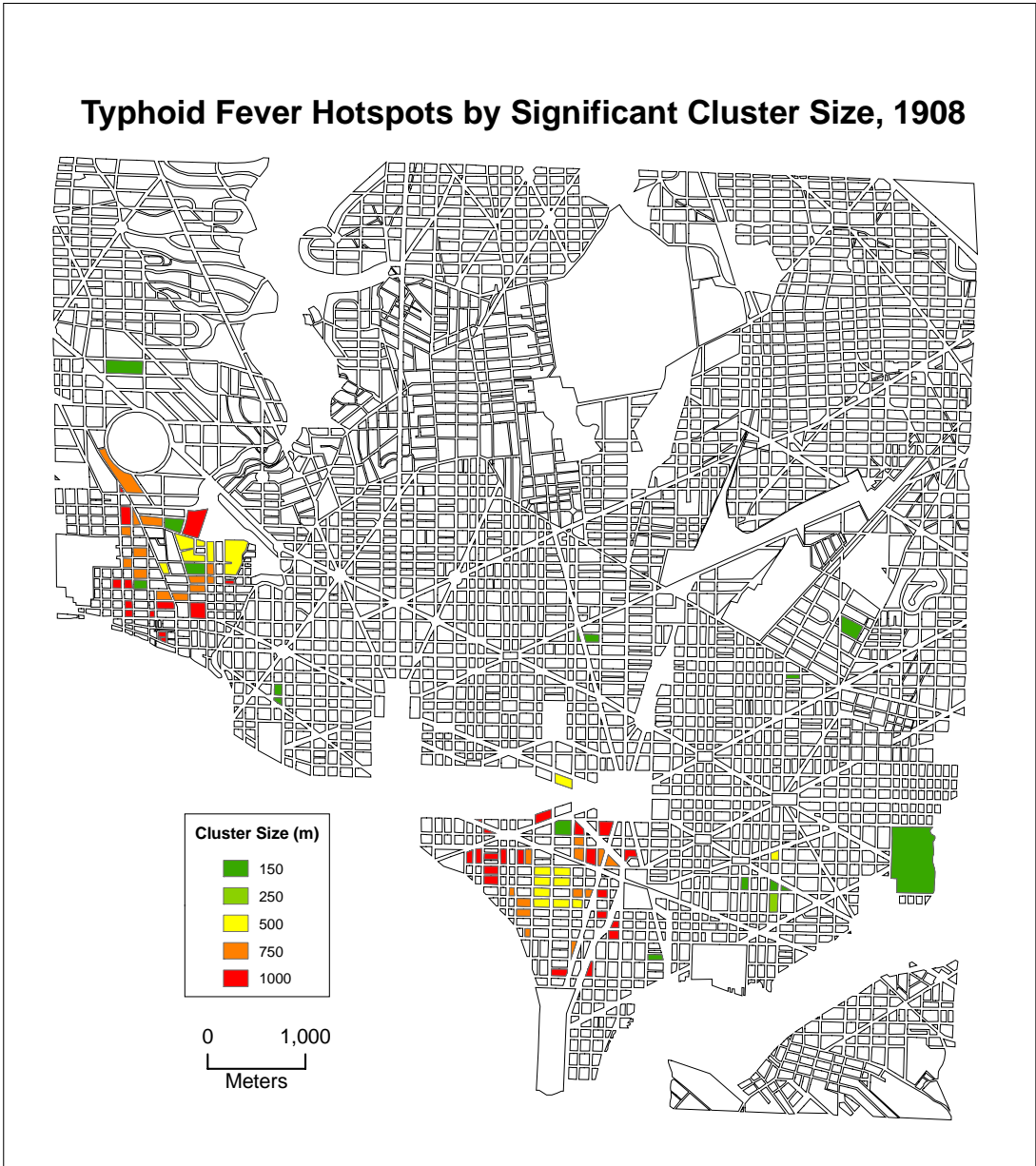
Typhoid Fever Hotspots by Significant Cluster Size, 1908

Cluster Size (m)
- 150
- 250
- 500
- 750
- 1000

0   1,000
Meters

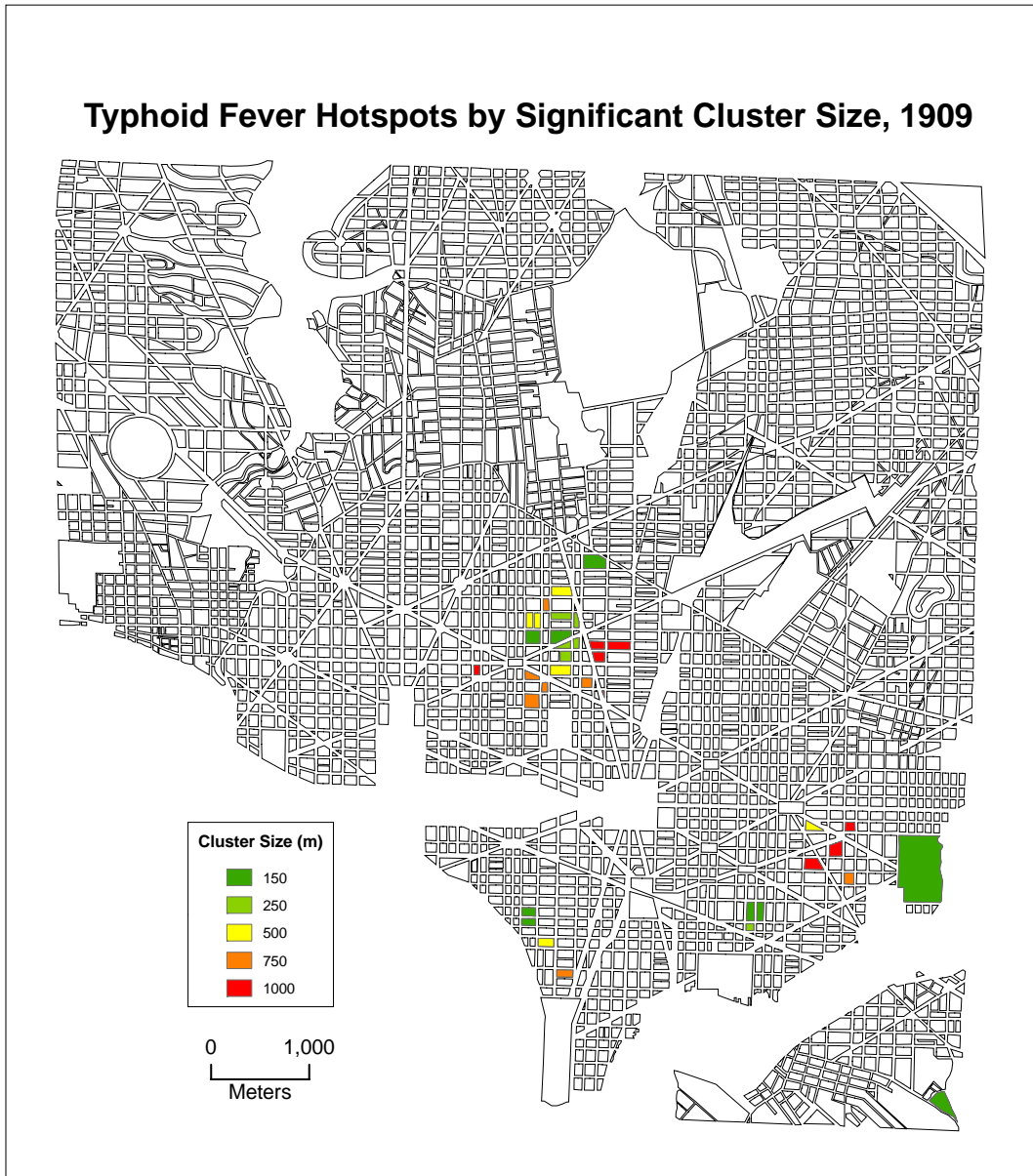**Figure 4.8** G_i* results for 1908

87

Figure 4.9 G_i* results for 1909

Temporal Results

To evaluate temporal stability in spatial hotspots between the years, in other

words to identify areas that were conducive to typhoid in all or multiple time periods, the

total number of years any block had a significant G_i* value at each distance were tallied.

All of these tallies are displayed in Table 4.4. Interestingly, at all distances most blocks

were only significant in a single year. Only seven blocks were significant at the same

distance in two years, and no cells were significant at a particular distance for more than

two years. When significant clusters were considered regardless of distance, cells were

still only significant in one year. Only two cells were significant in three years, and no

cells were significant cluster members for the duration of the study period regardless of

cluster size.

Table 4.4 Summary of temporal results, 1906-1909

Distance (m)	Number of blocks significant in years				Total
	1	2	3	4	
150	60	3	0	0	63
250	25	1	0	0	26
500	32	2	0	0	34
750	35	1	0	0	36
1000	45	0	0	0	45
Number of cells significant regardless of distance	179	13	2	0	194

Discussion

The K-function analysis identified that the original point data for each year studied here does cluster globally. This result indicates a need to identify the specific location and scale of those significant clusters. The hypothesis, that typhoid clustered locally, was supported in the results of the G_i^* statistic. The G_i^* statistic results also demonstrated that the disease clustered at multiple scales. These findings refute the conclusions drawn in all four PHS typhoid reports concerning the distribution of cases where the only mention of possible large concentrations of typhoid cases was within the 1908 report. Results of this current study indicated both widespread and localized clusters occurred in all four years (Table 4.3; Figures 4.2 through 4.9).

The second question addressed whether or not clusters varied in size during each year. Again, variation existed in the ratios of cluster sizes within and between the years (Table 4.3). The results for 1908 skewed toward clusters of 1000m suggesting that large areas of the city were infected over the course of that study period. However, 1906 and 1907, showed a more localized pattern of infection, with more clusters at a critical distance of 150m in these years. Finally, in 1909 an almost equal number of significant blocks at each distance indicated a probable lack of a single source of the disease. When assessing the question of temporal stability across the study period, we can consider each region of the city separately (Figure 1.2). Each region of the city appears to have had varied degrees of typhoid intensity within and between outbreak years generally indicating a potential lack of a consistent source of the bacteria (Tables 4.3 and 4.4).

Georgetown was impacted in 1906 and 1908, and lacked cases in 1907 and 1909. The 1908 milk related outbreak in this area was investigated by Rosenau et al. (1909) in

order to find the source of a rapid increase in typhoid cases during a month when cases had been decreasing that year. Through the investigation, the PHS officers found that only households purchasing milk from two different dairymen during September 1908 were reporting new typhoid cases. Other households on the same streets using different milk suppliers did not report new typhoid cases. The detailed investigation by the PHS narrowed the possible sources of typhoid carried into the city through milk to a single dairy in Maryland, which supplied both dairymen on a daily basis. The owner of the Maryland dairy was a typhoid carrier, having contracted the disease 18 years previously, and although not displaying symptoms since that time was still shedding bacteria. Following the discovery of a typhoid carrier at a dairy, the sale of milk from that dairy was discontinued and eight days later the last case along the routes of the two milk suppliers was reported (Rosenau et al. 1909).

Foggy Bottom was most dramatically impacted in 1907. In 1908, this area displayed two localized clusters. No clusters appeared in 1906 or 1909. Between 1900 and 1920 a small African-American residential enclave developed in this part of the city. Since more black residents contracted typhoid than white residents overall it is plausible that there would be a disproportionately high number of clusters in these neighborhoods (Rosenau et al. 1909; Groves and Muller 1975). This geographic variation associated with race is at least partly explained by poverty and that a majority of the city's alleys, unsanitary or otherwise, being home to African-Americans rather than whites (Groves and Muller 1975; Groves 1974; Borchert 1980; Groves 1973-74).

Northwest was most notably impacted in 1907 and 1909. Part of Northwest was home to the city's elite as was discussed in chapter 1. An example of how affluence

impacted disease presence is that no blocks were significant cluster members within a five-block area around Dupont Circle, the central point of this elite district. Other social groups lived in Northwest as well, including an African-American enclave and other blocks with dense alley residences to the west of Union Station (Groves 1974, 1973-74; Groves and Muller 1975; Borchert 1980). Note the reoccurrence of clustering in this general area around Union Station.

Northeast displays few significant clusters in any of the years. Initially it would seem likely for this to be related to more limited residential settlement in this area (Myers 1973-74), but results seen in chapters 3 and 7 indicate that maybe it is only post-1906 typhoid that does not cluster in this area. Only four clusters appeared in Northeast during the entire study period, two 500m clusters in 1907 adjacent to Union Station, one 150m cluster in 1906 and 1908 respectively. The smaller and more diffuse population in Northeast could help to explain the limited and localized clusters in this region, but these results are intriguing by virtue of the lack of significant clustering and will be explored in future research.

Significant clusters exist in Southeast in all four years, with the greatest number of significant cells occurring in 1906 and 1909. Only in 1909, when one block was part of a 750m cluster and three blocks were 1000m cluster members, were the significant clusters greater than 500m in size. This indicated a strong tendency towards highly localized sources until the very end of the study period in this area.

Southwest was impacted in all four years. This region of the city was known for its poverty, and high rates of disease throughout much of the city's history (Borchert 1980). The PHS report claimed that in 1908 there were more cases of unknown source in

the Southwest than in the north of the city (Rosenau et al. 1909). Figure 4.8 displays large

typhoid clusters in Southwest, while in comparison Northwest displayed only one 150m

cluster that year. The report attributed the clusters of typhoid in Southwest to poor

sanitary conditions. The relationship between the widespread distribution of typhoid in

Southwest compared to the localized clustering in Northwest in 1908 helps to support the

conclusions of Rosenau et al. that a lack of sanitation may have been at work in that part

of the city (Rosenau et al. 1909). If the conditions in Southwest were linked to poor

sanitation it is likely the disease situation would have reoccurred across the years. Indeed,

an outbreak from any other location in the city could possibly lead to secondary

outbreaks within this area through its underlying infrastructure, as long as route ways

such as worker movement, existed between the two locations. The precise determination

of what constituted poor sanitation from the perspective of the PHS doctors is unclear,

but there is good evidence from other secondary sources that Southwest housed residents

of lower socio-economic status regardless of race (Groves 1973-74). Additionally, the

largest enclave of African-American residents in the city lived in the area bounded by the

Capitol grounds to the north and extending about four blocks to the west from South

Capitol Street (Figure 4.8) (Groves 1974, 1973-74; Groves and Muller 1975). A racial

dichotomy in health care might help to explain this particular grouping of significant

typhoid clusters in largely African-American parts of the city (Moldow 1980; Green

1963). Nonetheless, a more detailed analysis of the social and environmental geography

of the city at the time would be needed to accurately address this matter.

The suburbs to the north of the city displayed significant typhoid clusters in 1906

to 1908, but none in 1909. The significant clusters varied in size and diminished in

number over time. Two aspects of these clustering patterns are interesing. First, the group of 11 significant blocks near Florida Avenue in 1906 coincide with a black residential enclave identified by Groves and Muller (1975). Second, George Kober, in his 1895 typhoid report (chapter 3) mentioned that typhoid appearred quite often in the suburban areas primarily because of the less sanitary conditions in that region. The "less sanitary conditions" were the result of the comprehensive water supply and sewerage systems not having been extended to those areas in 1895. With these two aspects of suburban typhoid clusters at hand, an African-American enclave and a lack of sanitary infrastructure, a possible explanation is found for the decrease of typhoid clustering in the suburbs between 1906 and 1909. First, as has already been mentioned the black population in Washington, tended to comprise of lower income groups. One reason, not mentioned in the literature about Washington, for the enclave's development on the edge of the city proper could have been housing cost. As housing prices in the city rose, those who could not afford them sought cheaper housing on the periphery particularly before public transportation expanded to those parts of the city (Harris 1996; Kellogg 1977). So, at the end of the nineteenth century it is likely that a poorer population inhabited the outskirts of the city using well water and box privy/cesspools. Maybe, then between 1906 and 1909 the water distribution system was extended to these suburban regions and brought the suburban areas up to the same sanitary standards as the rest of the city. Green (1963) does refer to the extensions of these systems during the 1890s and early 1900s, and in the 1920s the extentions of the sewerage system into the part of Maryland just north of these areas. So it is conceivable that city services were just reaching the near suburbs in the first decade of the twentieth century so that by 1909 the services were in place and

94

typhoid eliminated from the outlying areas. Again, this is just speculation, but it is a plausible conclusion based upon the clustering results at hand. Future research using archival resources and newpapers from Washington, D.C. might help to shed light on this matter, but for the time being it is beyond the scope of study.

The general lack of consistently reoccurring clusters between outbreak years indicates a lack of stability in disease origin as indicated by Figures 4.6, 4.7, 4.8, and 4.9 and Table 4.4. These shifting areas of disease intensity are suggestive of a lack of a common and universal source of infection. Between the visual instability of typhoid clusters in Figures 4.6, 4.7, 4.8, and 4.9 and a general understanding that by 1906 most of Washington, D.C. had the expected levels of sanitary infrastructure to prevent typhoid fever, it is logical that the disease would not remain spatially stable between years. Instead, it seems that the patterns produced here indicate that typhoid affected different communities and possibly households, indicating a need to better understand both neighborhood dynamics, and more importantly the specific practices of each individual household in order to understand the distribution of typhoid. Of course such "shoe-leather epidemiology" is now impossible, though one could state that neighborhood level interconnections were more likely to be causative than any overall city-level sanitary conditions. This conclusion was hinted at in the 1909 report as even though the disease surface was described as being uniform, it was suggested that typhoid persisted in the District of Columbia through the influences of multiple causes (Lumsden and Anderson 1911).

All four typhoid reports (1906 – 1909) attributed between 30 and 50 percent of the cases to contaminated milk, contact with a person carrying the disease, or to

contracting the disease outside of Washington, D.C. The source of infection for the remaining cases was unknown. Given the "general distribution" of the disease as described by the PHS, the water supply was suspected as the source of infection by the investigators (Rosenau et al. 1907). This hypothesized source of typhoid was particularly favored since over 90 percent of those individuals who contracted typhoid in any of the years studied regularly drank un-boiled tap water. With this idea in mind, and as part of their mandate, the PHS tested water related to the city's supply at all stages from the Potomac River, through the various settling basins and filters, to household taps. The results of these bacteriological investigations were inconclusive in 1906 and during the remaining years continued to be considered "free" of contaminants (Lumsden and Anderson 1911).

Although the overall finding of the three reports was that the city's water supply was not the source of infection for cases of unknown cause, a few more specific statements about the distribution of typhoid were made in 1906, 1908, and 1909. The 1906 report included a geographic study of shallow and deep well locations compared to the location of typhoid cases of unknown cause. The conclusion of this particular study was that there was no unusual concentration of cases around well locations. The condition and water quality in wells were not included in the subsequent reports (Rosenau et al. 1907, 1908).

While all four reports described a fairly general distribution of typhoid, the 1908 report made direct reference to two specific concentrations of typhoid cases, the milk related outbreak in Georgetown and the group of cases of unknown cause in Southwest, both discussed above (Rosenau et al. 1909). In 1909, the report summarized the findings

from all four years and discussed the possible associations between the disease and a contaminated water supply (Lumsden and Anderson 1911). The PHS remained reluctant to state that the water supply was the source of the continued presence of typhoid since not only was the disease nearly uniformly distributed, but this distribution included households not hooked up to the city water system. Additionally, in 1909 the PHS investigators mapped the distribution of diphtheria and scarlet fever, two diseases not associated with water, and found a general and uniform distribution of these diseases that resembled the distribution of typhoid. Given these three disease distributions, the authors believed that some mechanism besides water was at work in distributing typhoid around the city, this comparison will be investigated further in chapter 7 (Lumsden and Anderson 1911). It should be remembered, however, that all of these distributions were assessed visually between 1906 and 1909, and were not the result of any spatial analytical approach.

The spatial analysis of the datasets contained in the PHS reports for the years 1906 to 1909 should only be seen as the first step in investigating the disease surfaces of early twentieth century Washington, D.C. Further temporal precision can be built into the cluster analysis. In addition, other spatial layers, such as urban and social structure, can be overlaid in a search for a neighborhood association with cluster location.

Conclusions

The methodology applied here was useful for evaluating the spatial distribution and inter-annual patterns of typhoid outbreaks in Washington, D.C. from 1906 until 1909. Ripley's K-function identified global clustering in the typhoid case data, in contrast to the findings of the original PHS reports. Recent advances in local spatial auto-correlation

techniques allowed this study to go beyond a global investigation and explore the possibility of local clusters which are hard to distinguish within the complete spatial distribution of cases. The Getis and Ord statistic indicates that clustering occurs at multiple spatial scales which refutes the original PHS conclusions that typhoid's distribution was evenly distributed. While analyses of historical data sets must be interpreted with caution, this study does suggest that there is utility in these types of analyses, and provides new insights into the urban patterns of a series of typhoid outbreaks. Further the next chapter will build upon this foundation and explore the stability of clusters within years. The current chapter and its confirmation of localized clusters was the first step towards that end.

CHAPTER 5
TEMPORAL ANALYSIS OF TYPHOID IN WASHINGTON, D.C., 1906-1909

Introduction

The chronology of cases within an epidemic is an important consideration when studying most epidemics. Many diseases display a relationship between season and different epidemic stages, such as initiating, peak and cessation (Wills 1996; Edling and Liljeros 2004; Cromley and McLafferty 2002). In addition, the interconnection between time and space among the cases may reveal both spreading mechanisms and barriers within the epidemic. A number of disease diffusion studies have used historical data to investigate the movement of a disease throughout an epidemic (Cliff et al. 1981, 1986; Smallman-Raynor and Cliff 2001; Patterson 1994; Pyle 1969; Curson 1985). Temporal complexity is added to a disease study when considering the spread of a disease from the index case to other susceptible members of a population. It is spatially naïve to believe that a contagious disease spreads from the index case in even waves radiating outwards in every direction. Obviously the underlying environmental, social and urban structure of an afflicted location will both aide and hinder the diffusion of disease. A single house, an ethnic enclave, a large road, could all impact the spread. These types of variations will not only appear as spatial patterns, but also as temporal signatures. Epidemic data at a fine spatial and temporal resolution can help develop models of this complex surface (Watts et al. 2005).

Typhoid fever, like many diseases, typically followed a seasonal pattern throughout the nineteenth and early twentieth centuries (Whipple 1908; Parker et al. 1907; Kober 1895). The annual cycle began with low rates of the disease persisting over

the winter months, typhoid increasing in prevalence in the late spring and early summer, with a peak in cases and deaths in the late summer or early fall, and finally a fairly rapid decline to "baseline" levels over the course of the fall. This cycle of typhoid fever prevalence was not only noted in Washington, D.C., but in other U.S. cities during the time period in question (Whipple 1908, Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911; Parker et al. 1907). Other diseases display seasonal patterns, often with a summer peak in prevalence, including cholera, diarrheal diseases, yellow fever, malaria, and while not a disease, infant mortality (Patterson 1992, 1994; Mooney 1994; Wills 1996). Many of these diseases were either interrelated or associated with one another in the late nineteenth century. For example, much of the increase in infant mortality was related to the increase in diarrheal disease or other ailments associated with contaminated water or contact between an ill individual and a child's caretaker who might not have maintained the most stringent of sanitary practices (Mooney 1994). Understanding this seasonality was one step towards understanding the epidemiology of disease in the late nineteenth and early twentieth centuries, and was one of the means that cities used to determine the correct mitigation strategies for different diseases (Whipple 1908; Wills 1996).

Other gastro-intestinal disorders besides typhoid also displayed a summer peak in prevalence. Cholera provides a good example of a disease that tended to peak during the warmest summer months in the years when it swept through Europe and North America (Patterson 1994; Wills 1996). For both typhoid and cholera, the similar means of transmission through water helps to explain a similar tendency to peak during warm periods of time. The bacteria of both diseases survive in greater quantities in water

systems when it is warm (Parker et al. 1907; Wills 1996; Sedgwick 1893). Also, as water evaporates from river systems and other water sources during warm months, contaminants are then concentrated in the water supply (Parker et al. 1907). Finally, people tend to alter their water drinking habits when temperatures are high, by not boiling potentially unclean water before consuming it people increased their risk of infection (Wills 1996; Sedgwick 1893). All three of these summer-related factors contribute to elevated cholera and typhoid prevalence.

Of somewhat more interest in a United States Geological Survey report were the comments concerning the connection between low river flow and increases in typhoid fever cases (Parker et al. 1907). The relationship, which initially was seen as counter-intuitive to the authors, was one of increased bacterial concentration in the water column at times of low water in the Potomac. Thus, if a person were to drink river water under these low volume conditions the individual would consume more typhoid bacteria (and other contaminants) per cup than if the river flow were higher (Parker et al. 1907). Generally, the contamination load remained reasonably constant in the Potomac and other rivers, while the amount of water flowing through the river channel fluctuated (Parker et al. 1907; Wills 1996).

Parker et al. (1907) were careful to comment that not all periods of low river flow in the Potomac resulted in typhoid outbreaks, but that most typhoid outbreaks did coincide with low river flow. This, as discussed in their report, was an indication that typhoid fever still needed to be introduced into the river, which if at a time of low water flow could then result in a visible increase in typhoid fever cases and/or deaths (Parker et al. 1907). If both these events happened, then "the excreta of a single typhoid patient

[that] washed into a stream [and] caused over a thousand cases" of typhoid (Kober 1895, 259; Sedgwick 1893).

Once introduced into a river, if there was not enough water to thoroughly dilute the bacteria before reaching the water-supply intake pipe of a downstream city, then typhoid could spread quickly and widely through a community, as happened in the case the 1890-91 typhoid epidemic in Lowell, Massachusetts (Sedgwick 1893). Following the start of the epidemic in Lowell, a typhoid epidemic developed in the downstream city of Lawrence. The investigation of both epidemics indicated that typhoid had not entered Lowell or Lawrence until a typhoid sufferer in North Chelmsford used a privy overhanging a tributary stream to the river supplying drinking water to both cities, initially causing an outbreak in Lowell. This in turn resulted in the Lawrence epidemic as untreated sewage was released into the Merrimac River eventually reaching the down-stream city's drinking-water distribution system (Sedgwick 1893).

The original Public Health Service reports indicated a uniform distribution of typhoid cases for all four years studied, with the exception of two "clusters" identified in 1908. The results of the previous chapter demonstrate that typhoid did cluster in particular parts of the city, and that those clusters varied in their spatial scale, possibly indicative of more than one mechanism of disease spread. This calls into question whether typhoid truly clustered within an outbreak, or whether spurious clusters (actual high intensities of cases without interconnection) were revealed when looking at the outbreak as a whole? In order to better understand the geography of typhoid in early twentieth century Washington, D.C., the temporal structure of the clusters identified in chapter 4 were further investigated.

The temporal elements of typhoid in Washington, D.C. can be investigated in numerous ways. One approach is to explore and compare clustering patterns of the disease for finer temporal periods. A second approach is to investigate the seasonal patterns of typhoid within the entire outbreak. The first of these temporal considerations is more exploratory in nature, with the intention being to better understand the space and time complexities within the epidemic rather than simply producing the traditional outbreak-wide epidemic curve (Watts et al. 2005). An investigation of the seasonality of the epidemic allows for further comparison with other similar outbreaks, particularly the confirmation of a late summer peak in cases suggested by the PHS reports and other contemporary sources (Rosenau et al. 1907, 1908, 1909, Lumsden and Anderson 1911; Whipple 1908; Parker et al. 1907). The temporal scale at which the typhoid case data were collected by the PHS allow the use of a GIS and other spatial analysis techniques to better understand each outbreak year and to more closely consider the spatial aspects of the disease's seasonality.

The United States Public Health and Marine-Hospital Service typhoid reports for 1906 through 1909 include case data displaying both the location of cases and a temporal recording (first sign of disease onset), at two-week (1906-1908) or one-month (1909) intervals. In order to consider the interrelationship of space and time on typhoid two questions are explored here. First, did typhoid cluster at finer temporal resolutions, in other words were there mini epidemics within the annual summer outbreak. This question will be addressed by breaking the data into two 3-month time periods for each year (May to July; August to October), as there were a sufficient number of events at that temporal scale to identify true typhoid clusters using the G_i^* statistic. More specifically, this

temporal break will explore the question: did typhoid cluster during each half of the outbreaks, and if so, where were those clusters? A second question addresses the seasonality of typhoid by asking: for each of the years, was there an identifiable seasonality to the key stages of the epidemic curve?

Methods

The results in chapter 4 illustrated that typhoid clustered globally in all of the four years reported by the PHS, and that typhoid cluster locally within that global clustering. The current chapter temporally subdivides the data used in chapter 4 in order to better understand the temporal structure of the previously identified local clusters. While the 3-month analysis uses "different" data than the previous study, the specific interest in the subdivided data is intended to further explore the local clusters already identified using each full outbreak's worth of cases, and therefore Ripley's K-function is omitted from the temporal analysis.

As the purpose of the study was to explore the patterns of typhoid within each outbreak and to determine whether or not the disease clustered within smaller spans of time than the whole outbreak, further data aggregation was required. The G_i* statistic is not a test of presence or absence, therefore if a temporal unit were to have only blocks with either one case or no cases, the statistical assumptions would be violated and the test made invalid (Getis and Ord 1992; Ord and Getis 1995). When using the month-level data collected for each year, less than half of the months could be validly tested, thus not providing the comprehensive results sought for this study. To avoid this problem of small numbers with the G_i* statistic the data were temporally aggregated to 3-month (May through July; August through October) groups and the outbreak years analyzed as halves.

The aggregation was achieved by summating the number of cases that occurred in the months of May through July and August through October. In order to prevent any spurious findings based on the artificially defined break between July and August, the G_i^* statistic was run again on the typhoid cases that developed in July and August of each year with the intension of capturing that intermediate part of the spatial pattern of the disease across the city.

The methodology described in chapter 2 outlines the study design here. This includes aggregating the typhoid cases to the block-based grid surface and testing the five cluster distances of 150, 250, 500, 750, 1000 meters.

In addition to using the G_i^* statistic to identify where local clusters existed during both halves of each year, the number of cases in each of eight regions of the city identified in chapter 1 (Figure 1.2) were graphed as histograms of cases per month. This approach can be used to answer the second question posed, addressing the possible distinctive seasonality to the geography of typhoid in Washington, D.C. The eight regions of the city were digitized as polygons and used as "bins" for the histograms. The number of cases in each region per month were summarized with the Polygon in Polygon Analysis Tool available in the Hawth's Analysis Tools extension for ArcGIS 9.0 (spatialecology.com). To standardize the visual representation of these case numbers per region in the histograms created, each count of cases is presented as the percentage of the total regional cases per month.

Results

The results of the 3-month level cluster analysis using the G_i^* are summarized in Figures 5.1 through 5.8 and in Table 5.1. With the exception of the August to October

time period in both 1908 and 1909, all time periods display more clusters at the 150m distance than any other distance utilized in the analysis.

3-Month G_i* - 1906

The results of the analysis for 1906 are summarized in Figures 5.1 and 5.2 and Table 5.1. During the May to July time period most of the clusters in Georgetown were larger, with critical distances of 500m or greater. The significant clusters outside of Georgetown in Southwest and Northwest were 250m or smaller with one exception. During the August to September time period a geographical shift occurred in the location of the significant clusters; only one cluster appeared in Georgetown and none in Southwest. Instead, most clusters were in Northwest and in the suburbs adjacent to Northwest. Only three clusters were 750m in size or greater in the entire city during this time period, and of these the only 1000m cluster was located in Northwest (Figures 5.1, 5.2; Table 5.1).

3-Month G_i* - 1907

Figures 5.3 and 5.4 along with Table 5.1 summarize the G_i* results for 1907. During the May to July time period the clusters in Northeast and Southeast remain quite close to East Capitol Street extending approximately seven blocks to the east from the Capitol Building. All of the 1000m clusters were in the group of significant blocks in Northeast and Southeast, with the exception of one 1000m cluster in Anacostia (Figures 5.3, 5.4; Table 5.1). None of the significant clusters from the first time period in this year remained significant in the August to September time period. The two significant blocks in Foggy Bottom during the second time period neighbored significant blocks from earlier in the year. Most blocks were significant at smaller distance values during the

106

August to September time period, with only two 1000m clusters both of which were in

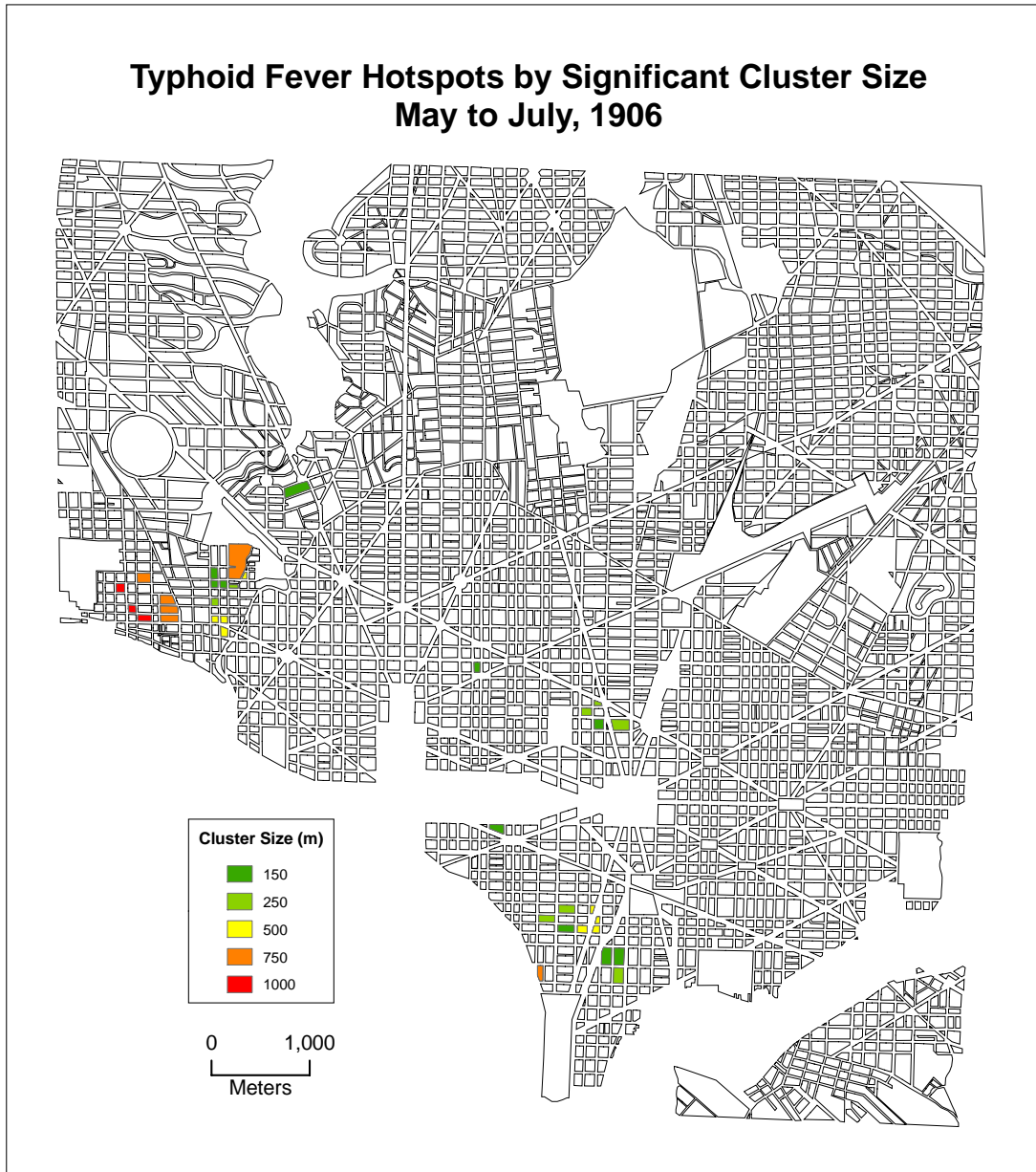Northwest (Figures 5.3, 5.4; Table 5.1).

Figure 5.1 G_i* results for May through July, 1906

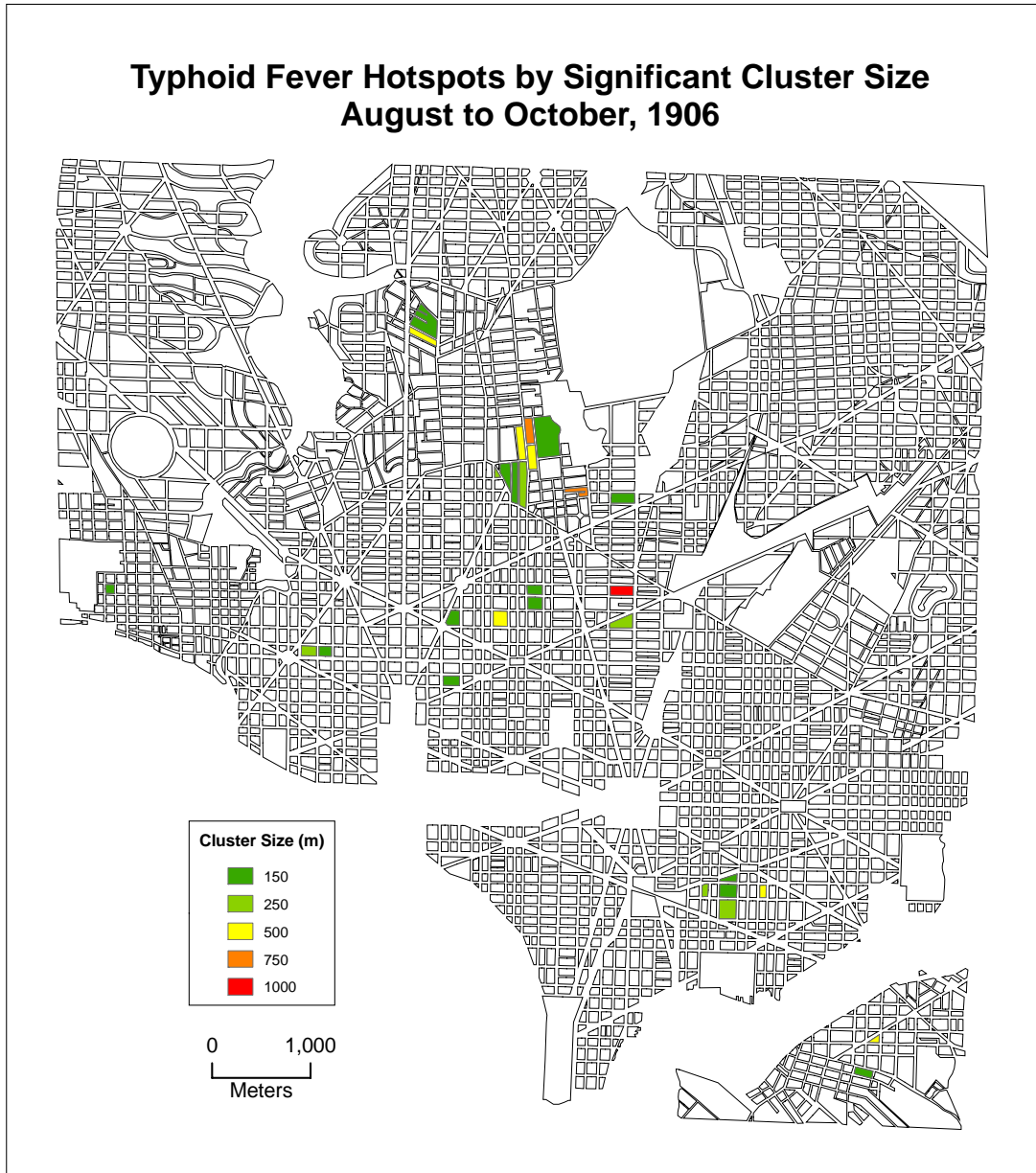**Typhoid Fever Hotspots by Significant Cluster Size**
August to October, 1906

Cluster Size (m)

150
250
500
750
1000

0   1,000
Meters

Figure 5.2 *G$_i$** results for August through October, 1906

Table 5.1 Summary of temporal Gi* results, 1906-1909

Year	Distance (m)	Cells	min G_i^*	max G_i^*
May to July - 1906				
	150	10	2.92	4.58
	250	8	2.10	3.13
	500	8	2.12	3.00
	750	6	2.13	2.31
	1000	3	2.06	2.26
August to October - 1906				
	150	14	2.75	4.36
	250	6	2.29	3.90
	500	6	2.13	2.92
	750	2	2.08	2.21
	1000	1	2.11	2.11
May to July - 1907				
	150	11	2.24	4.81
	250	1	3.18	3.18
	500	2	2.42	2.75
	750	5	2.28	3.39
	1000	6	2.12	2.74
August to October - 1907				
	150	9	2.75	6.15
	250	6	2.03	4.74
	500	6	2.20	4.77
	750	7	2.22	2.77
	1000	2	2.02	2.80

Table 5.1 (continued)

Year	Distance (m)	Cells	min G_i^*	max G_i^*
May to July - 1908				
	150	17	2.35	5.09
	250	2	2.50	3.33
	500	4	2.28	3.17
	750	7	2.47	4.74
	1000	8	2.01	4.74
August to October - 1908				
	150	6	2.88	6.24
	250	2	2.41	3.05
	500	15	2.08	4.74
	750	18	2.04	4.53
	1000	7	3.38	4.58
May to July - 1909				
	150	10	2.44	5.41
	250	0	0.00	0.00
	500	1	3.05	3.05
	750	4	2.11	3.11
	1000	4	2.05	2.84
August to October - 1909				
	150	1	5.59	5.59
	250	4	3.67	4.40
	500	3	3.86	4.51
	750	4	2.80	4.53
	1000	4	2.36	3.86

**Typhoid Fever Hotspots by Significant Cluster Size
May to July, 1907**

Cluster Size (m)
150
250
500
750
1000

0   1,000
Meters
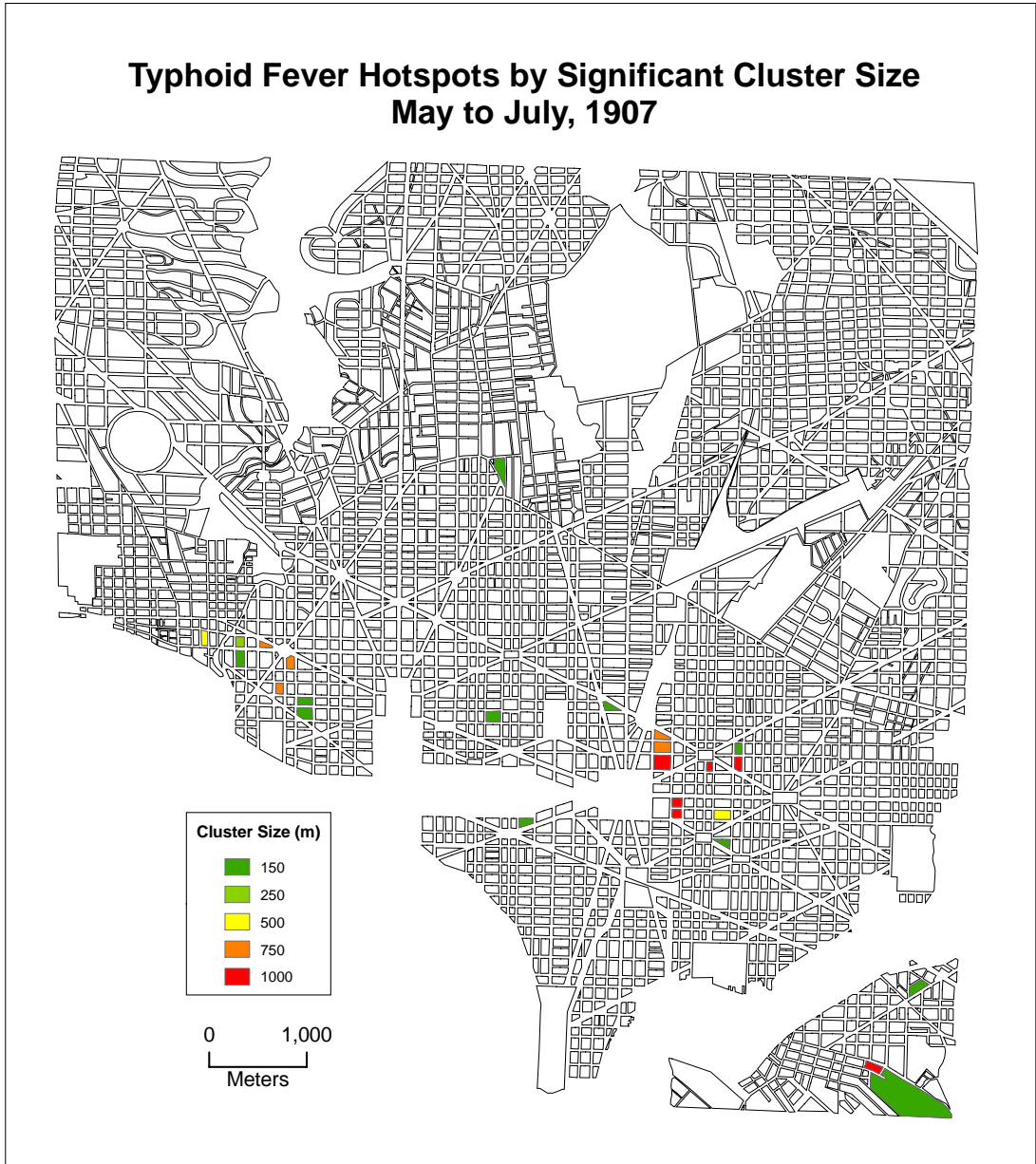
Figure 5.3 G_i* results for May through July, 1907
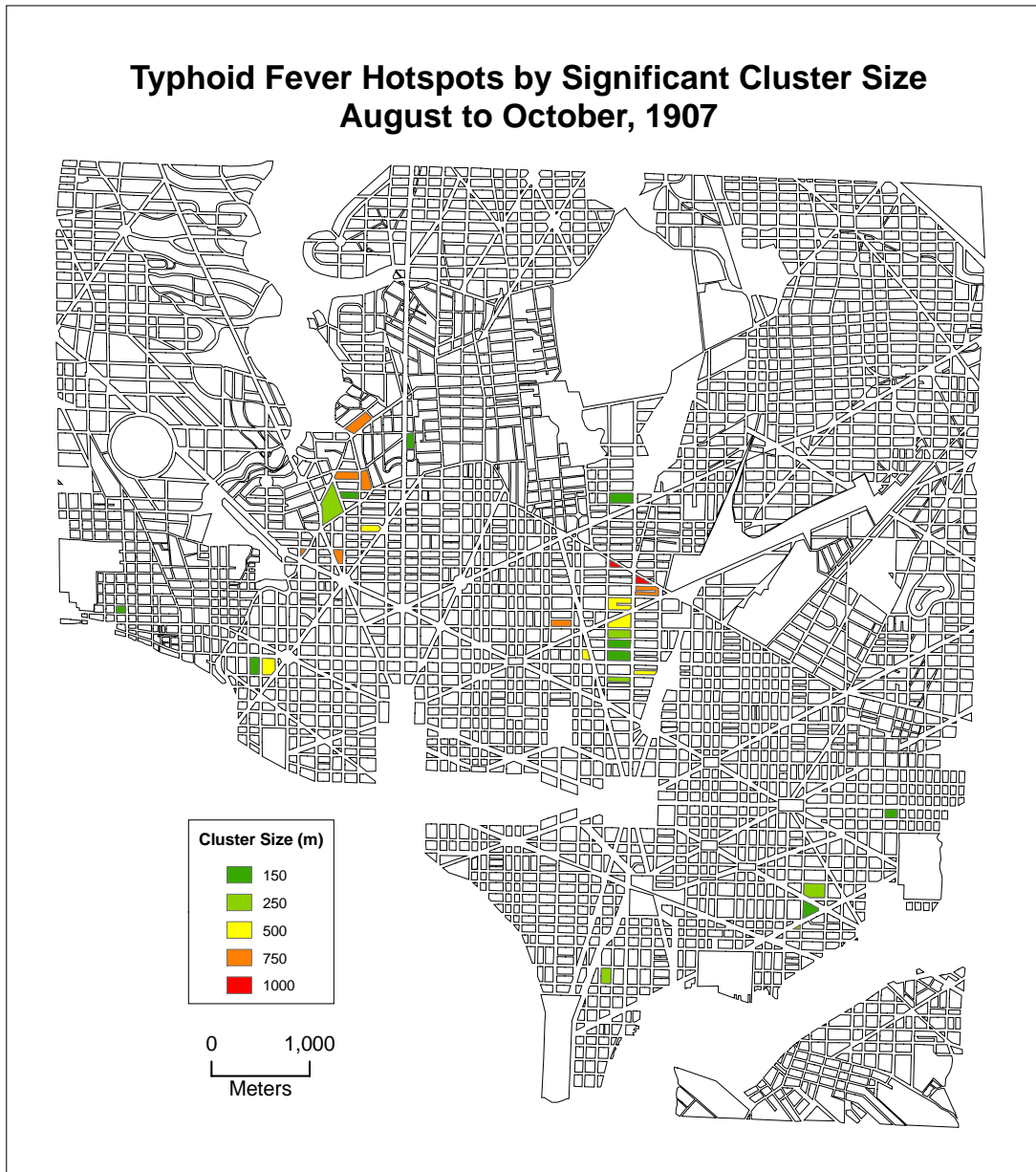
Figure 5.4 G_i* results for August through October, 1907

3-Month G_i* - 1908

The results of the G_i* analysis for 1908 are summarized in Figures 5.5 and 5.6 and in Table 5.1. Most of the significant clusters appear in Southwest during the May to July time period in 1908. Of all of the significant blocks most were smaller in size

(150m). Clusters larger than 150m only appear in Southwest and in the Suburbs (Figure 5.5). The geographic location of the most significant blocks shifts in the August to September time period. The largest group of significant blocks is in Georgetown, spilling into the Suburbs north of that region. The group of significant clusters in Southwest contracted and was primarily composed of 500m clusters during the second study period. For all of the 3-month level analyses the highest G_i^* value was 6.24, and this value occurred in the one 150m cluster in Georgetown during the August to September time period in 1908. Additionally, most of the significant blocks were either part of 500m clusters or 750m clusters during the second half of the 1908 study period. 1908 is the only year in which almost no clusters appear in Northwest during either of the study periods (Figure 5.6). There were two 150m clusters in the May to July time period, but these occur further to the south and west than the significant blocks in this region during the other three years (Figure 5.5 and Table 5.1).

3-Month G_i^* - 1909

The results of the G_i^* analysis for 1909 are summarized in Figures 5.7 and 5.8 and in Table 5.1. There were few significant clusters at any scale during either time period in 1909. During the first time period, most of the significant blocks were 150m clusters, and most of these blocks were located in Southwest. All of the clusters outside of Southwest were at the 150m distance. During the August to September time period all of the significant blocks were located in Northwest. Only one of the clusters during this time period was at the 150m range with the remaining distances appearing in approximately equal numbers of significant blocks. The highest G_i^* across both time

113

periods in 1909 was 5.59 and this was the G_i* value of the single 150m cluster in

Northwest, indicating a rather strong localized cluster (Figure 5.8 and Table 5.1).
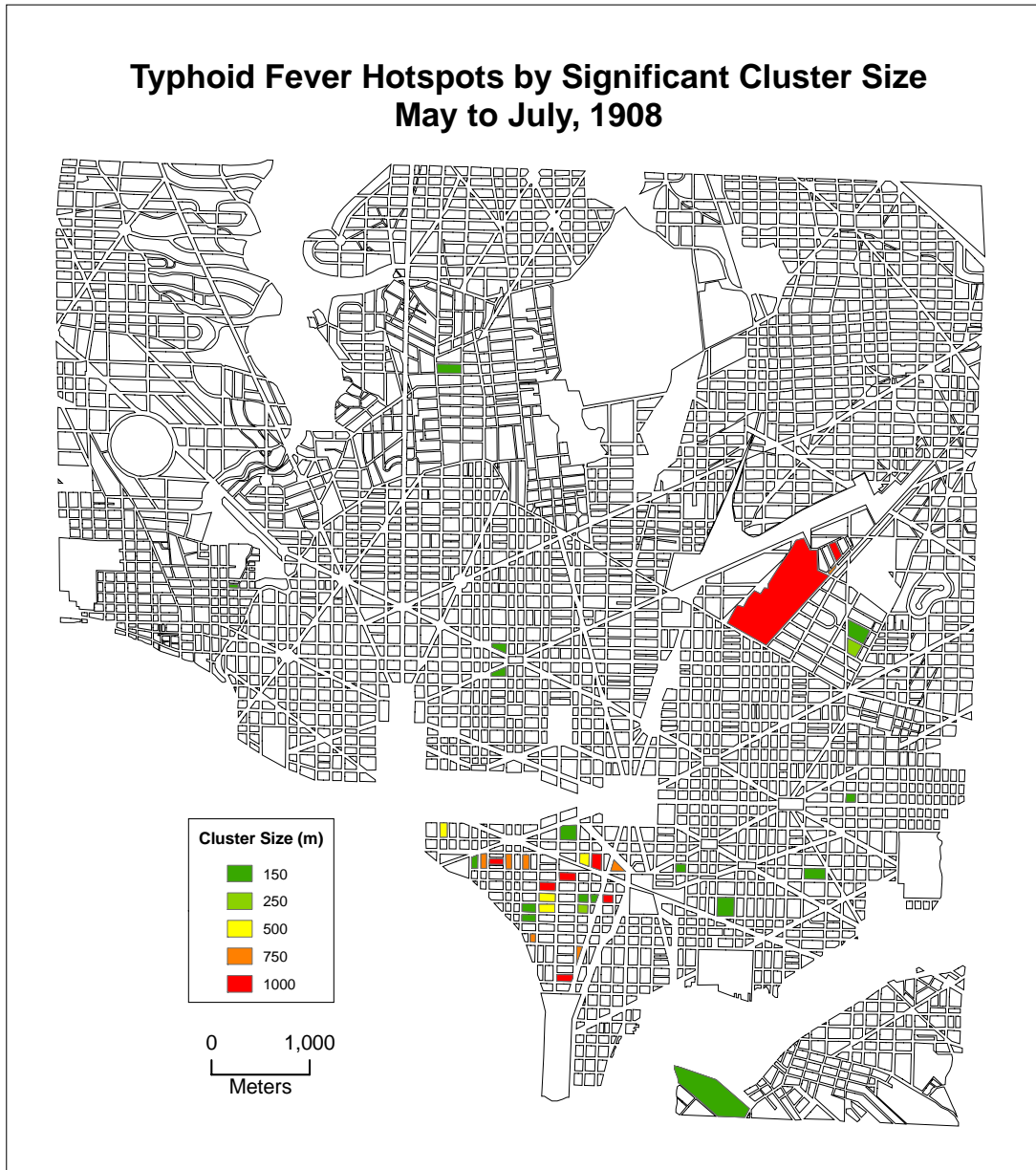
**Typhoid Fever Hotspots by Significant Cluster Size
May to July, 1908**

Cluster Size (m)
150
250
500
750
1000

0   1,000

Meters

Figure 5.5 G_i* results for May through July, 1908

114

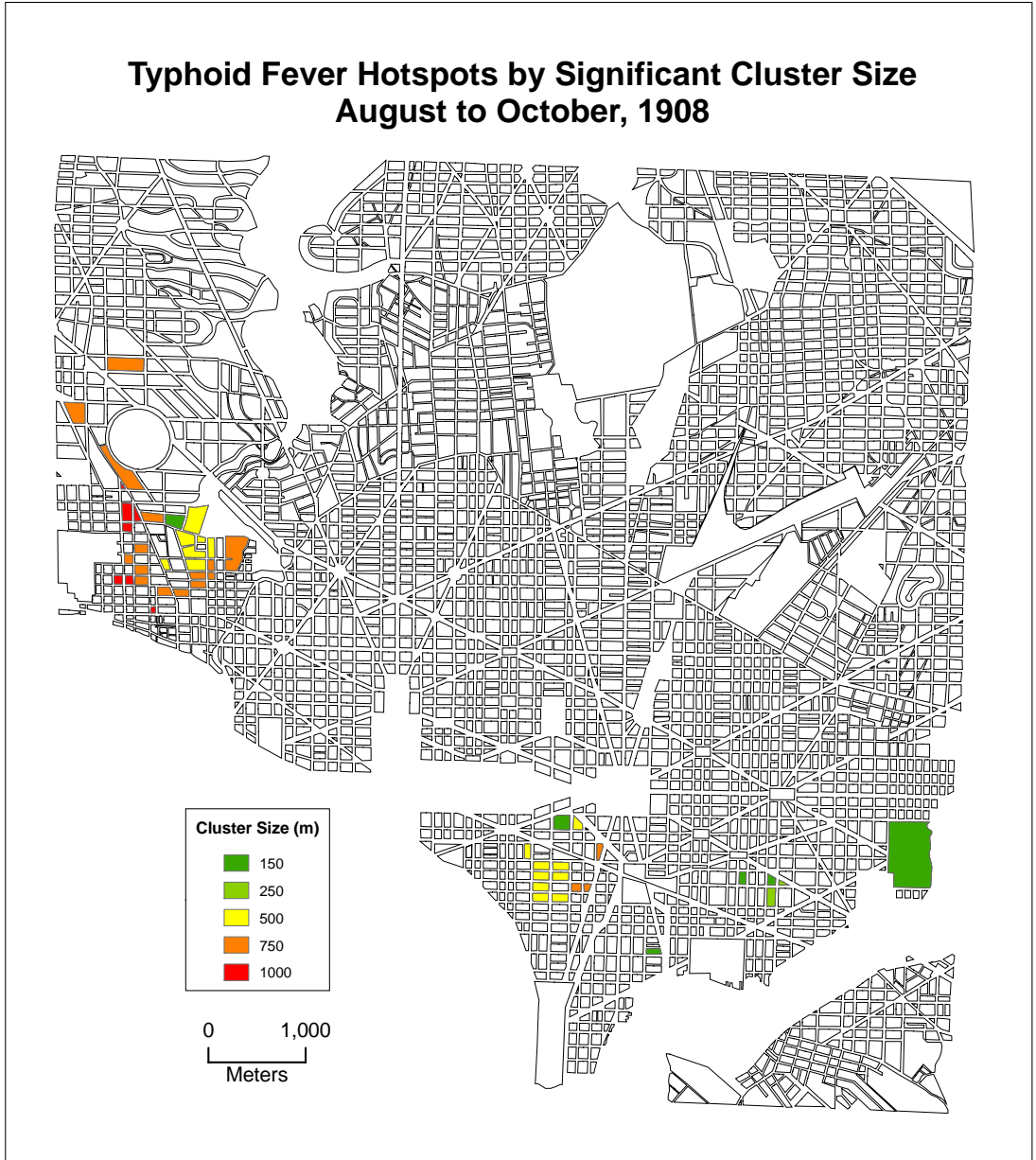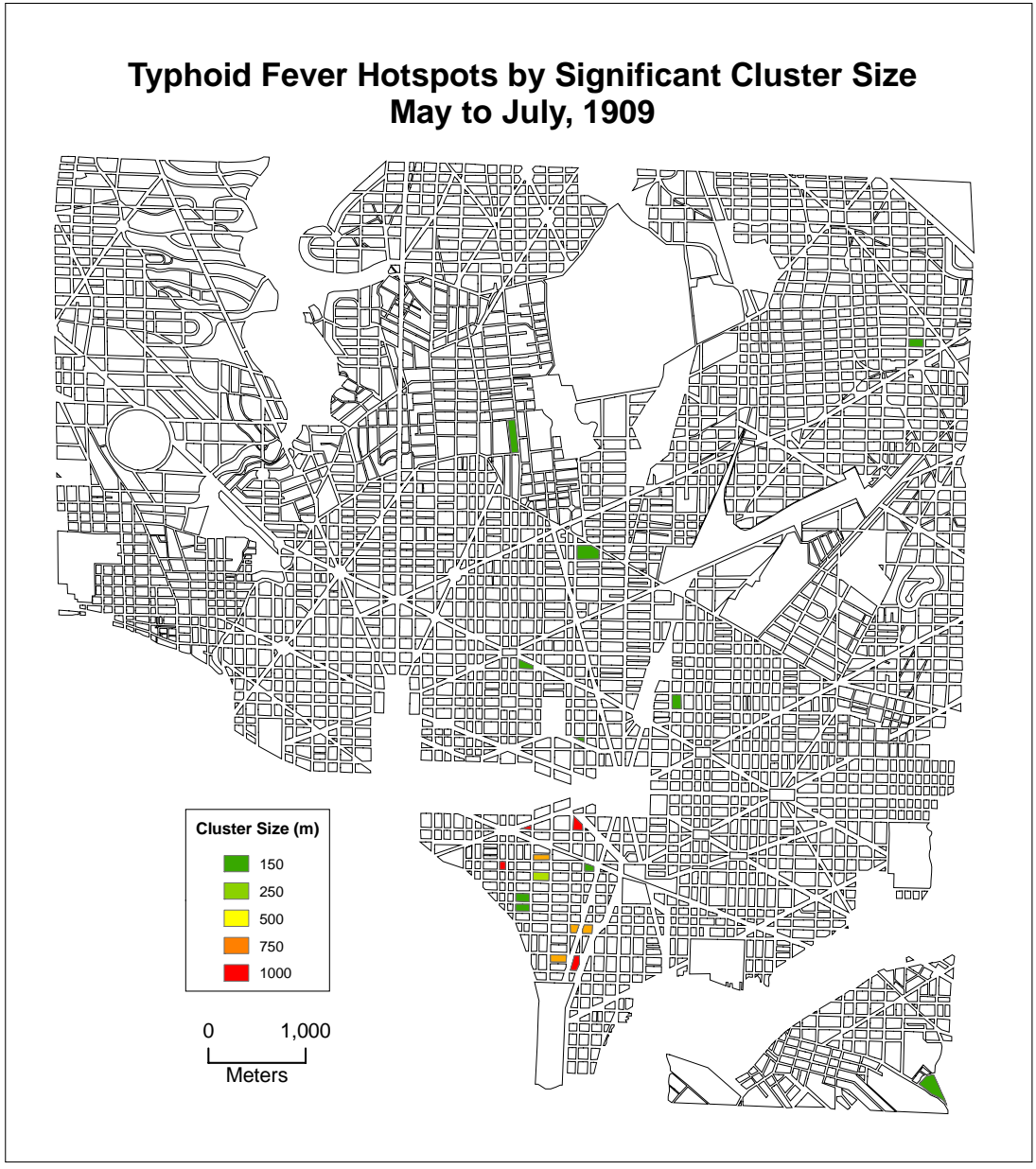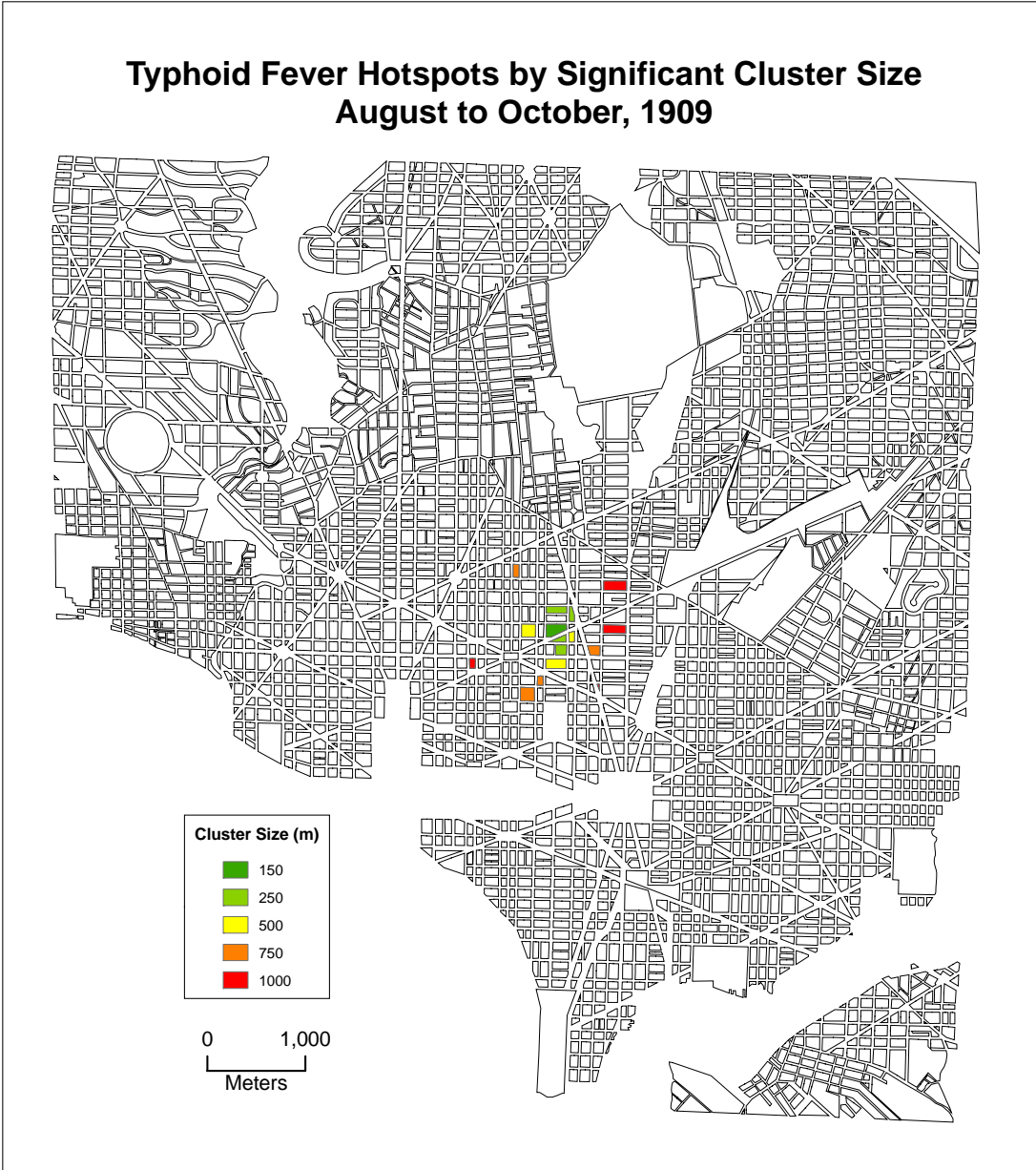**Typhoid Fever Hotspots by Significant Cluster Size
August to October, 1908**

Cluster Size (m)

- 150
- 250
- 500
- 750
- 1000

0   1,000

Meters

Figure 5.6 *G_i** results for August through October, 1908

Figure 5.7 *G$_i$** results for May through July, 1909

**Typhoid Fever Hotspots by Significant Cluster Size
August to October, 1909**

Cluster Size (m)
- 150
- 250
- 500
- 750
- 1000

0   1,000
Meters

Figure 5.8 G_i* results for August through October, 1909

Focusing on the clusters just in July and August of 1906 displays a number of small clusters scattered across Northwest, Northeast, Southwest, and Southeast (Figure 5.9). In Georgetown, four groups of blocks that neighbor one another display significant clusters of sizes ranging from 150m to 750m (Table 5.2).
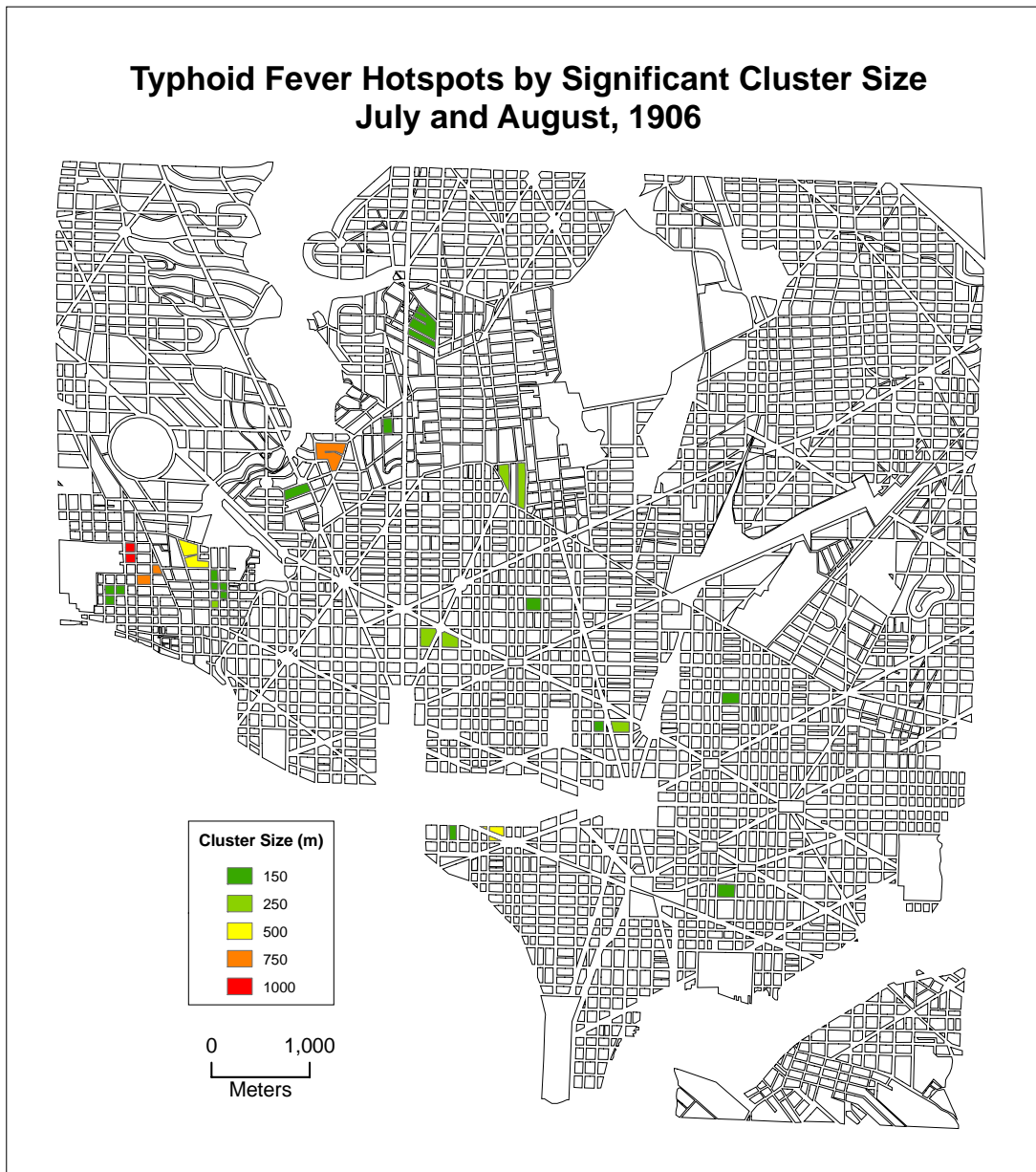
Figure 5.9 *G$_i$** results for July and August 1906

Table 5.2 Summary of July and August G_i* results, 1906-1909

Year	Distance (m)	Cells	min G_i*	max G_i*
1906				
	150	16	2.42	4.02
	250	6	2.55	2.72
	500	4	2.05	2.27
	750	3	2.02	2.09
	1000	2	2.03	2.09
1907				
	150	9	2.96	5.37
	250	6	2.27	3.46
	500	3	2.23	3.02
	750	2	2.75	3.68
	1000	1	2.47	2.47
1908				
	150	11	2.00	5.01
	250	6	2.00	3.51
	500	12	2.00	3.43
	750	8	2.17	3.44
	1000	8	2.02	2.73
1909				
	150	6	2.90	2.90
	250	1	3.18	3.18
	500	2	2.18	3.66
	750	4	2.09	2.86
	1000	1	2.09	2.09

July and August G_i* - 1907

The July and August analysis for 1907 is summarized in Figure 5.10 and Table 5.2. Clusters of varying sizes are scattered across the city and in only three places are neighboring blocks members of clusters, in Foggy Bottom, Southwest, and Southeast. Of these groups of clusters, only the group in Foggy Bottom was made of more than two blocks and a block with a cluster size greater than 250m. This single larger cluster was at the 500m distance (Figure 5.10). The remainder of the significant blocks during these months were mostly in Northwest and ranged in their significant cluster size across the entire spectrum tested.

July and August G_i* - 1908

Table 5.2 and Figure 5.11 summarize the results of the 1908 G_i*analysis for July and August. In 1908, the mid-summer clusters of typhoid are focused in the southern part of the city; in both Southwest and Southeast (Figure 5.11). While cluster sizes range from 150m to 1000m, most of the clusters during these two months are skewed towards being 500m in size or greater (Figure 5.11). Only eight significant clusters are scattered through the rest of the city during these two months.

July and August G_i* - 1909

The July and August patterns for 1909 are displayed in Figure 5.12 and summarized in Table 5.2. During this particular time period the significant clusters are scattered throughout the city and display little in the way of an overall pattern. Most of the clusters are at larger distances, greater than 500m (Figure 5.12). Only in Southeast and Southwest were significant clusters adjacent to one another and then with no more than three blocks comprising a group of clusters.
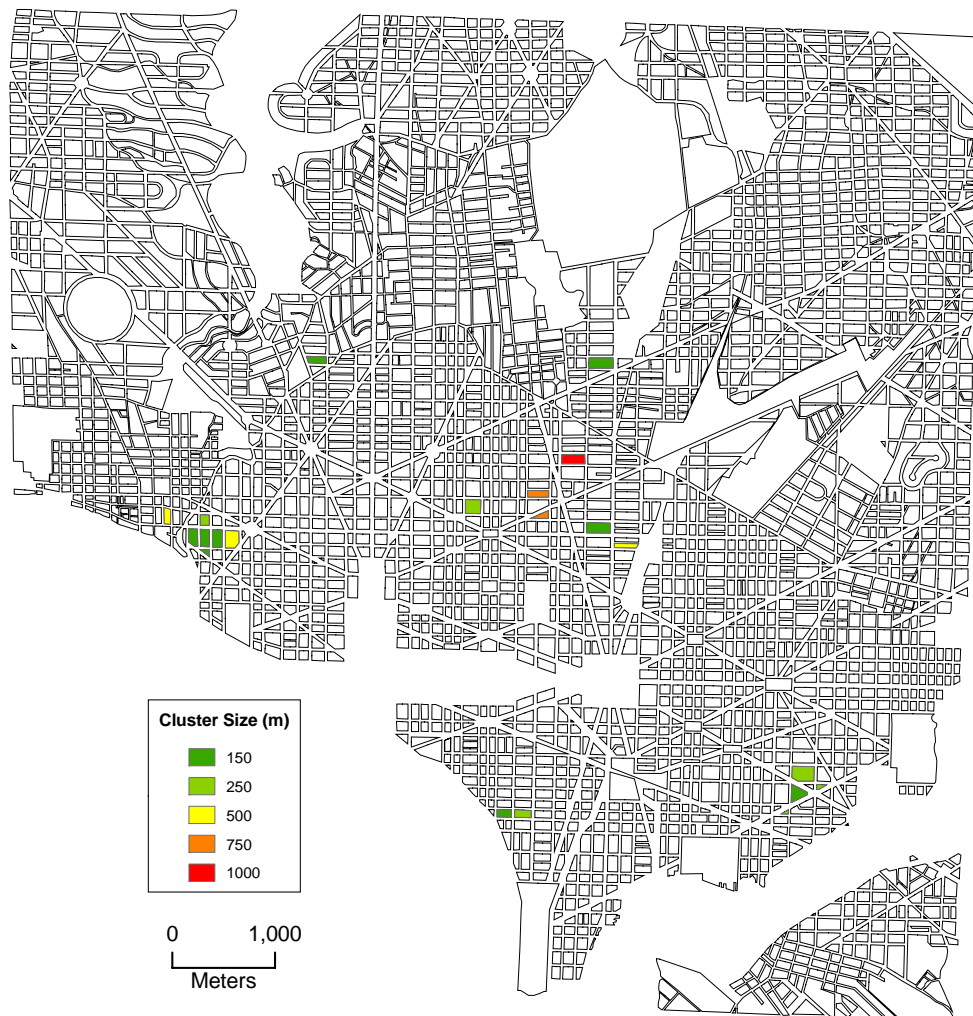
Figure 5.10 G_i* results for July and August 1907

Figure 5.11 G_i* results for July and August 1908

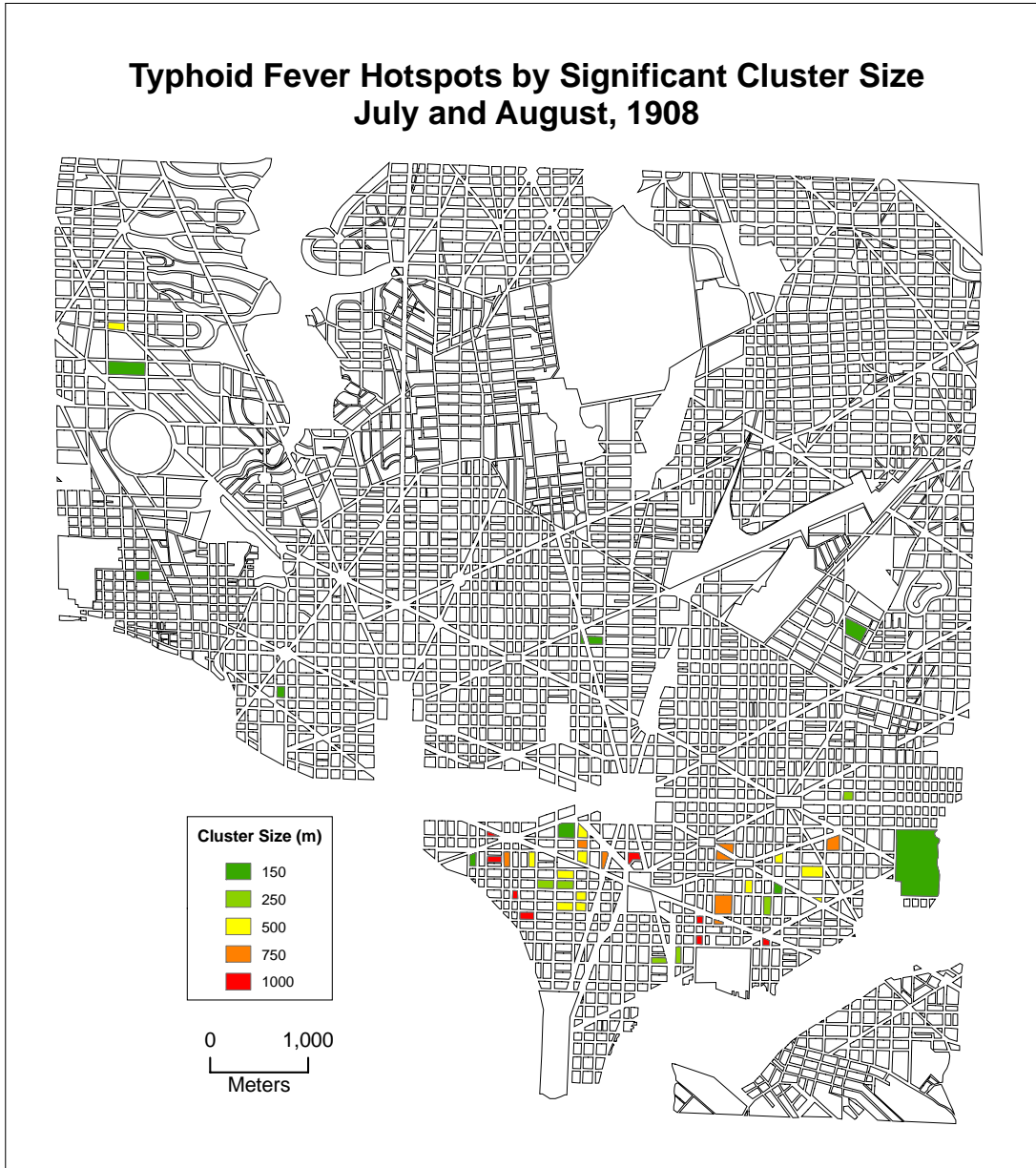**Typhoid Fever Hotspots by Significant Cluster Size
July and August, 1909**

Cluster Size (m)

150
250
500
750
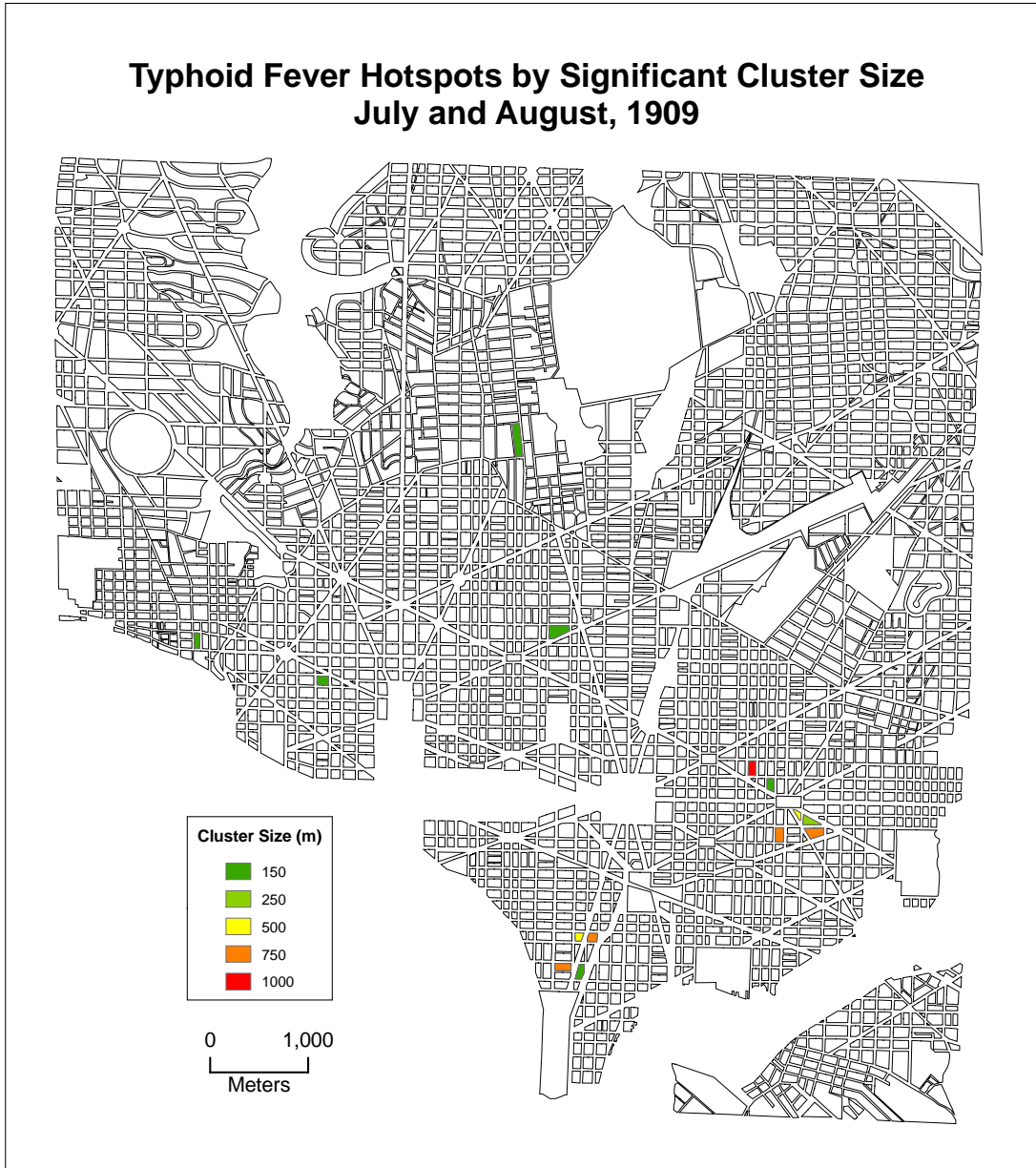1000

0   1,000

Meters

Figure 5.12 G_i* results for July and August 1909

<u>Monthly Histograms</u>

Most typhoid cases occurred in each region during July and August, whether looking at the city as a whole or the individual regions. There were some exceptions to this pattern that are discussed below. In some regions, in some years, there was an increase in the percentage of cases during the month of October. These results are summarized in Figures 5.13 through 5.17. The standard error bars on each graph generally indicate that there is little or no difference between the number of cases that occur in each month.
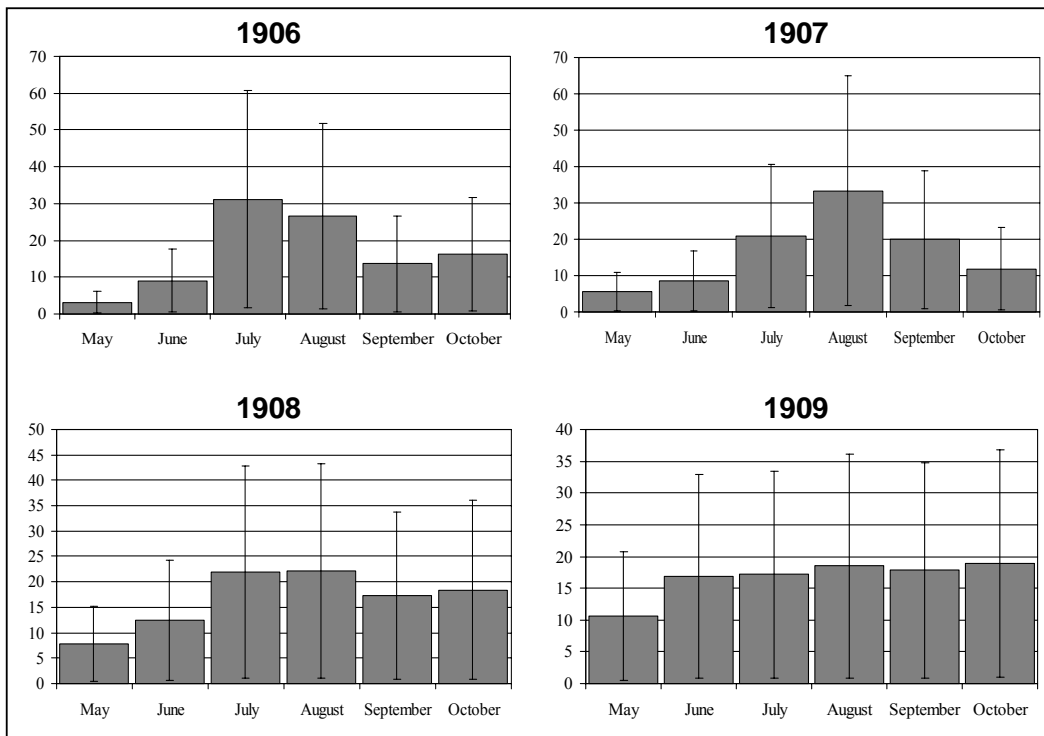
Figure 5.13 Histograms of summer typhoid showing the percentage of the total number of cases of typhoid that occurred in each month.

<u>1906</u>

Over 30 percent of the entire city's cases developed during July (Figure 5.13). Between 22 percent and 37 percent of cases showed signs of onset during July or August

in each region. The exception was Georgetown, where nearly 50 percent of the region's cases developed in July alone and another 36 percent in August. Northwest, the Suburbs, and Foggy Bottom all display an increase in cases in October. Anacostia, the most isolated of these regions, displays a strong peak in new cases in August, which is later than the rest of the city (Figure 5.14).

1907

More regions experienced their largest percentage of case onset in August rather than July during 1907, except in Anacostia, where 36.4 percent of cases showed signs of onset in July with a further 31.8 percent in August (Figure 5.15). Over 50 percent of the cases that occurred in Foggy Bottom developed during August, with another 21 percent having occurred in July. Only in Foggy Bottom was there an increase in cases in October 1907, which is a month when a continuation of epidemic decline is expected, yet did not happen in this case (Figure 5.15). The histogram of the whole city over the course of the six months studied in 1907 displayed a classic epidemiological curve (Figure 5.13).

1908

The city as a whole displayed only a slight increase in the percentage of cases that developed in July and August in 1908 (Figure 5.16). Only a few regions exceeded 30 percent of their cases beginning in a single month. These regions included Northeast and Anacostia in July and Southeast, Anacostia, and Foggy Bottom in August. Nearly 50 percent of the cases occurring in Georgetown that year occurred in October (Figure 5.16). This peak is explained in the original PHS report as an outbreak associated with contaminated milk distributed through this part of the city.
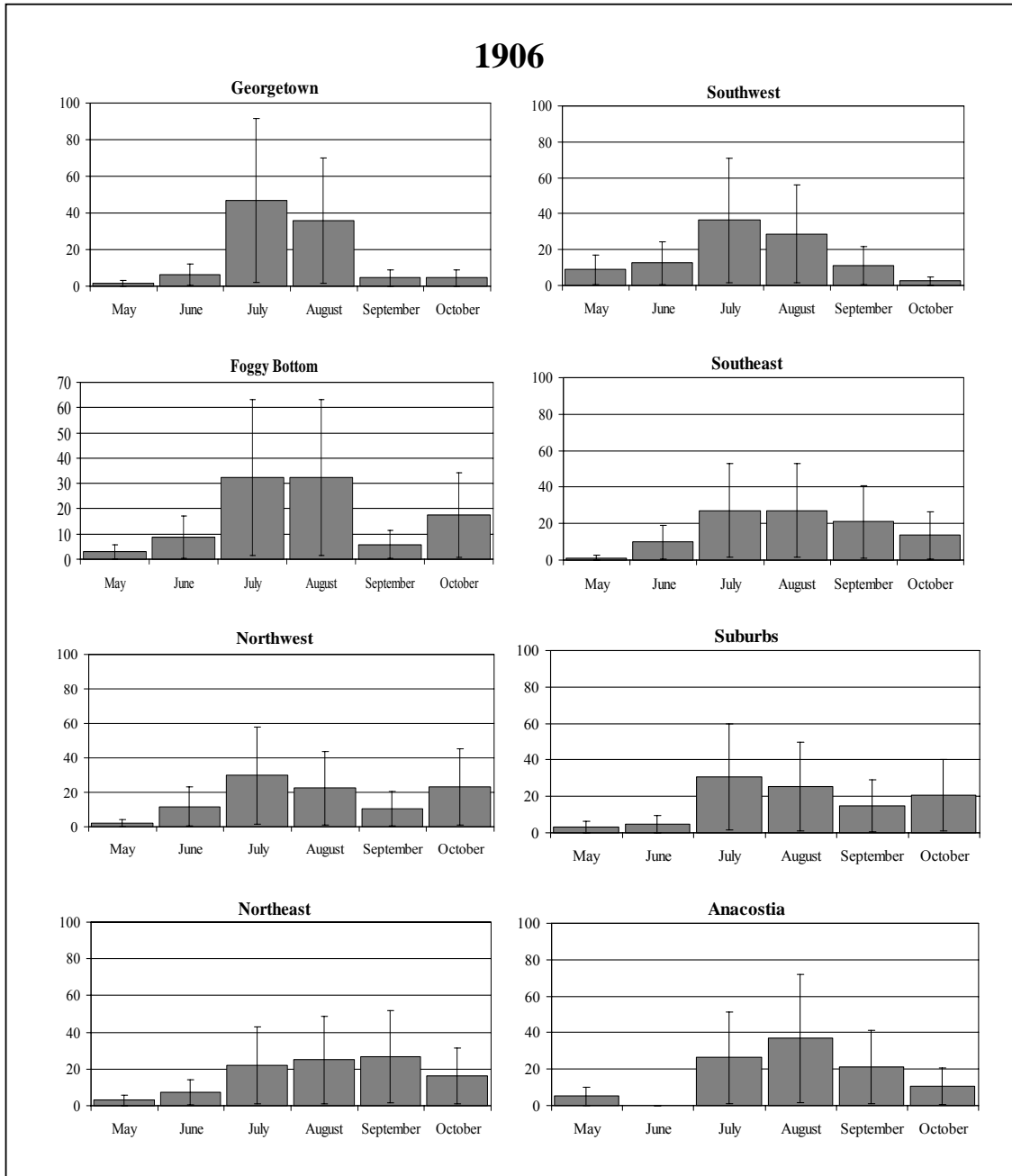
Figure 5.14 Regional histograms for 1906. Percentage of summer typhoid cases that occurred in each month.
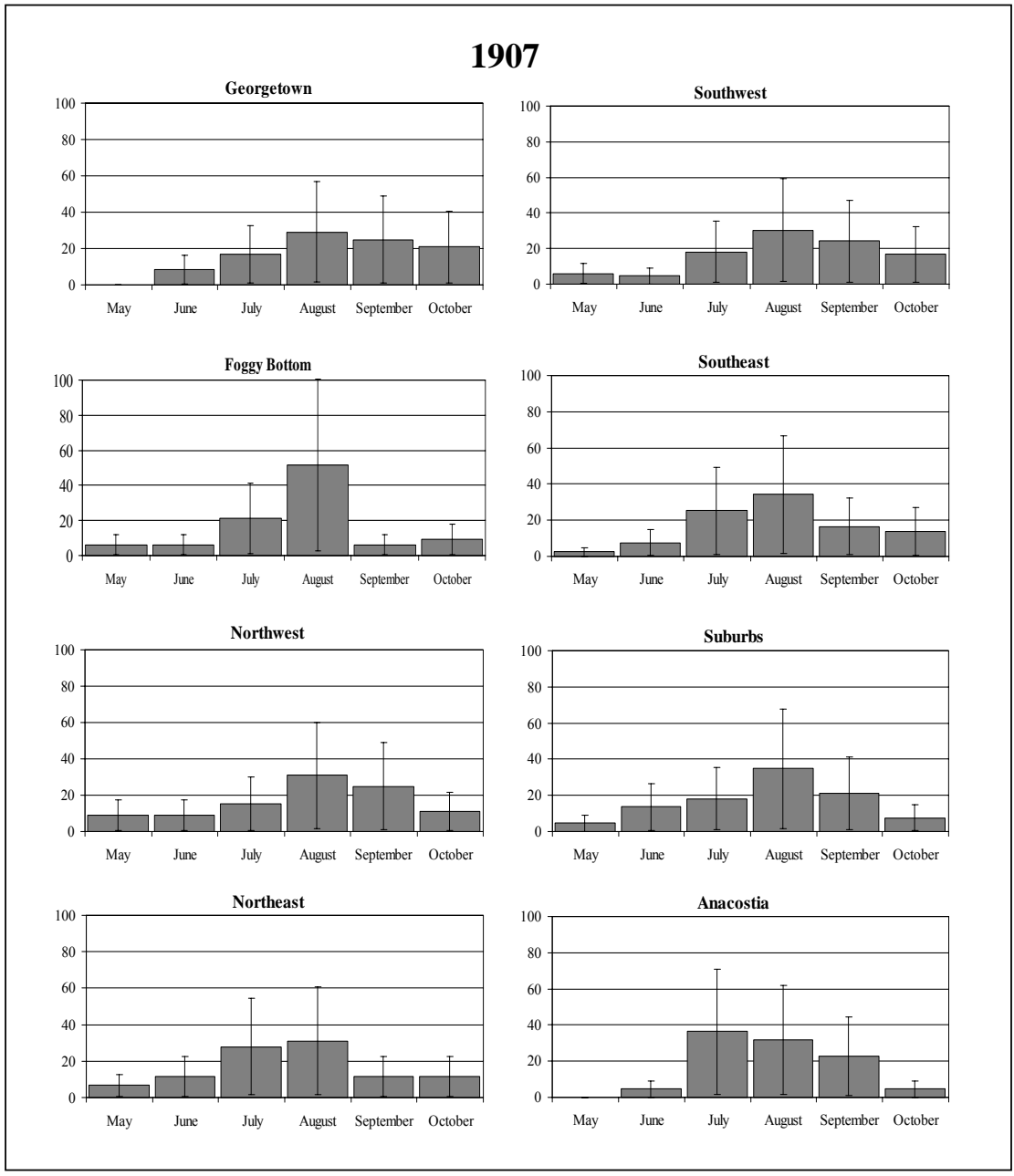
126

Figure 5.15 Regional histograms for 1907. Percentage of summer typhoid cases that occurred in each month.
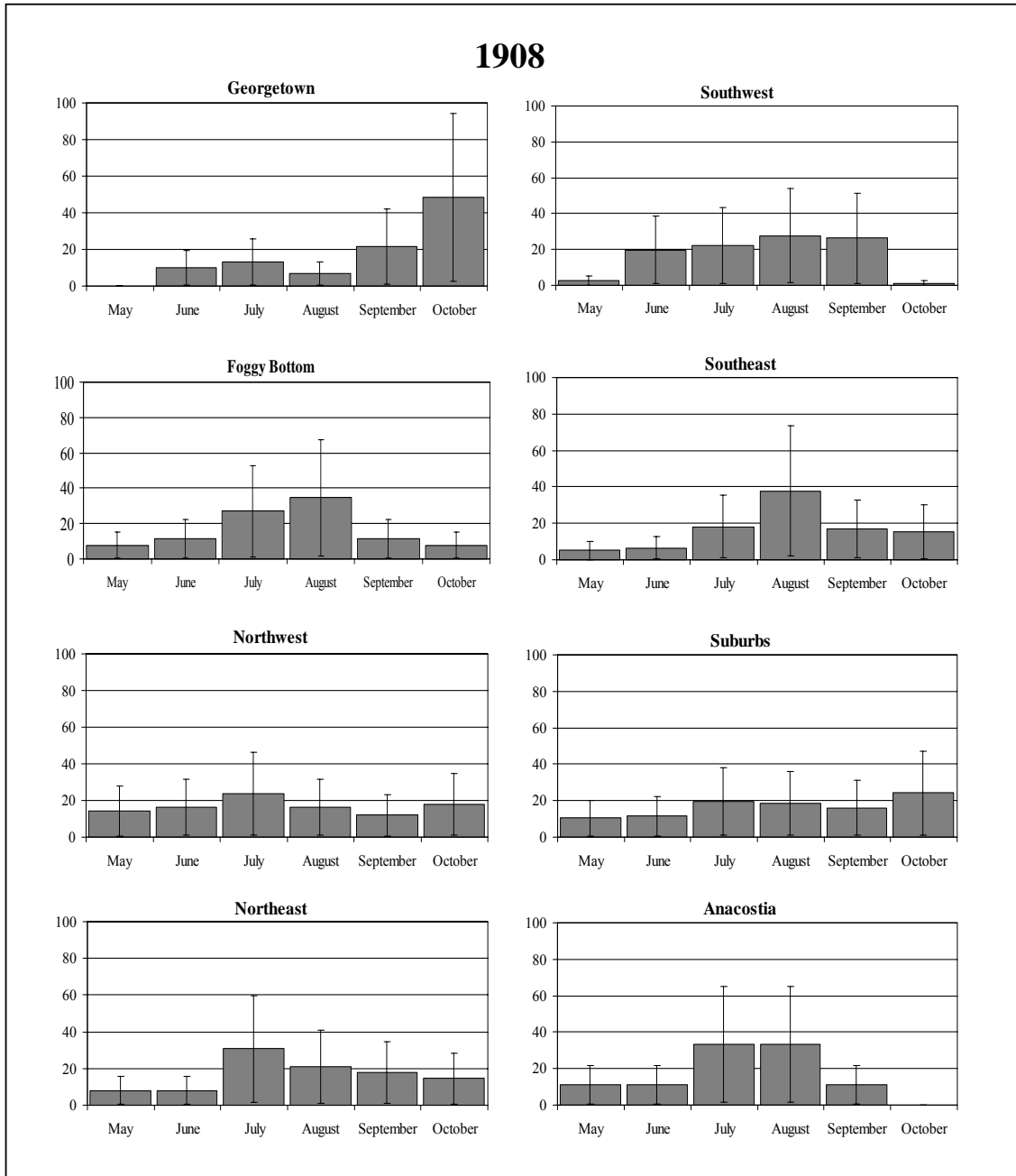
Figure 5.16 Regional histograms for 1908. Percentage of summer typhoid cases that occurred in each month.

1909

In the final year studied, fewer typhoid cases developed in July and August

(Figures 5.13 and 5.17). Additionally, only in October did any region (Northeast) display

more than 35 percent of its cases showing signs of onset in a single month (Figure 5.17).

When considering the city as a whole, there was an increase in cases between May and

June, and then the percentage of new cases in each month remained stable for the rest of

the study period (Figure 5.13).
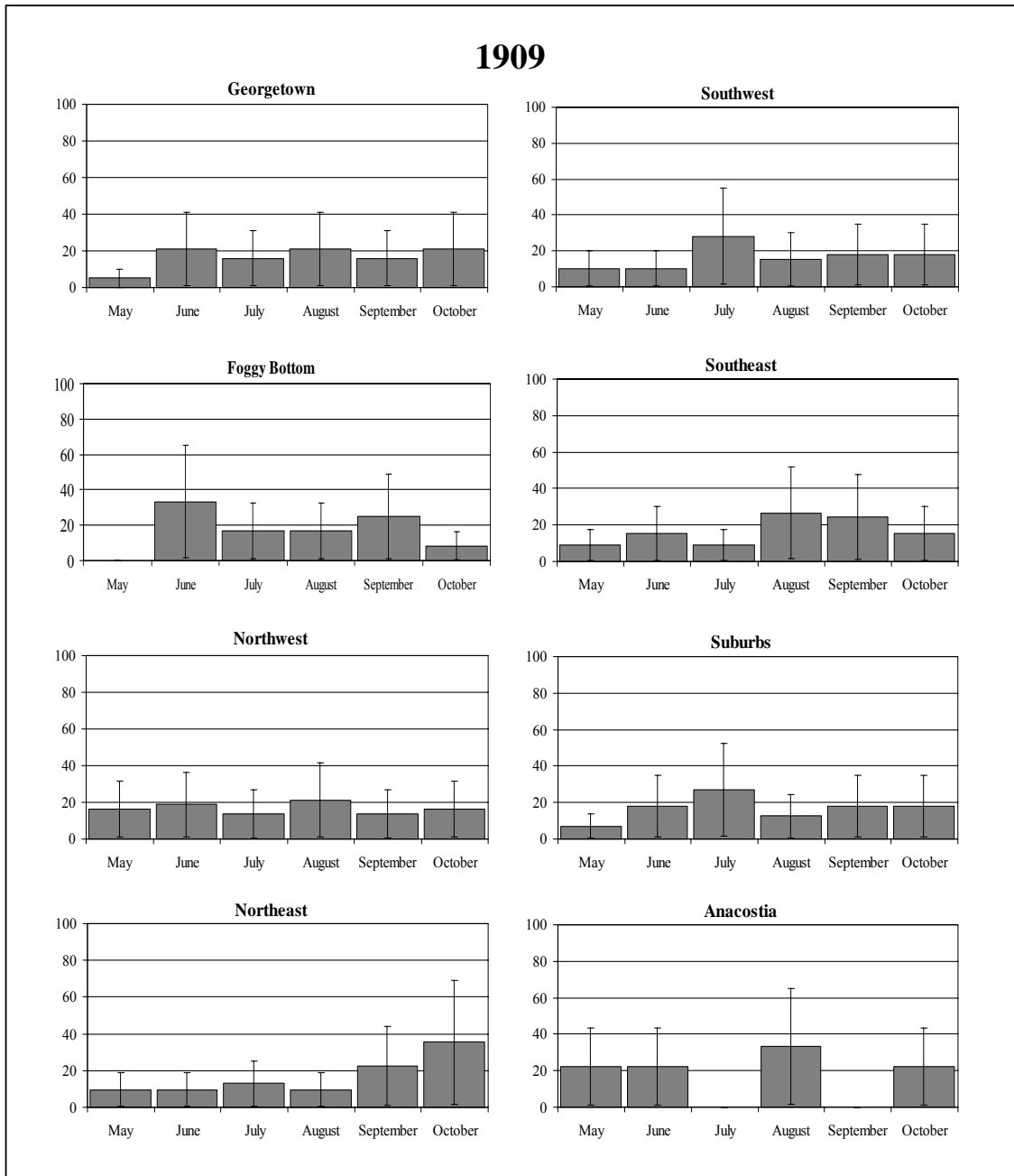
Figure 5.17 Regional histograms for 1909. Percentage of summer typhoid cases that occurred in each month.

Discussion

The first question posed in this chapter considered whether typhoid clustered within temporal phases of the outbreaks. In all eight 3-month and the four 2-month time periods investigated there were statistically significant clusters of typhoid fever, although the number of clusters and their sizes varied within and between years. Additionally, only in a few instances did a block remain a member of a cluster from the first half of an outbreak to the second, and relatively few blocks were members of clusters in more than one year.

The histograms of the percentage of each region's cases showing signs of onset by month address the second question posed in this chapter. Although it was not always as apparent on the 3-month maps, regardless of region and year, most cases show signs of onset during July and August. The July and August trend did vary slightly, such as in 1907 when August and September shared the highest percentages of new cases in Southwest, Georgetown, Northwest, and the Suburbs (Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911; Whipple 1908). The histograms provide a clear visual display of typhoid's seasonal trends during each 6-month study period.

It is interesting to note in Northwest, clusters appear in the August to October time period in every year, except 1908. A number of blocks in Northwest annually had typhoid cases and three neighboring blocks in particular were part of a cluster for two of the years (Figures 5.2, 5.4, and 5.8). Therefore, it seems strange that not a single one of these blocks or one of their neighbors were part of a cluster in 1908. After further investigation, and an extra run of the G_i* statistic to make sure that the milk related outbreak in Georgetown was not "washing out" clusters in Northwest it was determined

130

that there were no significant clusters in this location (Figure 5.18). This is explained by the drop in case numbers in Northwest in comparison to the global mean and variance applied in the G_i^* statistic. For that year, those blocks did not have enough cases to constitute a cluster compared to the rest of the blocks considered. Nonetheless, this is a good reminder that a lack of a significant cluster is not an indication of a lack of events in that location. In other words, a consistently endemic area may fluctuate between clustered significance and non-significance depending on whether it is an epidemic year or not, and additionally the size of the epidemic year. A finding such as this might lead to further investigation as to which local conditions are perpetually suitable for disease presence. This geographic area and some of the ideas expressed above will be explored in more detail in chapter 6.

<u>Typhoid Seasonality</u>

The PHS reports and other sources comment upon a summer peak in typhoid cases and deaths (Rosenau et al. 1907, 1908, 1909; Lumsden and Anderson 1911). This trend continued across all four years studied, but the degree of that summer peak lessened with each subsequent year (Figure 5.13). The 1906 PHS report included a chart of monthly typhoid *mortality* from the previous 30 years. This chart indicates that during most years typhoid mortality peaked during July, August, or September, following the warmest annual temperatures by approximately a month (Rosenau et al. 1907). The 1906 and 1907 reports also attribute the rise in typhoid cases during mid to late summer to a relationship between warmer temperatures and typhoid fever, although they do not elaborate further on what the link could be (Rosenau et al. 1907, 1908). Typhoid rates in 1909 remain fairly constant throughout the study period and the PHS report indicates that

this trend continues into 1910 (Figure 5.13) (Lumsden and Anderson 1911). Other contemporary sources further support the trend seen in the graphs for Washington, D.C. between 1906 and 1909. Cities world-wide, when graphed, display a peak in typhoid deaths during the late summer or early fall seasons (Whipple 1908). The shift away from the height of the outbreaks occurring during the warmest months by 1909 could be an indication of effective changes in the city's infrastructure or in general human behavior that kept outbreaks localized and limited spread to epidemic proportions.
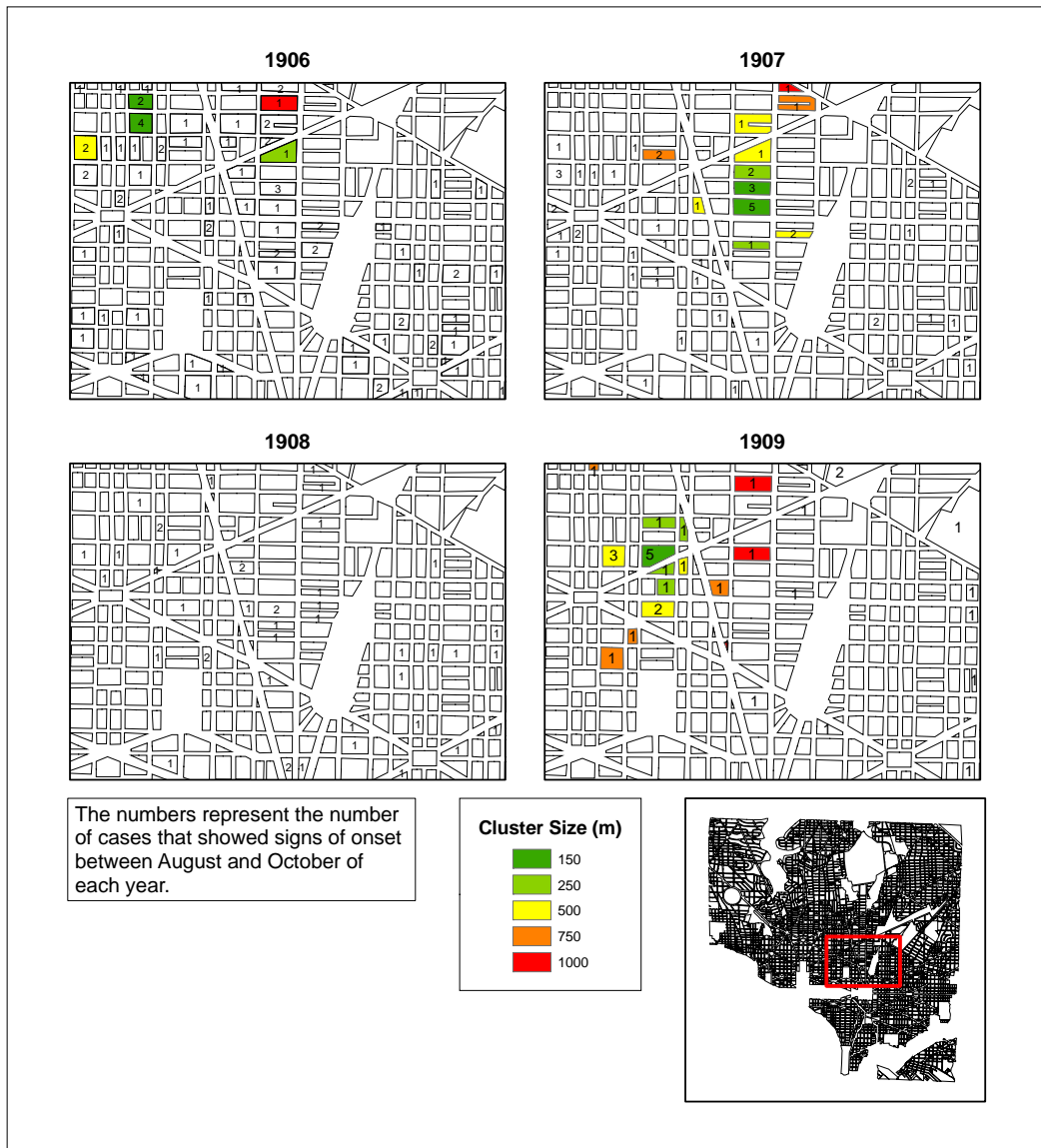
Figure 5.18 Map series of Northwest

By 1909, there was no consistency in when regions experienced the height of the "outbreak" (Figure 5.17). A rather interesting element of these inconsistent regional peaks in case numbers is that while a region had its largest number of cases, these cases did not cluster (Figure 5.7, 5.8 and 5.17). This combination of no clusters, but a high percentage of the disease in a region indicates that those infected with typhoid probably were not contracting the disease through the same sources, as there were many cases, but they were widely distributed. Additionally, the lack of clusters in 1909 is an indication that the number of overall cases had decreased (716 cases in 1909 compared to 936 cases in 1908) to a level such that few true clusters existed, which was generally supported by the 1909 typhoid report and the Health Officer's report (Lumsden and Anderson 1911; Health Officer 1911).

Rosenau et al. (1909) attribute a later peak in typhoid cases in 1907 to a cooler than usual month of June that year. In the 1908 report the authors describe in detail the 1906, 1907, and 1908 summer typhoid curves and their parallels to the temperature curve for those summers. Usually, the typhoid curve would follow the temperature curve closely, but offset by three weeks, which accounts for the approximate longest incubation period for typhoid (Rosenau et al. 1907). While unsure about what factors cause the link between typhoid prevalence and temperature, the authors did comment upon a few related aspects in their discussion of the temperature curves in 1908. One of the related elements in the discussion was an abundance of houseflies and their possible relationship to typhoid prevalence. This was of particular interest since the number of flies tended to increase when temperatures rose. The PHS investigators noted that while typhoid rates

remained stable in 1906, 1907, and 1908 during the cooler months (January through April and November through December), and although case numbers increased during the summer for all three years, these increases became progressively smaller This was explained as being due to the opening of the sand filtration plant (Rosenau et al. 1909).

In a report concerned with the Potomac River watershed, the United States Geological Survey (USGS) also discussed the consistent pairing of typhoid and high temperature, as well as the pairing of typhoid and low stream flow (Parker et al. 1907). The report suggested that warm temperatures create an optimal environment for typhoid's survival in water until the bacteria could be consumed by an unsuspecting individual(s). Again in this report, typhoid mortality tended to reach its peak about a month following peaks in both air temperature, and the temperature of the Potomac River, both of which were closely related to one another (Parker et al. 1907).

Spatial Patterns of Typhoid

With the exception of 1908, where the pattern of clusters is quite different to that of any other year, a potential pattern of disease progression can be seen. During the May to July time period typhoid clusters appear in the more peripheral parts of the city (Georgetown, Southwest, Anacostia, Foggy Bottom) while during the August to October time period the clusters centralize in the same neighborhood in Northwest with a few clusters in the Suburbs near Boundary Street, although these Suburban clusters did not appear in 1909 (Figures 1.2; 5.2 through 5.8).

In 1906, the May to July time period display a number of clusters in Georgetown (Figure 5.1). By the latter part of the study period the typhoid fever clusters shift to the

north and east. The clusters north of Boundary Street are located in the enclave of black residences discussed in chapter 4 (Groves and Muller 1975).

During 1907, the group of clusters identified in chapter 4 in Foggy Bottom are once again evident, now the movement of the clusters from block to block is visible in Figures 5.3, 5.4, and 5.10. This part of the city includes another black residential enclave identified by Groves and Muller (1975). Generally, during the August to October time period in 1907 there is a shift in the location of clusters throughout the city to the north of their locations in the early part of the year. The latter part of the study period display the only significant clusters in any year near the elite neighborhood around Dupont Circle (Green 1963). This is interesting as the literature strongly supports the finding that wealth provided access to certain amenities that helped to protect individuals from disease, such as cleaner water, good health care, the most modern facilities in the way of water closets, and the best quality diet. For the time period under study, black residents and other marginalized groups were less likely to live in the most sanitary parts of the city or to receive good quality health care, if they received health care at all (Green 1963; Groves and Muller 1975; Moldow 1980; Williams 1992; Thornton and Olson 1993; Borchert 1980; Condran and Crimmins-Gardner 1978; Wills 1996; Olson 1997; Mckeown and Record 1962).

The temporal spatial patterns in 1908, differed from the other years under investigation and in the extent to which the epidemiological investigation linked cases to causes (Rosenau et al. 1907, 1908, 1909). In the 1908 PHS report the researchers describe two areas of the city with elevated concentrations of typhoid. The first of these concentrations was easily explained in the report as being caused by contaminated milk

delivered to a number of households in Georgetown during September and October (Rosenau et al. 1907, 1908, 1909). This outbreak, which is clearly visible in Figure 5.6, began in late September and continued through October until the source of the outbreak was identified and that particular dairy closed. The second concentration was in the southern part of the city where there were more cases of unknown origin (Rosenau et al. 1907, 1908, 1909). This concentration of cases is clearly displayed in Figures 5.5, 5.6, and 5.11. While the G_i* statistic maps show the concentration of clusters in Southwest and parts of Southeast more information is needed to understand why typhoid clustered in those locations.

The clusters in Southwest during 1908 were nearly all in the predominantly white residential part of that region. To the west of 4½ Street Southwest, the racial profile was predominantly white while to the east of that dividing street the population was mostly black (Figure 5.19) (Groves 1974; Health Officer 1910). Black residents from all socio-economic groups lived in Southwest, while the white population was primarily of the working class (Groves 1973-74; Groves and Muller 1975; Borchert 1980). The Vital Statistics district west of 4½ Street had a lower overall death rate in 1908 than the two districts to the east of this dividing line. Again, without knowing what any household-level sanitary conditions were like, or knowing their sources of water and milk, it is impossible to accurately explain why there were more clusters of typhoid fever west of 4½ Street in 1908.

In 1909, there were few clusters in either time period. During May to July more significant clusters were located in Southwest, but in general clusters were randomly distributed throughout the study area (Figure 5.7). In August to October clusters only

136

appear in Northwest (Figure 5.8), and the size of this cluster is suspect due to statistical limitations of working with small numbers. In general the results for the temporal analysis of 1909 and particularly the later part of the study period, should be viewed with caution given the small number of cases.
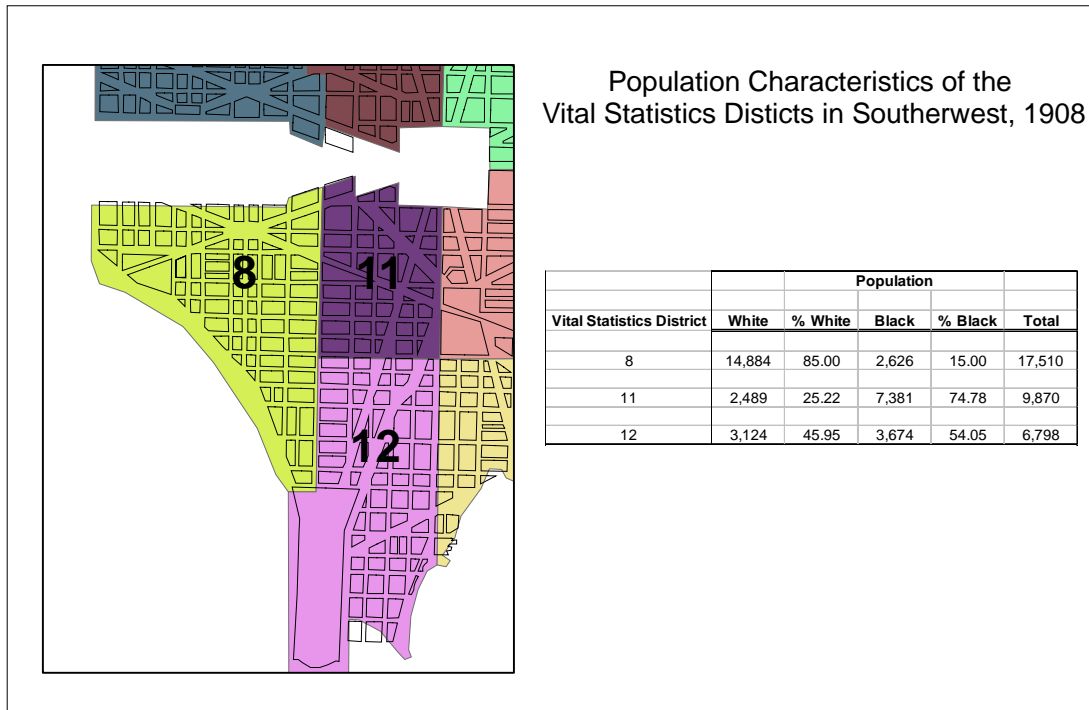

Population Characteristics of the
Vital Statistics Disticts in Southerwest, 1908

Vital Statistics District	White	% White	Population Black	% Black	Total
8	14,884	85.00	2,626	15.00	17,510
11	2,489	25.22	7,381	74.78	9,870
12	3,124	45.95	3,674	54.05	6,798

Figure 5.19 Vital Statistics Districts in Southwest, Washington showing population characteristics

The histograms presented in this chapter display that over the course of the four years the expected pattern of decreased seasonality of the disease following the introduction of filtration did occur. This indicates that the presumed primary cause of seasonal typhoid, most likely contaminated water, lessened or was eliminated as a cause. Whipple (1908) echoes this pattern for other cities. The maps of clusters for July and August help to show the movement (diffusion) of typhoid between a few small areas. For example, in 1907 the clusters in Foggy Bottom can be seen in two adjacent blocks in the

May to July time period (Figure 5.3). In July and August alone this group of clusters was at its maximum size, covering five adjacent blocks (Figure 5.10). Finally, in the August to October time period, there were only two clusters, but these were the two eastern most significant blocks of the July to August map (Figure 5.4). A similar trend is visible in Southwest during the 1908 outbreak (Figures 5.5, 5.6, and 5.11). In that case, the shift in cluster size and location is a little easier to track by viewing the July and August G_i* results with the results from the 3-month analysis.

It is unfortunate that the small number of cases per block made running the G_i* statistic on a finer temporal scale than 3-months impossible. The ability to visualize the monthly movements of typhoid clusters, if they existed, in Washington during these years could have been informative. The available data made such an enterprise impossible without violating the statistical assumptions of the technique.

One interesting item of note concerning the lack of clusters in Northwest during 1908, particularly during the latter half of the study period, is that cases still did occur in this area (Figure 5.6), and the peak of the outbreak in Northwest in that year was in July, which is early in the season for a regional peak (Figure 5.16). When only July and August were analyzed using the G_i*, a single 150m cluster did appear in Northwest (Figure 5.11). This result again indicates how both spatial and temporal scale can change an overall epidemiological impression of an area.

Without similarly scaled population and socio-economic data it is difficult to determine the reason why typhoid cases consistently cluster in the eastern part of Northwest. Parts of Northwest in question were two of the most populated Vital Statistics districts in the city (Health Officer 1910). Beyond sheer population numbers, it is

138

difficult to understand why that region consistently displayed clusters in the latter half of each study period. Since the small numbers problem limits the extent of the temporal exploration of typhoid fever with the G_i^* statistic, in chapter 6 a different approach to visualizing the temporal patterns of these data will be employed.

Introduction

The previous two chapters identified local clusters of typhoid fever at multiple spatial scales and for different temporal aggregations. The goal of this study thus far has been to better understand the spatial and temporal structure of typhoid fever in early twentieth century Washington, D.C. through the identification of typhoid hotspots in space and time. Chapter 5 identified some limitations in the use of the G_i* statistic as a measure of local spatial auto-correlation when the number of cases per block becomes too low, as is prone to be the situation when fine aggregations of time are analyzed. The hotspot surfaces generated in chapters 4 and 5 though insightful are also not designed to capture local patterns within and around the significant clusters. For example, a 150m cluster center block does not show, given the majority of the cases originated in that block, what was the distribution of cases in the adjacent blocks. This situation becomes even less clear as the significant cluster distance enlarges upwards towards 1000m. In order to visualize this neighborhood-level pattern of cases, the data manipulation capabilities of the GIS can be utilized. A descriptive diffusion can be achieved by mapping the bi-weekly progression of cases.

Smooth spatial and temporal waves usually can only be seen at the smallest spatial scale. Within outbreaks, local variations, such as social pathways and barriers in the built-environment will influence the direction and velocity of a disease through a city (Edling and Liljeros 2004). The agent itself, or diffusion mechanisms such as vector presence and activity, will also lead to inconsistencies in a space-time progression

(Edling and Lijeros 2004; Woolhouse et al. 2005). An area receiving scant attention in the literature is the nature of neighborhood-level spread, how different large scale patterns coalesce to form the traditionally identified hotspot surface.

Explorations of block-level typhoid cases in the previously described neighborhoods of the city can provide insight into how the disease may have moved through these areas, and how the clusters identified in chapters 4 and 5 developed through time. Since many aspects of daily human life influence where individuals go and how communities in different parts of a single city interact, small areas of a city may exhibit different spatial-temporal patterns of disease movement. In other words, the conditions leading to the development of a significant cluster as identified by the G_i* analysis may vary geographically. The space and time diffusion pattern for Foggy Bottom may be quite different to that of Northwest. The visual assessment by the PHS, and the results of the Ripley's K-function indicate a lack of global typhoid clustering, yet the local clustering measure, G_i*, reveal hotspots, and that typhoid was not evenly distributed throughout Washington, D.C. It is likely that when comparing statistically significant clusters that directional and temporal patterns will vary. It has been shown that temporal aggregation alters the size and location of many of the clusters that were identified when an entire outbreak year was studied. The next logical line of inquiry is: where did the actual cases present themselves in relationship to the significant block?

The lack of contextual information about each case; such as race, age, or water source; limits some of this part of the exploration process, but by considering the points themselves some clues about the spatial structure of typhoid may be revealed despite the limits on the human components of each event.

Methods

Using the results from chapters 4 and 5, four areas of the city were selected for further investigation. The following areas; Southwest, Foggy Bottom, a small part of Northwest, and what is referred to here as the Boundary area;[6] were selected based on the G_i^* cluster maps from the previous two chapters (Figure 6.1). Of the areas, only Northwest lacked clusters (1908), as was discussed in chapter 5, but clearly this situation also needed additional investigation. Further, although the selection criteria did not involve any social or racial aspect, all four of the areas contained all or part of an African-American enclave as identified by Paul Groves and Edward Muller (Groves 1973-74; Groves and Muller 1975).

To visualize the development of clusters in space and time, sequential maps of the case data for 1906, 1907, and 1908 are created for each of the four areas. Each map displays all new cases and previous cases (for that year only) for each two week time period. In other words, for Foggy Bottom the first map displays just those cases with typhoid onset between 1 May and 15 May 1906. The second map displays all of the cases from 16 May through 31 May 1906 over those cases from the first map in a new symbol (Figure 6.2a). In this way, the map reader can visualize where typhoid first appeared in each local area and where it diffused to by two-week intervals until the entire study period is displayed on the twelfth map in the series (Figure 6.2c). The 1909 data were excluded from this analysis since these data only appeared in monthly aggregations. The maps created by this methodology can be compared to one another for potential

---

[6] The Boundary area is an arbitrary name for the section of the city just north of Boundary Street, Northwest that is defined by the group of significant clusters in 1906 that were located in the black residential enclave in the suburbs just north of the city (Figures 1.2, 4.7, and 6.1).

The Four Selected Local Areas
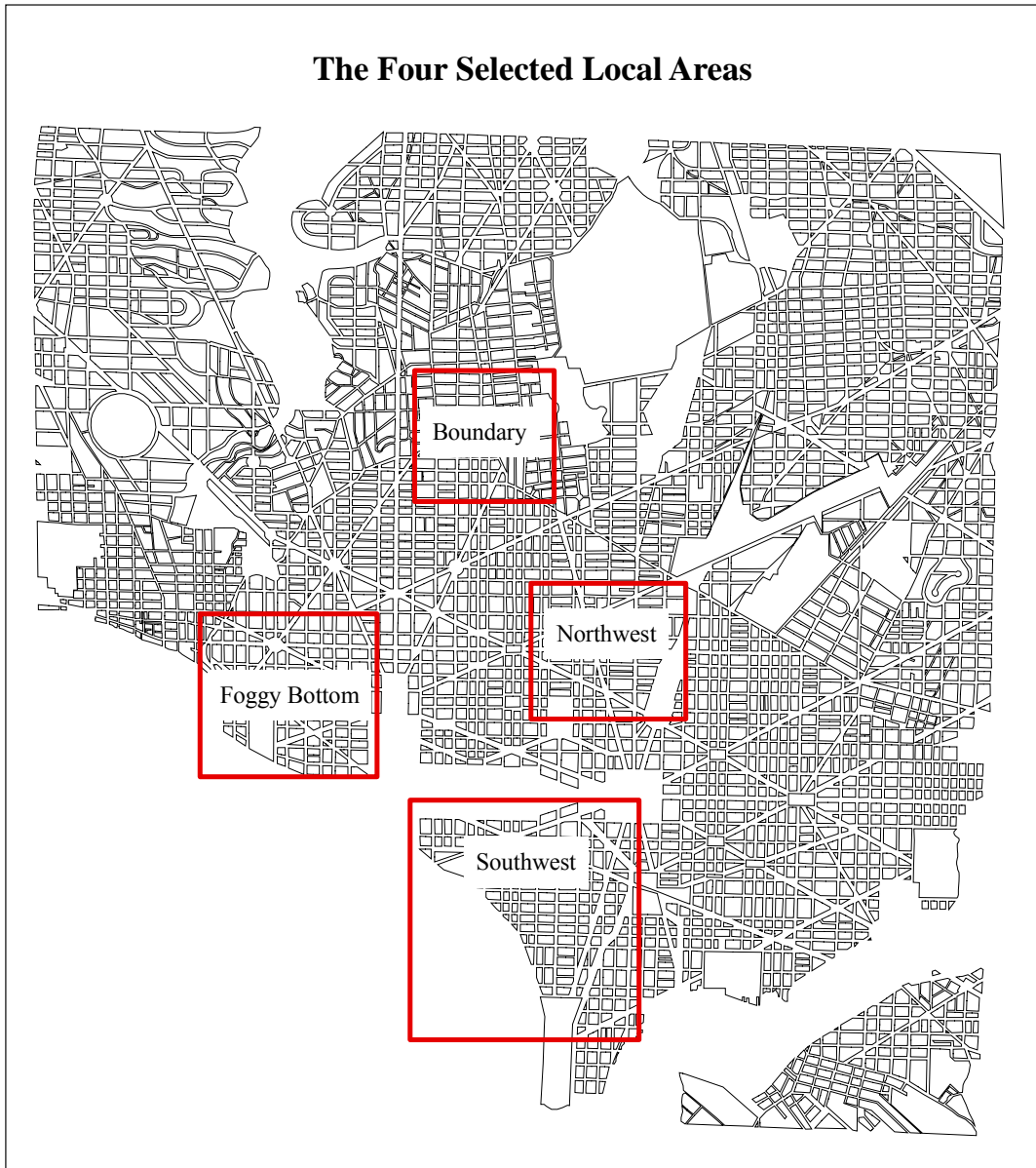
Boundary

Northwest

Foggy Bottom

Southwest

Figure 6.1 Selected neighborhoods

connections between existing cases and the new cases added with each subsequent time period. Drawing spatial connections between cases is not a new approach to understanding macro-scale diffusion. Indeed, contemporary reports would often describe the spatial progression of a disease between residences, with sometimes drawing connecting lines. Spider diagrams have also been used by other medical geographers working with historic epidemics (Curson 1985; Cliff et al. 1998; Wilson 1993; Edling and Liljeros 2004).

Results and Discussion

Twelve maps display this fortnightly sequence for each year under investigation, and for each area, resulting in 144 total surfaces (Figures 6.2, 6.3, and Appendix B). These results should only be viewed as a first step towards a more detailed understanding of neighborhood level space-time interactions of disease spread. The patterns vary between the four areas, particularly in how the timing of case onset impacted the development of the statistically significant clusters seen in chapter 4. To simplify the discussion here, only two examples from these results are presented, Foggy Bottom in 1908 and Northwest in 1907. The remaining maps can be found in Appendix B and will be used in future research as the starting point for a detailed study of neighborhood-level health, social, and environmental geography. The focus of such an investigation is currently outside the scope of this particular dissertation and therefore, will be a part of the future research agenda.

What stands out on the sequential maps of Foggy Bottom and Northwest is that the relationship between a significant cluster, and the temporal sequence of disease entry into the region need not be the same (Figures 6.2 and 6.3). For example, in 1908, two

blocks in Foggy Bottom were members of significant 150m clusters, yet when the cases for the six months are broken down into their two-week intervals, typhoid appears throughout this area and in fact did not even reach the two blocks identified as clusters until July, midway through the typhoid season (Figures 6.2b).
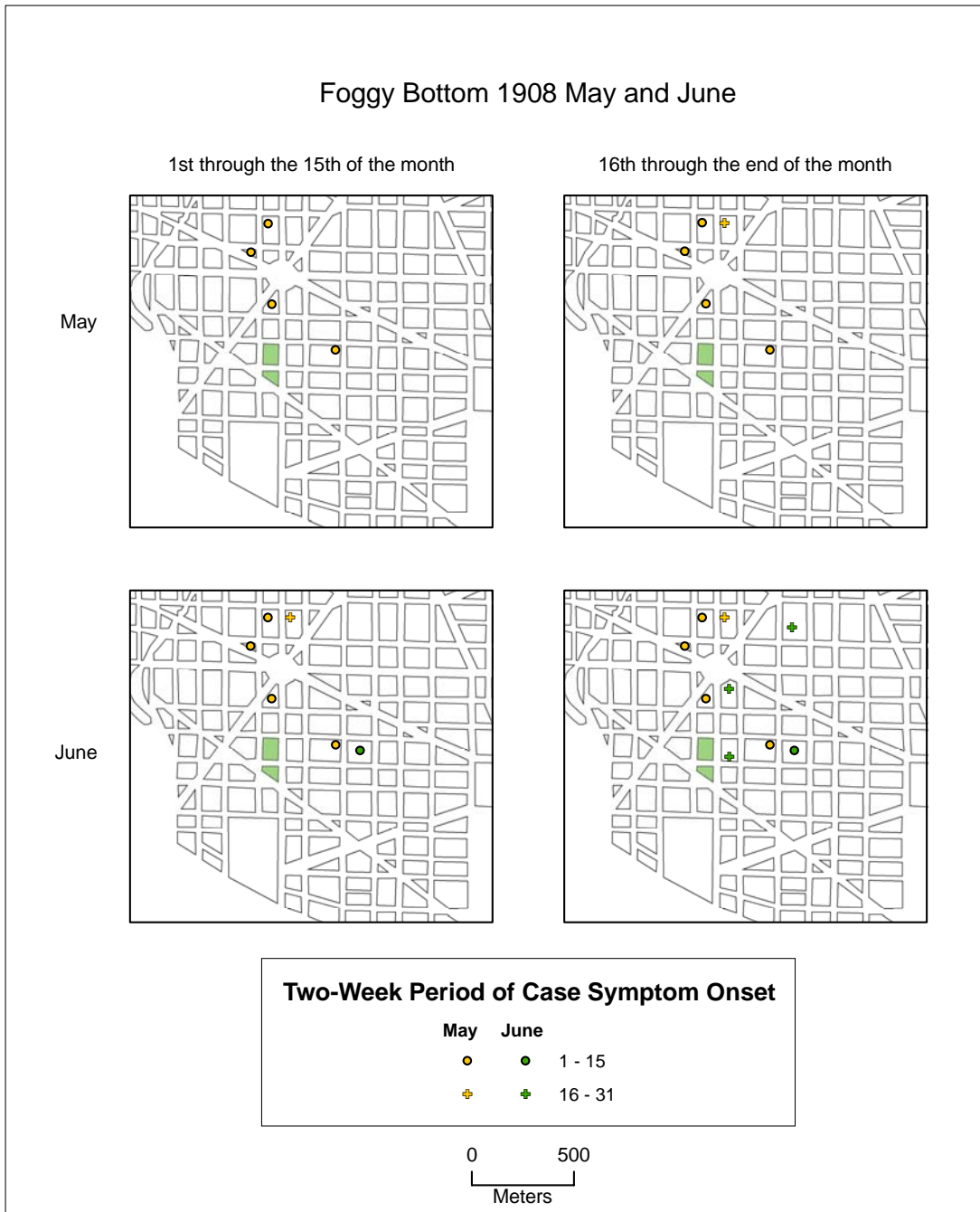

Figure 6.2a Temporal sequence of typhoid cases in Foggy Bottom
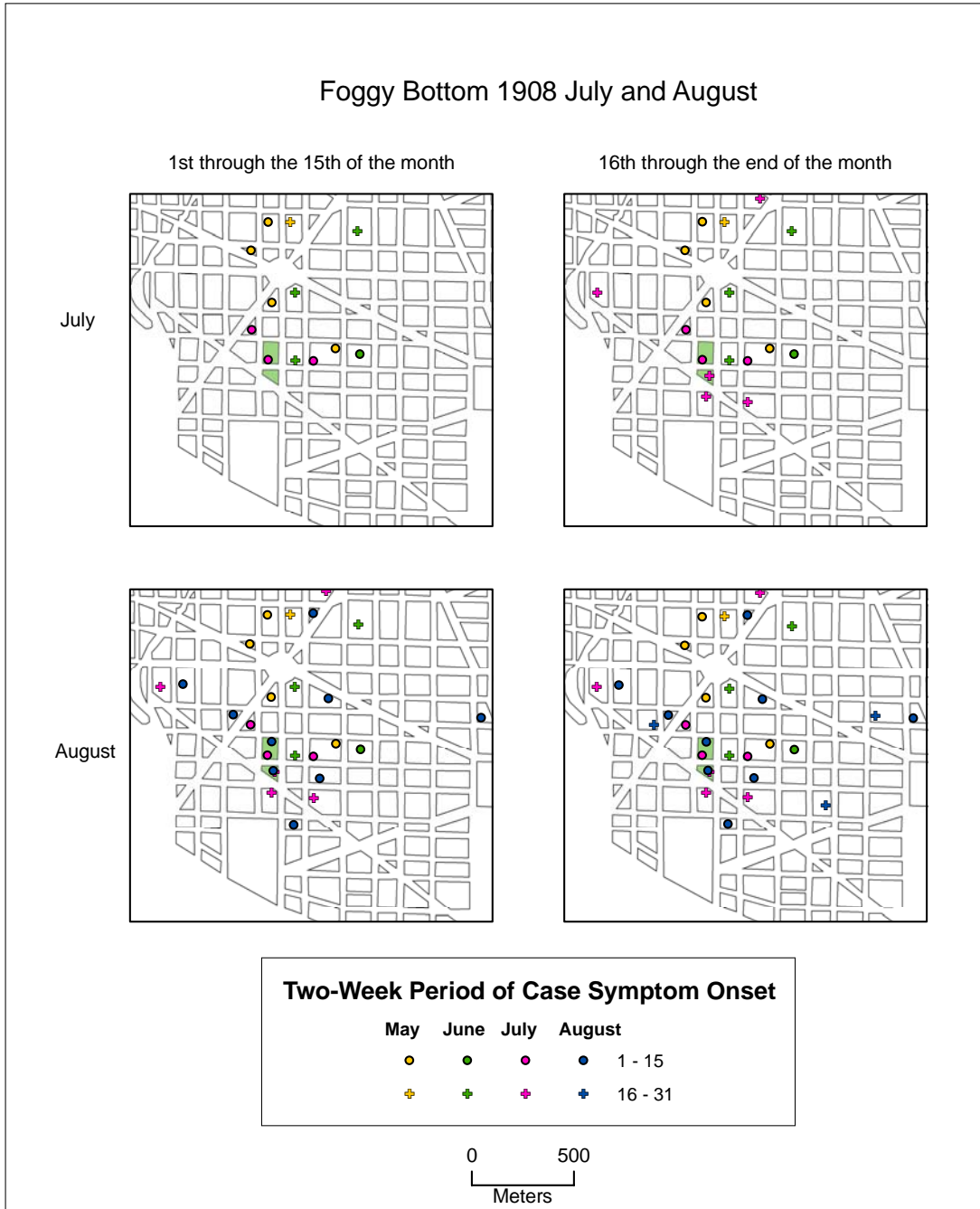
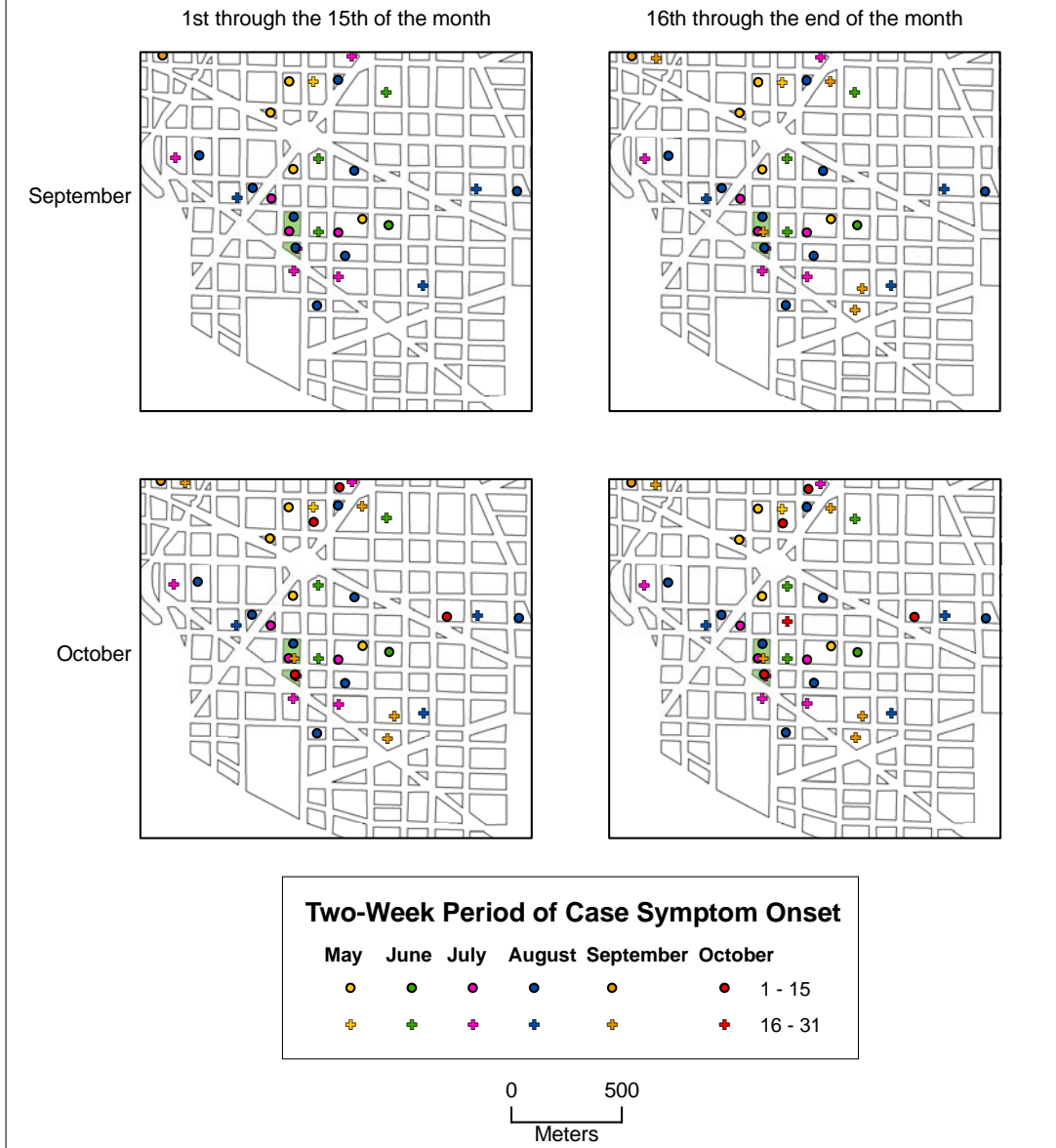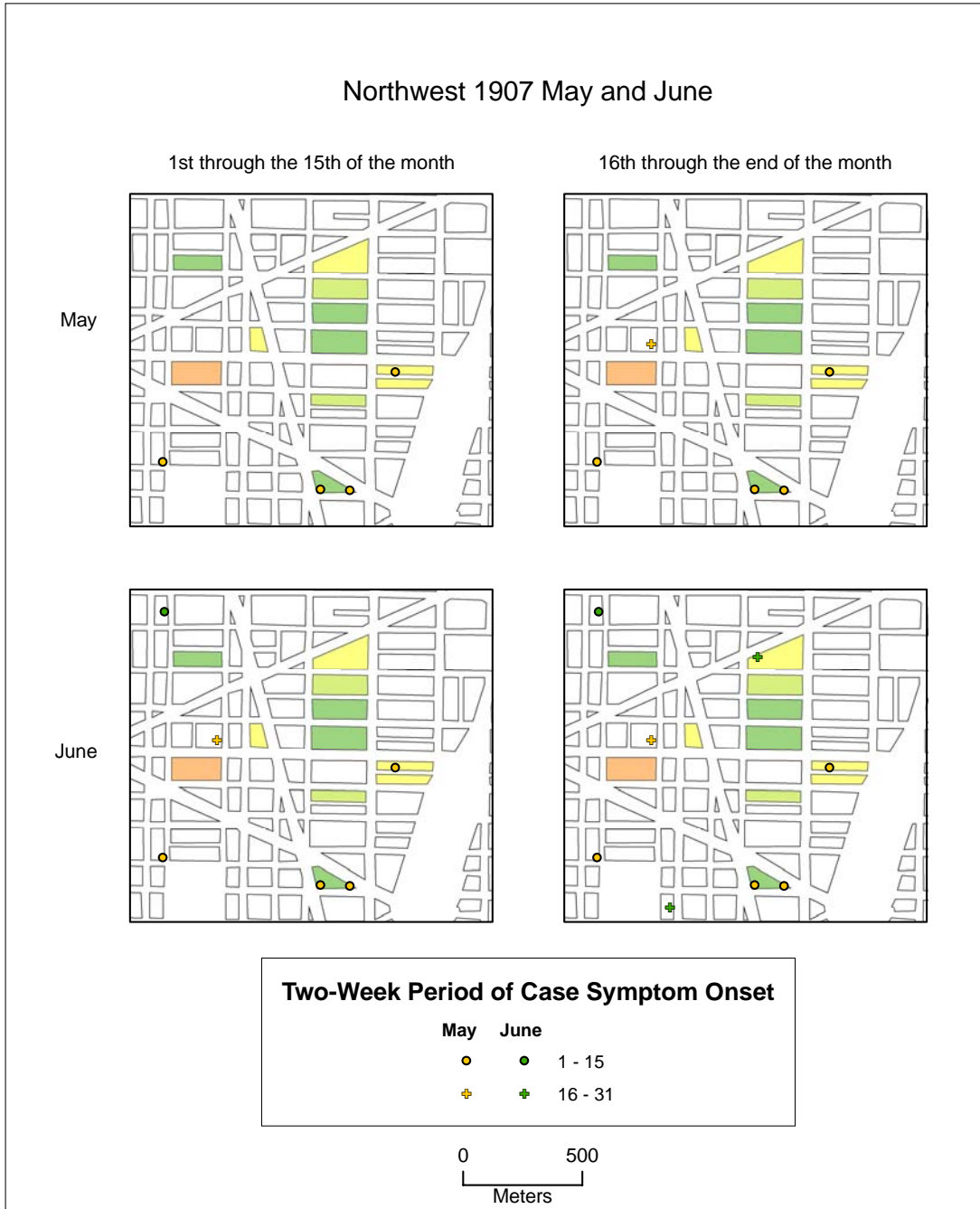Foggy Bottom 1908 July and August

Figure 6.2b

Foggy Bottom 1908 September and October

Figure 6.2c

Figure 6.3a Temporal sequence of typhoid cases in Northwest

Northwest 1907 July and August

1st through the 15th of the month   16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

May	June	July	August	
●	●	●	●	1 - 15
✚	✚	✚	✚	16 - 31

0   500

Meters
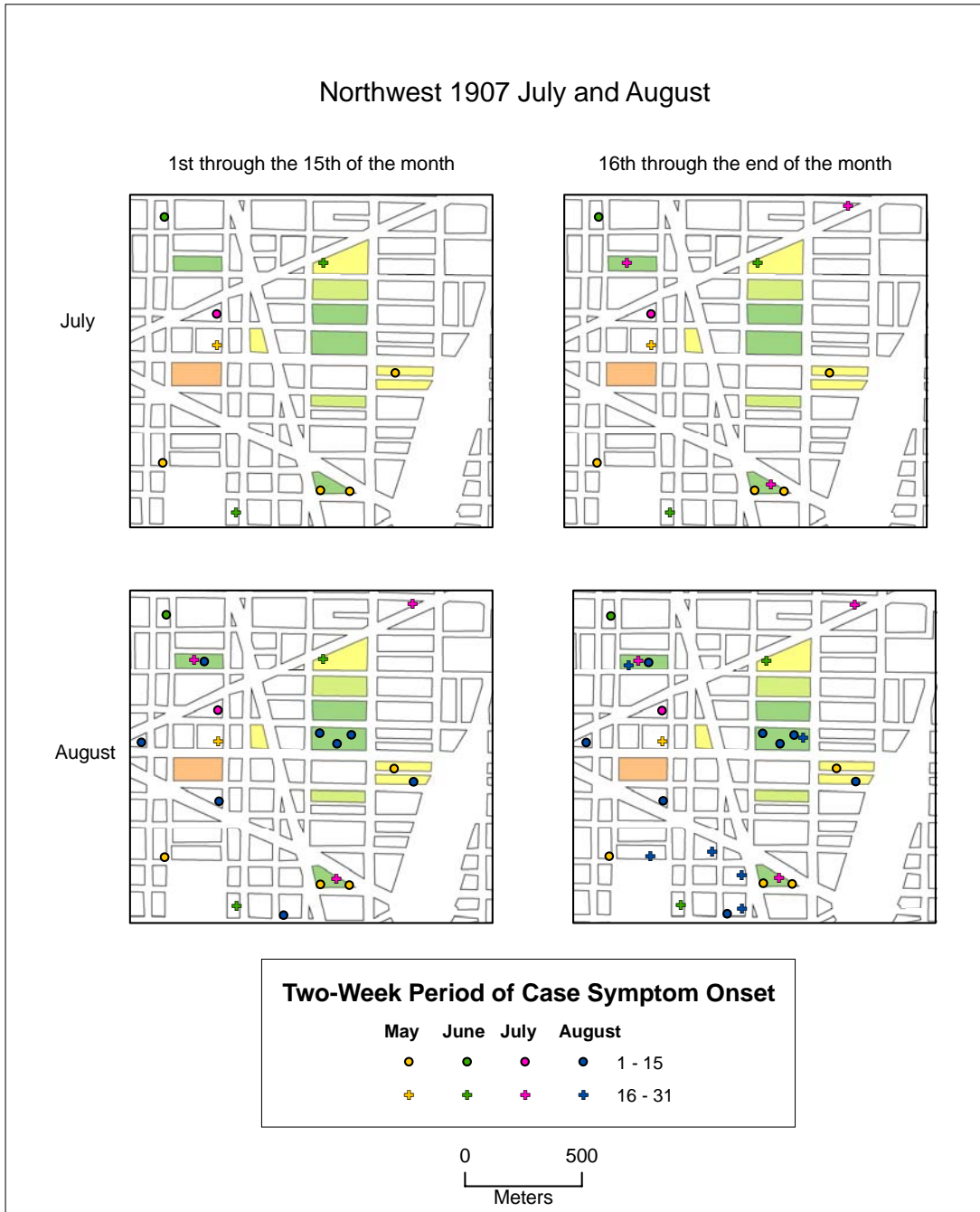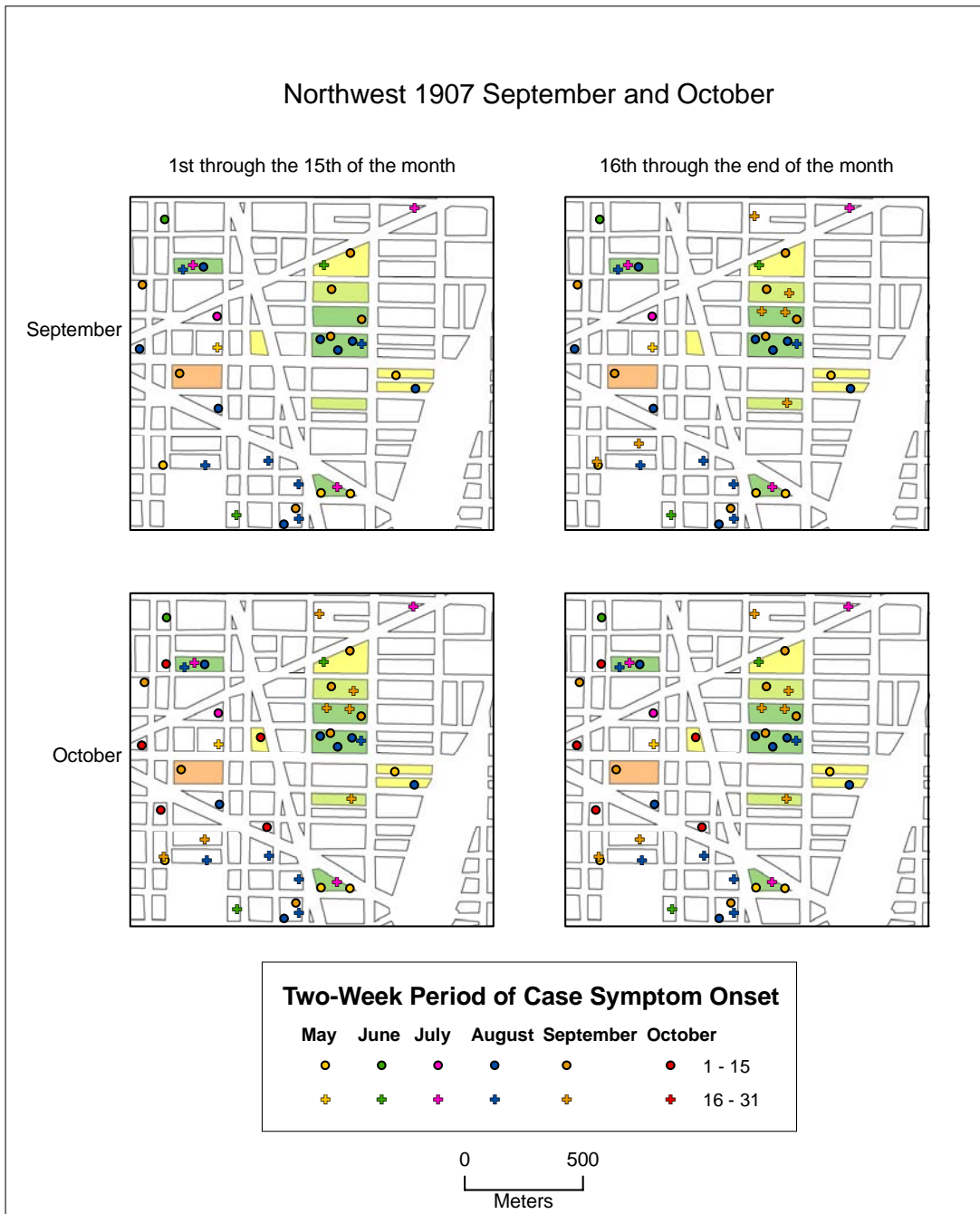
Figure 6.3b

Figure 6.3c

Foggy Bottom

This region in 1908 provides an interesting illustration of typhoid diffusion through a small area (Figure 6.2). Beginning with the first two weeks of May and progressing forward, most new cases of typhoid appear on neighboring blocks to those with existing cases. This is somewhat expected for a contagious disease. This proximate diffusion process eventually leads to the blocks identified as being significant for the entire outbreak. The first cases to appear on either of the hotspot blocks occurred in the first part of July (Figure 6.2b). A case had already appeared in the latter part of June on the block immediately to the east of this northern hotspot. New cases developed in the two hotspots in the following sequence: The northern hotspot center had three cases, one developing in early July, a second in early August, and the third in the later part of September. The southern of the two hotspot centers presented its first case in the latter part of July, possibly related to the case from early July on the block to the north, then a second case in the first part of August, and the third in early October.

The pattern of spread from the four index cases in Foggy Bottom, particularly the southern most one, is initially east/west and then southward. The index cases in the northern part of this region are weakly linked to the other cases of that area in terms of a temporal progression (Figure 6.2). The directionality of typhoid's spread may be associated with any number of socio-cultural factors. It is these types of questions raised by case-level observed patterns that will drive future research using these data.

Northwest

In contrast to the patterns found in Foggy Bottom, the cases in Northwest do not present an easily identifiable diffusion pattern. New cases appear on spatially separated

151

blocks until the first part of July in 1907 (Figure 6.3). Similar patterns can be observed for 1906 and 1908 (Appendix B). Of the three new cases in this area during the latter part of July 1907 only one developed on or neighboring a block with an existing case(s). Finally, in the second half of August multiple new cases appear on blocks or next to blocks with existing cases (Figure 6.3b). From this point in time through to the end of the study period most new cases develop on or neighboring blocks with existing cases (Figure 6.3c).

Much like the development of the two significant clusters in Foggy Bottom, most of the significant blocks in Northwest in 1907 did not have cases until later in the summer season. Yet, in contrast to Foggy Bottom, multiple cases develop simultaneously in Northwest, rather than scattered through the time periods.

Shifting the focus to the four significant clusters in Northwest that form a north-south line from the 500m cluster south to the southern 150m cluster. The two significant 150m clusters have large number of cases on the hotspot block, three and five north to south respectively (Figure 6.3). On the block with five cases, three of those develop in early August, one in late August, and the final one in early September. To its north, the block with three cases includes one in early September and the remaining two in the later part of that month. Further research could continue with these comparisons, to see whether the variation in case sequence was as a result of a larger population in Northwest or some other social factor.

The sequential maps of this chapter do not currently reveal any new patterns or findings concerning typhoid fever in the early twentieth century. However, they do clearly illustrate that hotspots identified through statistical analysis can, and must, be

investigated in more detail by returning to the original data. Particularly since, as was previously mentioned, analysis at fine temporal aggregations can lead to problems with small numbers. The results of hotspot identification for larger aggregations can still prove useful, however, by guiding local level descriptive investigations of the neighborhoods surrounding each cluster one can glean deeper knowledge of how the disease moved through space and time to create that cluster. The next step to be taken at this neighborhood level is to reconstruct the physical and social geography of these areas in terms of the numbers of houses, alleys, businesses, and the local inhabitants. Socio-economic status and race have all been mentioned as having association with disease patterns, ethnicity can play a role as well in terms of social contacts in most places (Thornton and Olson 2001; Condran and Crimmins-Gardner 1978; Higgs and Booth 1979). In Washington, D.C. the former two variables likely played a more dominant role influencing disease patterns as there was a very limited immigrant population in the city (Borchert 1980).

These results suggest a future line of inquiry: is there a geographic pattern or signature for every area, based on a susceptible population as well as other infrastructure and social conditions? Where a disease enters into an area, in combination with the number of cases as entry points, will affect the temporal sequence of the pattern. The question is, will the pattern still remain the same? To use an analogy, if the area is a concrete mold, pouring concrete into the mold at different rates, and along different sides will affect the speed of the molds completion, but will not affect the molds final shape. The analogy can be further extended to include irregular shaped molds, depending on where the concrete is poured, local shapes will be first filled in (local hotspots will

153

occur). If for some reason the concrete stops pouring (the intervention of a public health initiative or meteorological externality), the resulting mold will be incomplete, with only certain sections completed. However, if enough concrete is poured, the mold will always look the same. This analogy could also be extended to the whole city, the bottom of the mold being those neighborhoods with consistent hotspots through time.

The following chapter returns to the identification of typhoid clusters, but examines how typhoid is distributed in comparison to other diseases in Washington, D.C.

CHAPTER 7
COMPARING TYPHOID FEVER TO DIPHTHERIA AND SCARLET FEVER
IN WASHINGTON, D.C., 1909

Introduction

Typhoid fever morbidity and mortality have previously been analyzed in this dissertation by time and at multiple spatial scales. Until this point, no consideration has been given to how the spatial structure of the typhoid fever outbreaks and/or seasonality compare to the spatial structure of other prevalent early twentieth century urban diseases within the same city. Of particular interest would be the following questions: Even though different diseases have varying etiologies, does the common correlate of poverty and population density lead to similar spatial patterning? Would the weakening of a population by one disease lead to secondary disease outbreaks in the same location, or, does the removal of a susceptible population limit secondary disease occurrence in that location?

The opportunity to investigate such questions is afforded by the 1909 typhoid report which in addition to 12 months of typhoid maps, also contains distributions of diphtheria and scarlet fever in Washington, D.C. For only this year, the PHS doctors added to their own geographical analyses of typhoid cases by making comparisons to the locations of diphtheria and scarlet fever cases. No other reports or discussion of typhoid from this time period made such a comparison, either to these diseases or any other ailments that might have a spatial or temporal association with typhoid (Sedgwick 1893; Whipple 1908; Provincial Board of Health, Ontario 1912; Rosneau et al. 1907, 1908, 1909; Medical Society of the District of Columbia 1894).

The purpose behind including these additional diseases, according to Lumsden and Anderson (1911), was to compare the distribution of typhoid, a water-borne disease, to two diseases spread through contact. The PHS doctors hypothesized that if diphtheria and scarlet fever cases were evenly distributed throughout the city, as they also believed was typhoid's distribution, then it might be possible to conceive another mechanism beyond the water supply that caused the general distribution of typhoid. If their hypothesis was true, then it would lend further support to the idea that typhoid spreads through contact, as well as being carried by water. Alternatively, if the typhoid surface differed greatly from those of diphtheria and scarlet fever then it was more likely that typhoid spread through the city's water supply or through a combination of the water supply and other transmission forms, such as contact, well water, milk, food, and/or flies (Lumsden and Anderson 1911).

The original 1909 report concluded that both diphtheria and scarlet fever maintained uniform patterns across the city, just as was described for typhoid.

> Diphtheria and scarlet fever – diseases not usually regarded as being caused in large part by water-borne infection – also appear to be quite generally distributed in Washington. If only the primary cases in the different residences are considered for all three disease it is evident that the distribution is about as general and nearly, if not quite, as uniform for diphtheria and scarlet fever as for typhoid fever" (Lumsden and Anderson 1911, 34).

As has been shown in chapters 4 and 5, a purely visual examination of the data is not sufficient to be able to identify spatial patterning, and that typhoid was not generally distributed throughout the city during the summer of 1909, as well as during the shorter temporal periods of May to July and August to October, but instead clustered locally.

Therefore, in order to consider both the questions posed at the beginning of this chapter, and the hypotheses originally proposed in the report, the same analysis techniques employed in previous chapters are employed to identify if local clustering occurs amongst the diphtheria and scarlet fever cases.

If all three diseases were to display hotspots in the same locations, then it is possible that a unique combination of urban/social/economic forces were at work creating a particularly unhealthy environment in those locations. Alternatively, if the three diseases display spatially distinctive clustering patterns, then some external factor, whether a variation in etiology, a spatial barrier to one component of a particular disease system, or even a competitive/interactive situation between the two diseases could be at work.

It is unclear why the PHS specifically chose diphtheria and scarlet fever as the two diseases to compare to typhoid. Nowhere in the report is an explanation given other than that both diphtheria and scarlet fever are spread by contact, and typhoid was presumed to spread primarily through the water supply (Lumsden and Anderson 1911). Nonetheless, there are a number of other non-water-borne diseases that could have been used for the comparison such as tuberculosis, pneumonia, or bronchitis. Additionally, the doctors could have considered other gastro-intestinal diseases, such as the generic disease category of diarrheal disease to determine if there was a similar pattern between different types of stomach ailments, possibly lending more credibility to the belief that the general water supply was the primary culprit in the spread of typhoid.

Possible reasons for the particular inclusion of diphtheria and scarlet fever, based upon knowledge of the time and trends in public health, are as follows. Diphtheria was a

157

common disease and considered endemic in most urban areas, and by the 1890s

mandatory bacteriological testing of all suspected diphtheria cases became commonplace

in large cities (Liebanau 1987; Hammonds 1993). Therefore cases were reported to local

health officers, not just deaths. Cities, like New York, mapped reported diphtheria cases

and began to visualize the distribution of the disease (Hammonds 1993). In addition, the

PHS typhoid reports were written in the early years of the bacteriological phase of

American public health and by some of the leading bacteriologists in the country

(Liebanau 1987; Hammonds 1993). During the early years of the bacteriological

movement, diphtheria was the first disease "attacked" and ultimately defeated using new

laboratory tests and treatments (Hammonds 1993; Liebanau 1987). Therefore, diphtheria,

as the first disease truly diagnosed and treated through this new public health approach

was topical. This, combined with the frequency of the disease, and availability of data

allowed for the creation of a comprehensive map of cases for 1909. The choice of scarlet

fever is harder to understand as it was not among the diseases that were well understood

in 1909. Given its similar symptoms to diphtheria, scarlet fever was often diagnosed by a

negative result in a diphtheria test (Hammonds 1993). That is, if the Klebs-Loeffler test

results were negative for the presence of diphtheria, a case might be counted as scarlet

fever (Hammonds 1993). In 1909, there was an epidemic of scarlet fever in Washington,

D.C. that began in October and continued until May 1910 (Health Officer 1911). It was

probably this coincidence that lead to its inclusion, and not because it was considered a

common public health problem (it typically had a low mortality rate). Therefore, it seems

likely that diphtheria and scarlet fever were the chosen contact diseases simply because

for the year in question there was sufficient morbidity data available to make a comparison with typhoid.

Disease Etiology

Diphtheria and scarlet fever, both of which have a bacterial agent, spread through contact between an infected individual and susceptible individuals, although the specific vehicle of transfer for each disease can differ. The initial symptoms of diphtheria and scarlet fever are similar, as they are for many diseases, these being a sore throat and fever. However, the manifestation of these symptoms is similar enough that, prior to the beginning of bacteriological testing for diphtheria in the late nineteenth century, the two ailments were often confused. One of the only means of differentiating between the two diseases was the type of membrane that would develop across the patient's throat, though even this was an imperfect measure as it was not present in all cases of either disease (Hammonds 1993; Brainthwaite et al. 1996). During the nineteenth century both diseases primarily affected children. In a similar fashion to typhoid, by 1909 mortality rates of both diphtheria and scarlet fever had diminished though their morbidity rates remained higher than was considered acceptable (Lumsden and Anderson 1911; Health Officer 1911; Hammonds 1993). To better understand diphtheria and scarlet fever for the purposes of comparison to typhoid fever, the current etiological understanding are provided below.

Diphtheria

Diphtheria, a respiratory infection, primarily affects the nose and throat. The bacteria are spread by contact through respiratory droplets from an infected person. Although it is a relatively rare disease now, diphtheria can appear in areas with

159

population crowding, poor hygiene, and a lack of immunizations, conditions which describe the environment of the 1993 diphtheria epidemic in the former Soviet Union (NIH 2004; Hammonds 1993).

<u>Scarlet Fever</u>

Caused by the bacteria group A streptococcus, a sore throat, fever, headache, and rash, characterize the symptoms of scarlet fever. The disease is transmitted by direct contact between people or by droplets exhaled by an infected person. By the early twentieth century, scarlet fever, which could still cause epidemics, was typically not fatal. This once serious childhood illness now is treated with penicillin and sometimes referred to as scarlatina (NIH 2005; Cliff et al. 1998).

Methods

As previously described in chapter 4 this study explores the conclusions of the PHS doctors, and is intended to determine if indeed diphtheria and scarlet fever were uniformly distributed in 1909 Washington, D.C. The methods used here are primarily the same as those described in chapter 2 and utilized in chapters 4 and 5. The difference here is the inclusion of typhoid morbidity data from all 12 months of 1909. These additional six maps (January to April and November to December) were also digitized as previously described (chapter 2). All 12 monthly typhoid surfaces were merged into a single shapefile for this analysis. The 1909 typhoid report also included a map of diphtheria cases and a map of scarlet fever cases; both maps were digitized at the block-level for compatibility with the typhoid fever surface. Ripley's K-function was used as a measure of global spatial auto-correlation, again as previously described.

The two new datasets, diphtheria and scarlet fever, had not previously been tested for global clustering and the typhoid surface includes additional data thus warranting another test using Ripley's K-function. The G_i* statistic is run for all three disease surfaces using the grid of city blocks and the same distances,150m, 250m, 500m, 750m, and 1000m described in chapter 2.

The histogram analysis presented in chapter 5 to explore the spatio-temporal relationships of typhoid is again employed to compare regional differences between the diseases. Histograms are used in two ways to further explore the geographic distribution of typhoid, scarlet fever, and diphtheria. First, each region (as identified in chapter 4; Figure 1.2) became its own bin for summarizing the number of cases in each region in order to visualize the proportion of the total load each disease contributed for a given region. Second, the contribution of cases from each region to the total number of cases of that disease is charted.

Results

Figures 7.1 through 7.3 illustrate the case distributions for each disease investigated in this chapter and used in the K-function analysis. The results of each type of analysis are described below.

Ripley's K-function

The global spatial auto-correlation test, Ripley's K-function, indicates global clustering for the typhoid fever, scarlet fever, and diphtheria case data (Table 7.1; Figures 7.4 through 7.6). The values for $\hat{L}(h)$ were fairly similar for each of the diseases at all distances, but the graphs of $\hat{L}(h)$ values never intersect with the envelope of complete

161

spatial randomness, regardless of the disease being tested thus displaying significant

global spatial auto-correlation.

Distribution of all Typhoid Cases, 1909

○ Dots represent individual cases

0   1,000
Meters
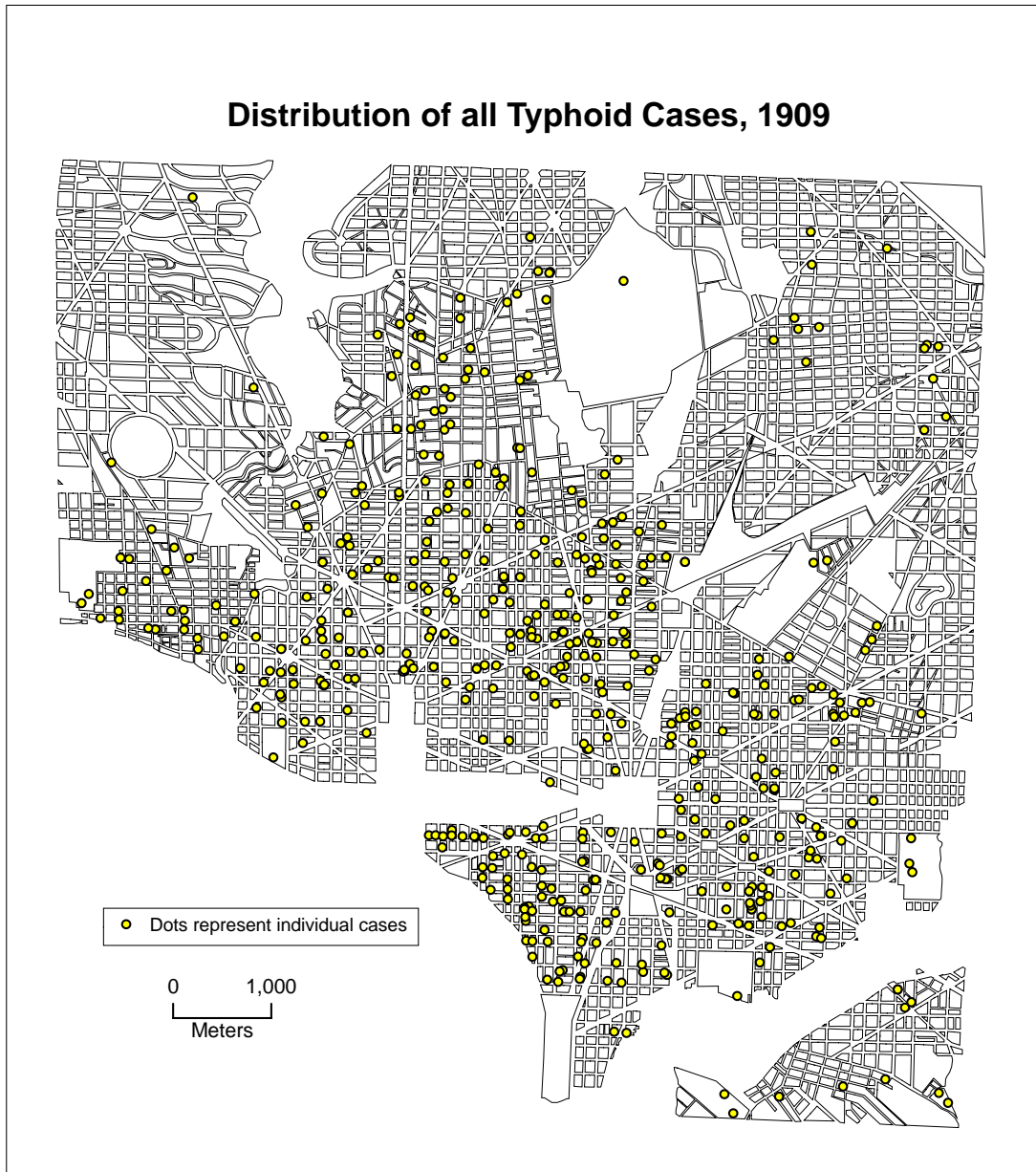
Figure 7.1 All typhoid cases in 1909

Distribution of Diphtheria Cases, 1909

Dots represent individual cases

0   1,000
Meters
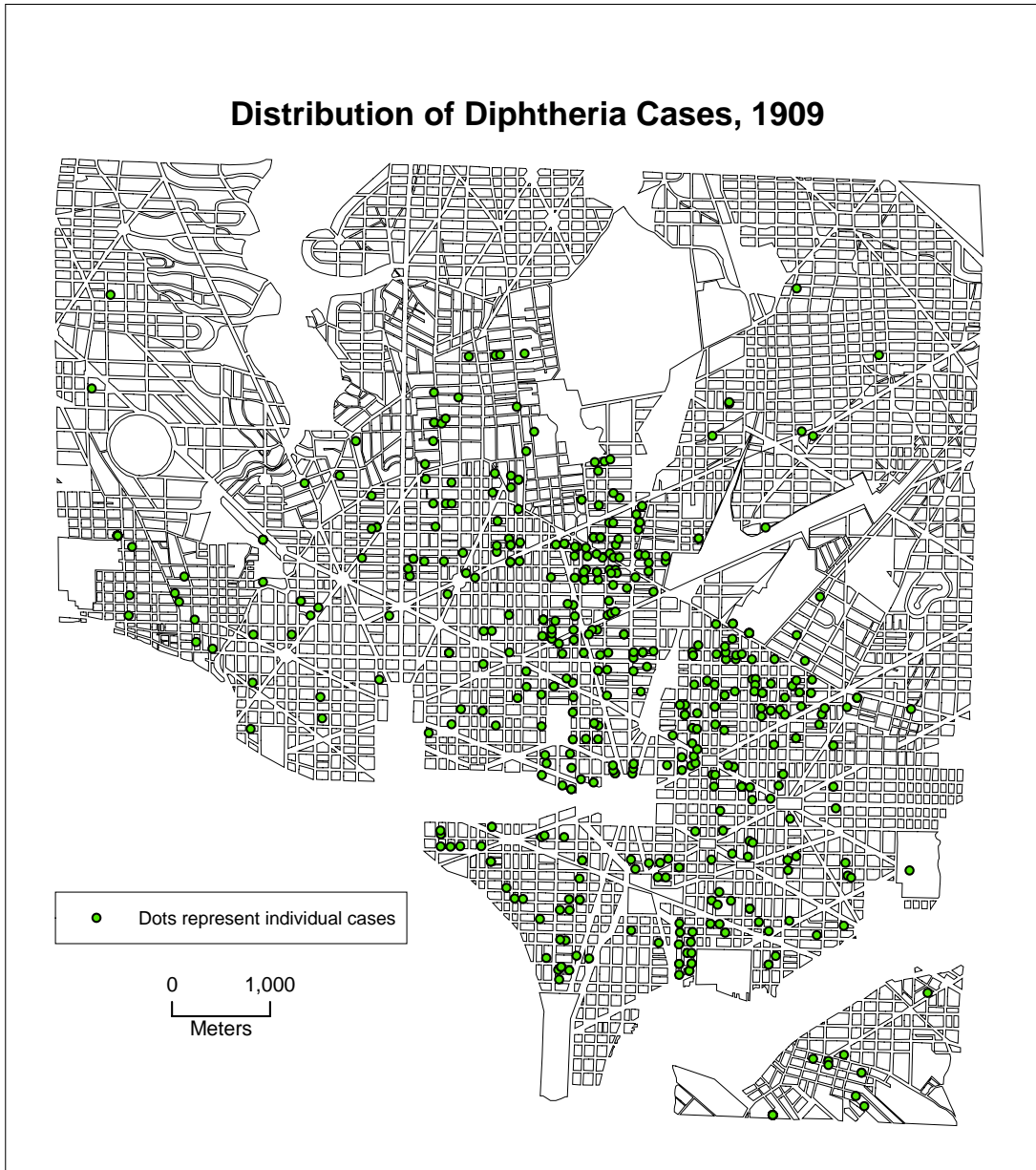
Figure 7.2 All diphtheria cases in 1909
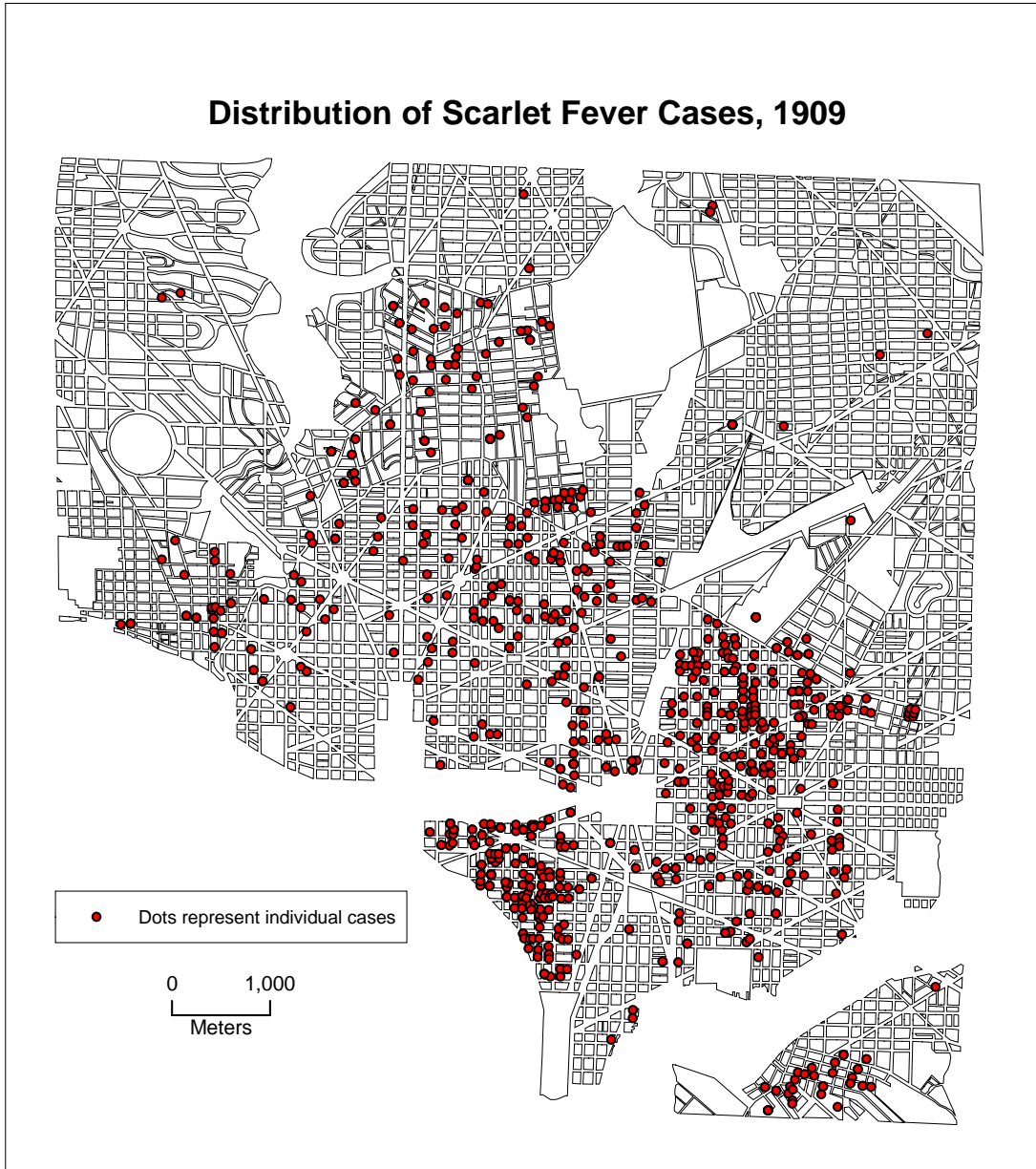
Figure 7.3 All scarlet fever cases in 1909

Table 7.1 Ripley's K-function results for typhoid fever, diphtheria and scarlet fever in 1909

Distance (m)	Typhoid	Diphtheria	Scarlet Fever
0	160.01	274.61	246.69
100	312.53	411.25	417.68
200	457.02	553.76	578.14
300	607.77	696.36	731.43
400	741.20	829.51	878.12
500	869.80	953.26	1026.81
600	990.93	1082.88	1169.38
700	1112.92	1211.43	1297.99
800	1235.92	1340.17	1421.68
900	1358.01	1473.40	1537.38
1000	1473.59	1597.87	1643.39
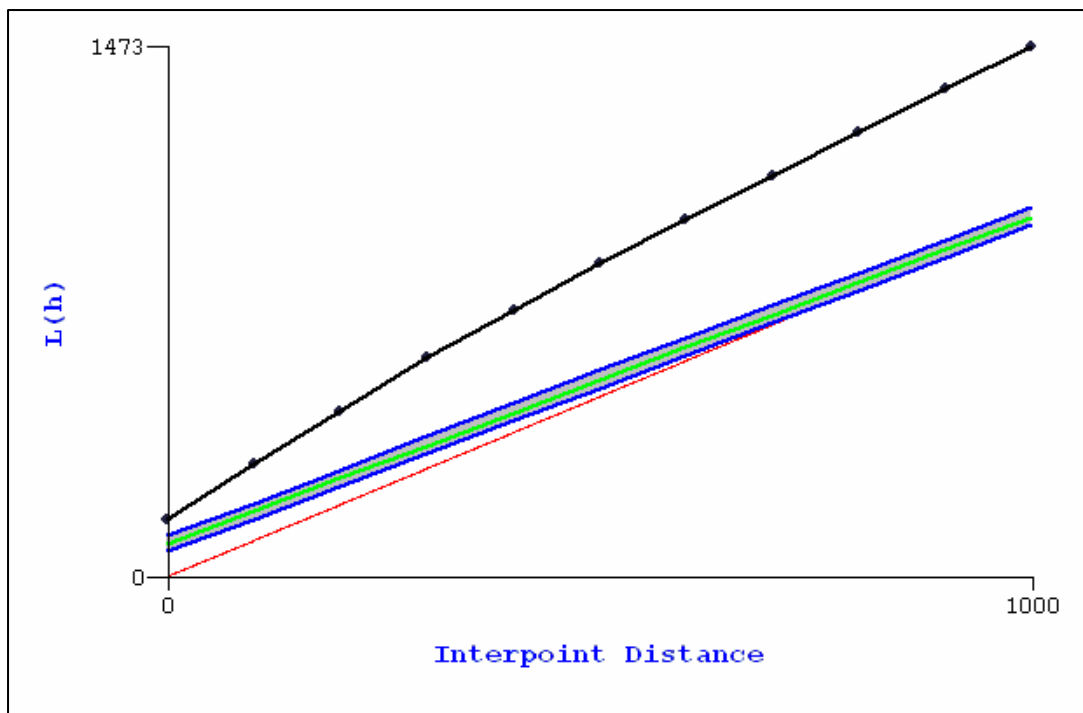
Figure 7.4 Graph of Ripley's K-function results for typhoid fever, 1909
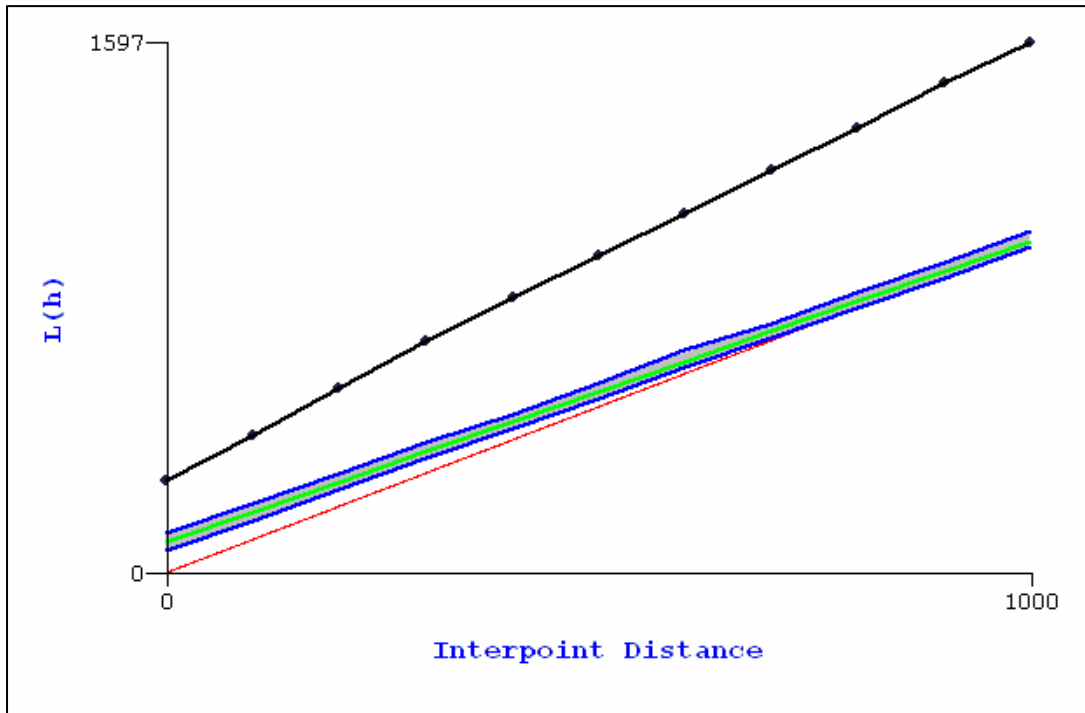
Figure 7.5 Graph of Ripley's K-function results for diphtheria, 1909
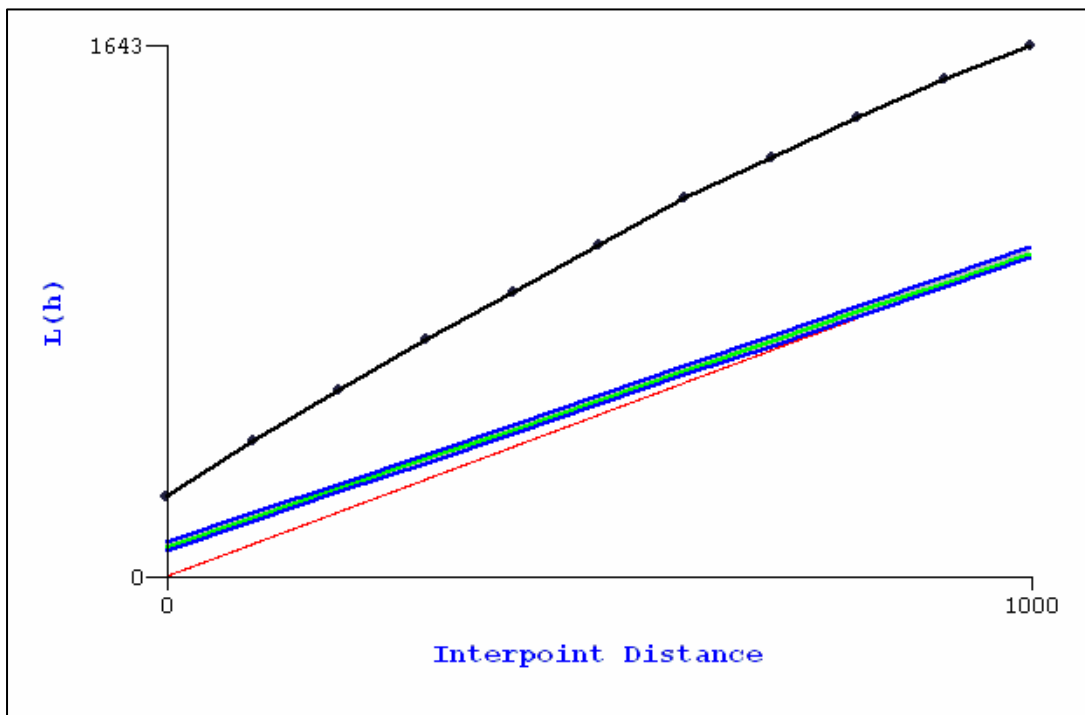
Figure 7.6 Graph of Ripley's K-function results for scarlet fever, 1909

G$_i$ - Typhoid

As expected from the results in previous chapters, local clusters of typhoid are statistically significant when using data for the entire year (12 months). The majority of the clusters are located in Northwest (Figure 7.7). Three groups of significant clusters also appear in Southwest and Southeast. There was a slight skewing of the cluster sizes towards more localized distances, 150m and 250m, but clusters did exist at all five tested distances (Table 7.2).

G$_i$ - Diphtheria

A large group of local diphtheria clusters is located on both sides of Boundary Street in Northwest extending into the Suburbs. Twelve other significant clusters are scattered throughout the city, although Southwest and Southeast are devoid of significant diphtheria clusters (Figure 7.8). The highest *G$_i$** value for this part of the study, or any of the previous *G$_i$** analyses in this dissertation, occurred in Georgetown. This 150m cluster had a recorded *G$_i$** value of 10.40 and is a block that has 15 cases of diphtheria occurring at a single address in 1909. The neighboring block was part of the strongest 250m cluster in the entire study, though this again was largely due to the multiple-case residence and one further observation in a neighboring block (Table 7.2).

G$_i$ - Scarlet Fever

Two distinct groups of localized scarlet fever clusters appear in the city. The first covers almost all of Southwest and the second is located in Northeast along Boundary Street (Florida Avenue), extending into the Suburbs (Figure 7.9). In both of these areas most of the clusters are significant at the two largest distances, 750m and 1000m. A

number of smaller clusters are scattered through the city, beyond the two large clusters (Table 7.2).
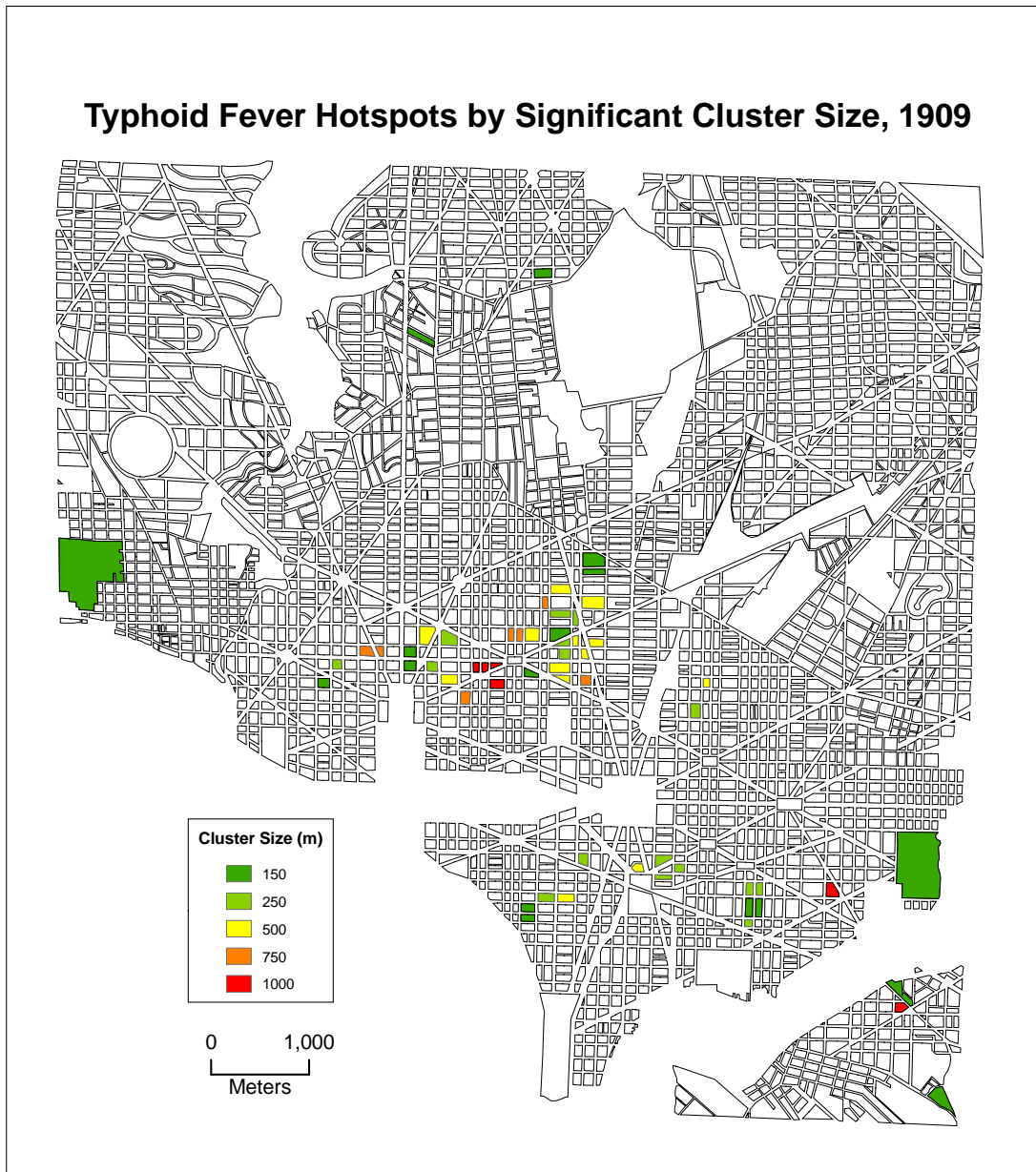
Figure 7.7 Typhoid fever G_i* results, 1909

Table 7.2 Summary of the G_i^* results for typhoid, diphtheria, and scarlet fever in 1909

Disease	Distance (m)	Cells	min G_i^*	max G_i^*
Typhoid				
	150	17	2.88	4.54
	250	16	2.01	3.50
	500	13	2.01	3.11
	750	7	2.04	2.96
	1000	6	2.01	2.65
Scarlet Fever				
	150	15	2.25	4.39
	250	7	2.48	4.45
	500	13	2.33	4.63
	750	28	2.06	4.68
	1000	58	2.01	4.96
Diphtheria				
	150	8	2.69	10.40
	250	3	2.69	7.07
	500	9	2.05	2.48
	750	10	2.04	2.80
	1000	8	2.02	2.53

Diphtheria Hotspots by Significant Cluster Size, 1909

Cluster Size (m)

- 150
- 250
- 500
- 750
- 1000

0   1,000
Meters
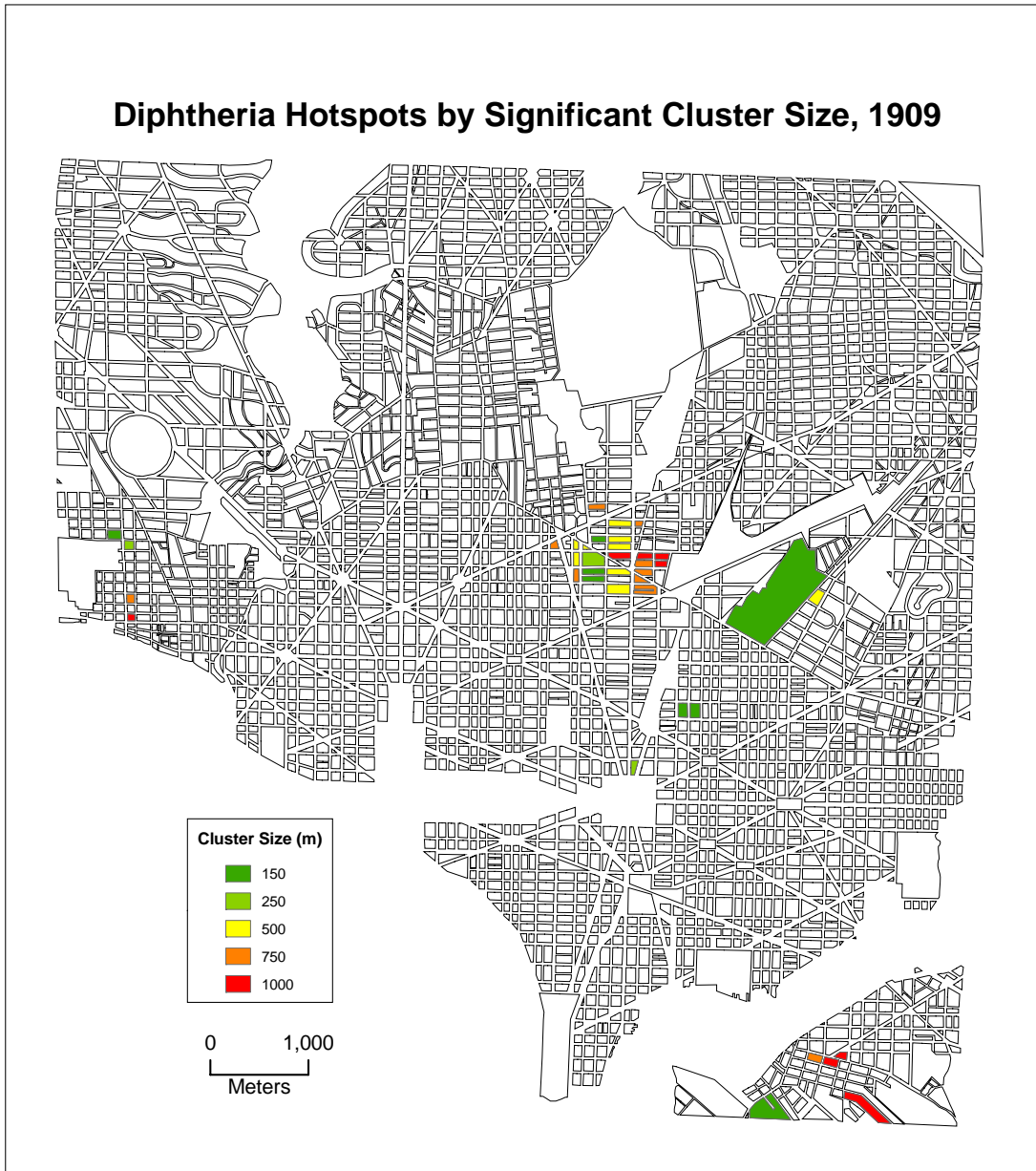
Figure 7.8 Diphtheria G_i* results, 1909

170

Scarlet Fever Hotspots by Significant Cluster Size, 1909

Cluster Size (m)
- 150
- 250
- 500
- 750
- 1000

0   1,000
Meters
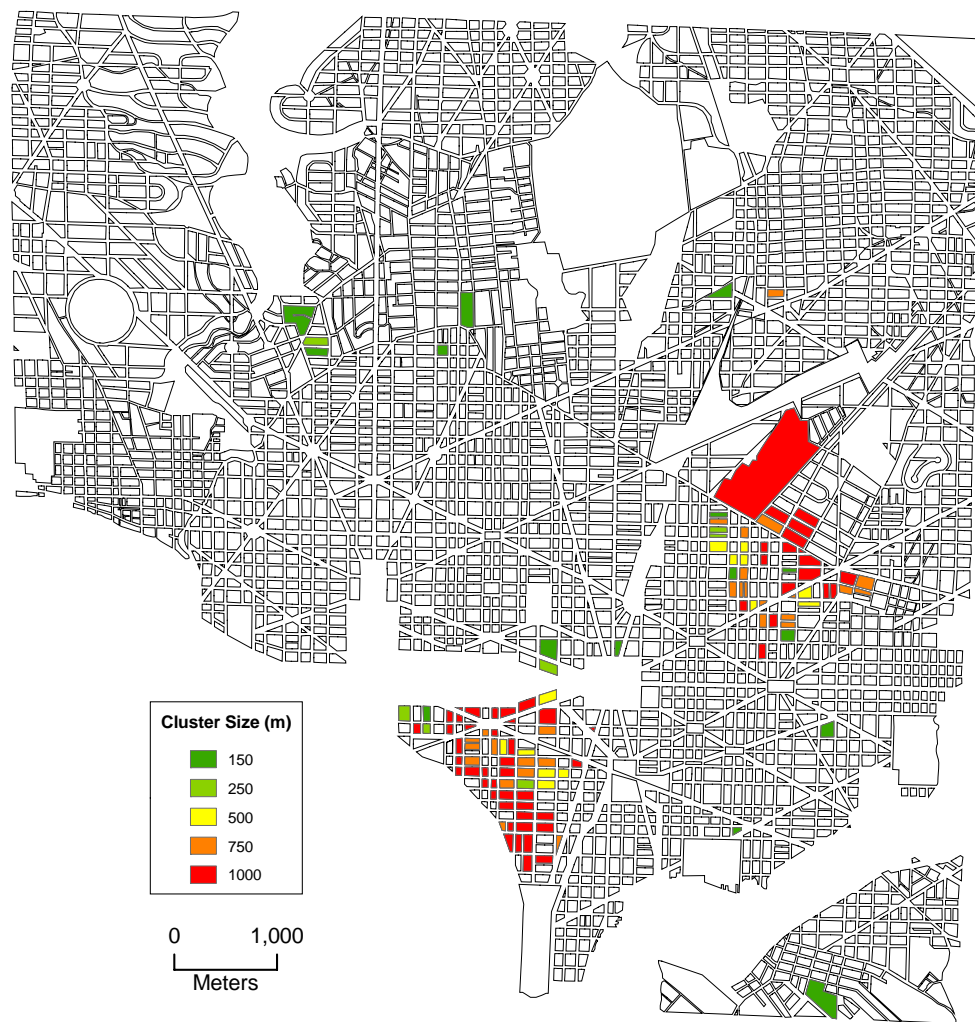
Figure 7.9 Scarlet fever G_i* results, 1909

Only nine blocks are significant for more than one disease. These nine are only significant for combinations of two of the three diseases, and for different cluster sizes. Four of these blocks were in the same part of Southwest and those were all significant clusters of typhoid and scarlet fever (Figure 7.10). No block was part of a significant cluster for all three diseases in 1909.
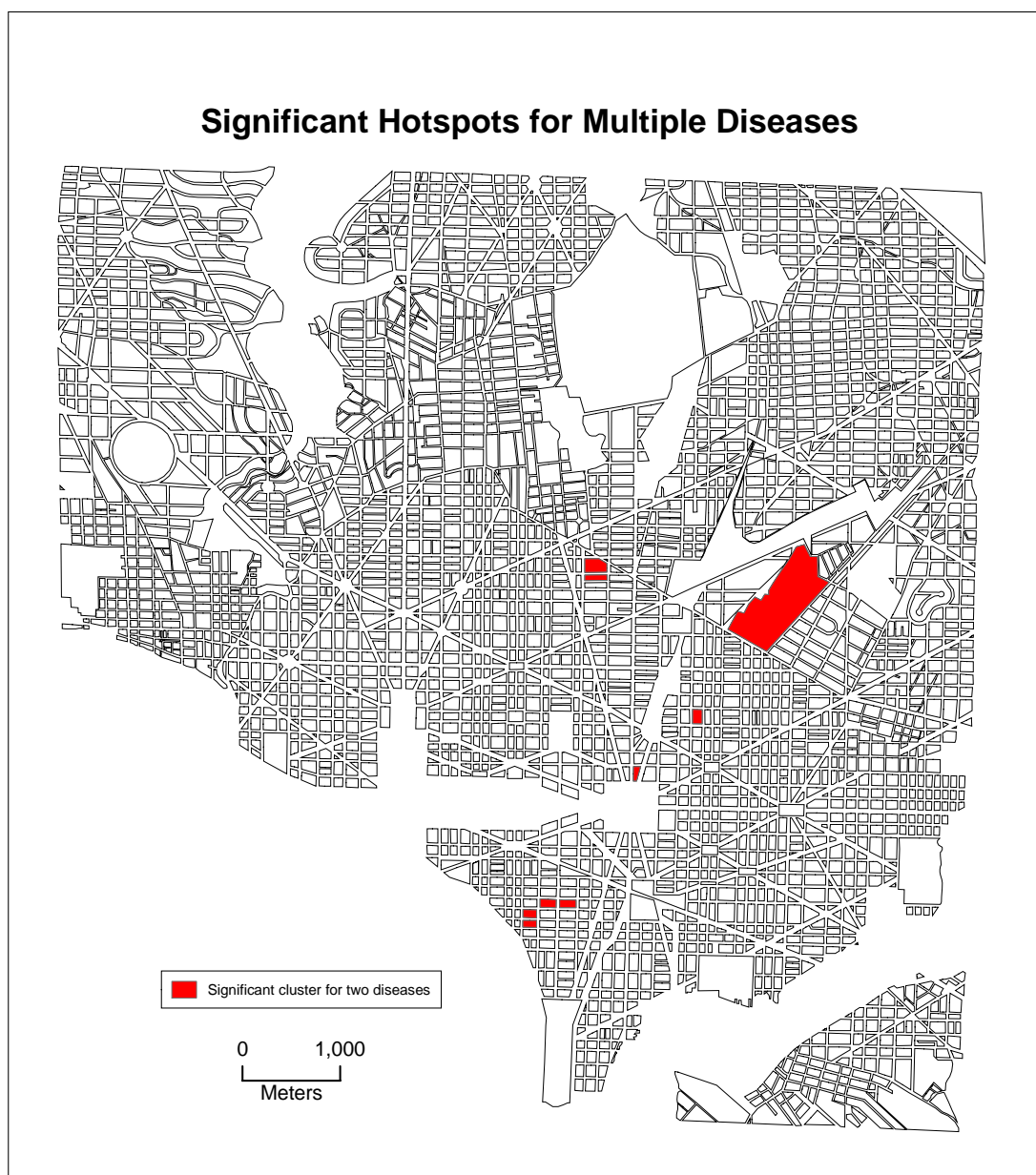
Figure 7.10 Blocks with significant clusters for two diseases

Table 7.3 clearly illustrates that there is approximately twice the total number of scarlet fever cases than either typhoid or diphtheria cases. Some parts of the city were disproportionately impacted by one disease. Of particular note is Foggy Bottom where of the 32 total combined cases for any of the three diseases, more than 60 percent (20 cases) were for typhoid. In Northwest and Georgetown there was an approximately equal distribution of cases between the three diseases, even though city-wide the number of scarlet fever cases was double that of the other diseases. Two other regions of interest are Northeast and Southwest. Northeast was dominated by scarlet fever with four times the number of scarlet fever cases than typhoid cases (Table 7.3; Figure 7.11). In Southwest, the dominant disease was once again scarlet fever, as also seen on the clustering maps, but here there were four times as many scarlet fever cases as diphtheria cases (Table 7.3, Figures 7.2 and 7.11).

Discussion

The use of Ripley's K-function returns the result that these data, regardless of disease type, do cluster globally, but this only presents a need to further explore the data for the location and scale of local clusters within the global clustering patterns. However, as previously stated, the contemporary researchers' ability to identify local clusters was limited.

The results of the G_i* statistic identify local clustering of the diseases as mapped in the 1909 PHS report (Figures 7.7 to 7.9). Each of the three diseases displays a distinctive geography to their clustering patterns that generally does not coincide with the

Table 7.3 Counts of cases by region for typhoid, diphtheria, and scarlet fever in 1909. Percentage of total number of cases are in parentheses

Region	Typhoid Fever	Scarlet Fever	Diphtheria
Southeast	68 (14)	99 (12)	59 (12)
Northeast	52 (11)	209 (25)	108 (22)
Southwest	72 (15)	161 (19)	38 (8)
Georgetown	24 (5)	29 (3)	25 (5)
Northwest	155 (31)	158 (19)	148 (31)
Suburbs	92 (17)	145 (17)	84 (10)
Anacostia	11 (2)	34 (4)	17 (2)
Foggy Bottom	20 (4)	8 (1)	4 (0.5)
Total	494	843	483

clusters of the other two diseases. Upon further investigation, only nine blocks overlap between different significant clusters, but never for all three diseases (Figure 7.10). Therefore, of 200 blocks (of the total 3036 blocks in the city) that are significant cluster members for any one disease, only nine blocks were significant cluster center for two diseases.

In the original report the PHS hypothesized that if scarlet fever and diphtheria were uniformly distributed across the city in a similar way to typhoid, then it was more likely that typhoid diffused through Washington via a different route than the general water supply. The PHS officers ultimately reported that all three diseases were generally distributed throughout Washington, and to some degree this is supported by the findings
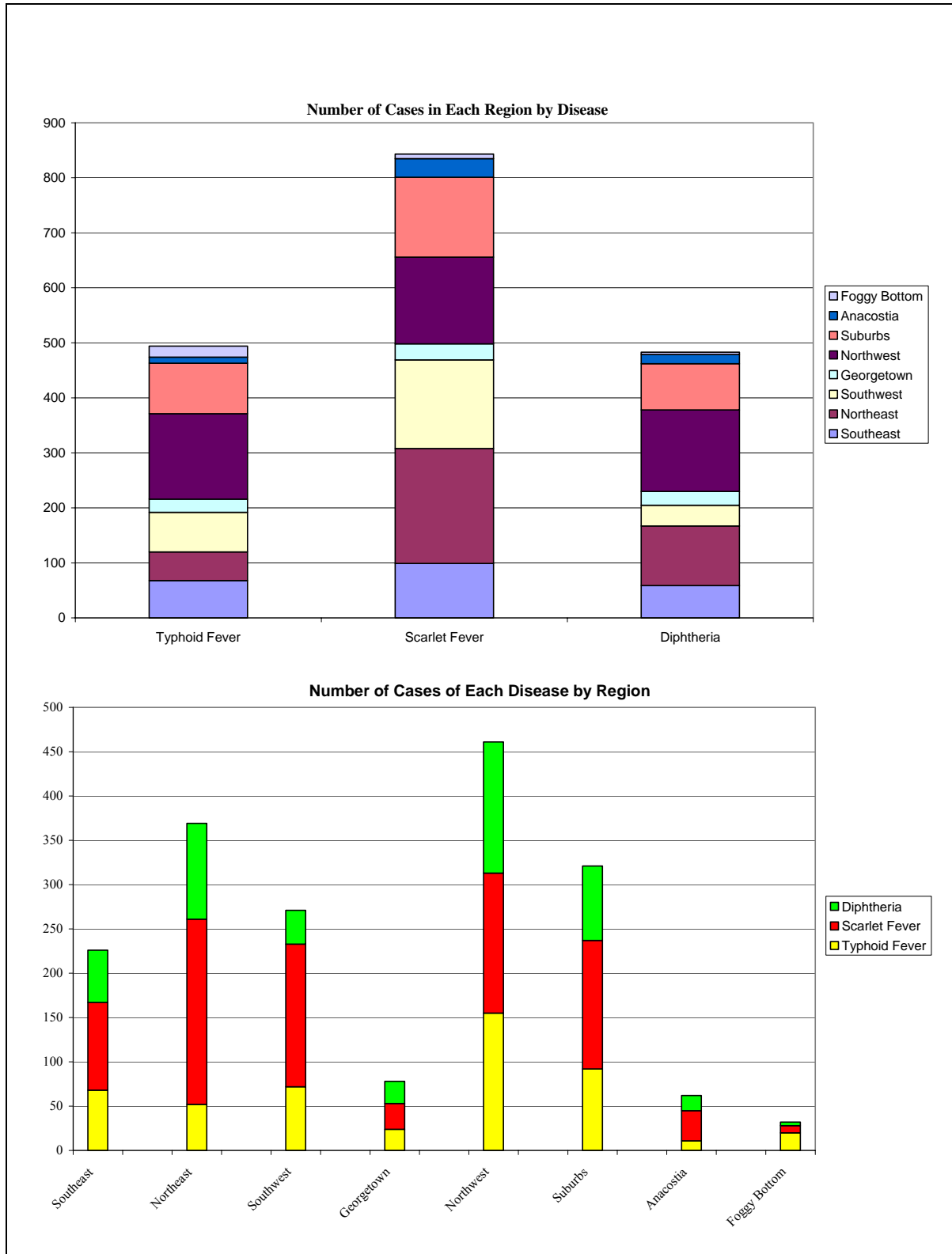
Figure 7.11 Histograms of the typhoid, diphtheria, and scarlet fever case counts

from using Ripley's K-function. Local cluster analysis contradicts these findings that there was no spatial pattern of disease in the city. Instead, the results suggest a neighborhood-level scale to endemic and epidemic diseases spread by contact. In other words, there is a geographic scale to these types of outbreaks, a scale which is most likely defined by neighborhood level interactions.

Although neighborhood factors can play a major role in typhoid spread, water-borne diseases are also likely to reflect an underlying physical structure, such as locations along a river, a milk route, or proximity to a well. The other two diseases are less likely to create such a pattern. Drawing concrete conclusions about the underlying structures and causes of the clustering patterns seen here is difficult. However, it is fairly clear that the general water supply was not the spreading mechanism for typhoid during 1909, or any of the years studied. Nonetheless, typhoid clusters are clearly more generally distributed throughout the city than the diphtheria or scarlet fever clusters.

Widespread infrastructure or lack thereof (ie. lack of sanitary sewers) is needed in order to efficiently disseminate typhoid fever throughout a community. For example, the water taken from the contaminated Merrimac River needed to be distributed through the general water supply for Lowell, Massachusetts in order for typhoid to reach epidemic proportions in 1892 (Sedgwick 1893). Typhoid-contaminated milk must be delivered to multiple homes in a single neighborhood for the disease to diffuse throughout a neighborhood, as was the case in Georgetown in 1908 (Rosenau et al. 1909). In both situations, a form of distributary infrastructure acted as the vehicle for contaminated consumer products.

176

Small localized clusters of typhoid can develop from highly localized transfer points, as well. Flies, for example, are rather inefficient vehicles for transferring typhoid bacteria throughout a large population since they cannot travel far nor can they transport large amounts of typhoid bacteria. Oysters are another inefficient distributor of infection since they were a luxury item and only consumed raw by a small portion of the population. Unless the general water supply becomes involved in diffusing typhoid, the disease is more likely to appear throughout the city as small, highly localized clusters of a few related cases.

Alternatively, both diphtheria and scarlet fever spread efficiently through contact between people, and one person can quickly infect others by simple proximity. One infected individual can have close enough contact with susceptible individuals in any social gathering place to create multiple new cases. Since proximity is the primary vehicle for transferring diphtheria and scarlet fever from person to person the diseases should spread as an expansive wave through a neighborhood until the disease reaches a physical, cultural or biological barrier. Most cases of both diseases occur in children, as parents are likely to be immune due to their own childhood infection. Therefore, social gatherings, cultural contacts, play spaces and physical barriers *of children* are also likely to affect disease spread. Overall, the spatial distinctiveness of the typhoid, diphtheria, and scarlet fever clusters, and in particular the limited spatial extent of the diphtheria and scarlet fever clusters is likely to be related to differences in how the two spreading mechanisms (food/water vs. contact) function in the urban socio-cultural landscape.

By considering the point distributions alone, each disease appears to have its own geography in the city. Diphtheria occurs mostly in the eastern part of the city and more to

the north. Scarlet fever also occurs more in the eastern half of the city and in Southwest. Comparatively, typhoid is more widely dispersed throughout the city, with many more cases appearing towards the west and center of the city (Figures 7.1 to 7.3).

As was found in the analysis of summer typhoid in the previous chapters, the three diseases investigated here vary in the scale of the significant clusters, with hotspots appearing at all distances for all of the diseases. The scarlet fever results differ from the other analyses in the dominance of 1000m clusters. Neither typhoid nor diphtheria display such a strong skewing toward a particular cluster size. Typhoid tends towards the smaller cluster sizes than diphtheria. When comparing the results of the G_i^* statistic for the entire year of typhoid data to those for just the six summer months it is interesting that there is a stronger skewing toward the smallest distances in the summer months.

This tendency towards more localized clusters agrees with the final "conclusions" suggested by the PHS concerning the origins of typhoid fever in the District of Columbia. In the 1909 report Lumsden and Anderson (1911) wrote that the previous reports focused too much on finding a single source of the disease through a process of elimination, and in so doing missed the possibility of multiple origins of typhoid. This concept of multiple origins could help explain the lack of global clustering of the disease, and tendency towards smaller cluster sizes. According to the 1909 report many of the causes of typhoid, contaminated oysters or contact, for example, would affect two or three people living in the same household or neighboring households. Therefore, if similar events occurred throughout the city this could explain the spatial pattern of typhoid with small clusters being scattered throughout the city.

The hypotheses of the PHS doctors when they compared the three diseases were that if scarlet fever and diphtheria cases shared a similar uniform distribution across the city to those of typhoid, despite the different means of transmission, then maybe the city's water supply was not the source of the generalized typhoid distribution. In the previous reports as no other explanation could be found for most of the general distribution of typhoid, the typhoid investigators attribute the source of typhoid to the city's water supply despite a filtration system and tests showing the water to be clean. The PHS doctors describe scarlet fever and diphtheria as nearly as uniformly distributed as typhoid in 1909, therefore ruling out the water supply theory. They discuss in the report the possibility that other factors besides the general water supply, including flies, contact, milk, and shellfish may have facilitated the continuation of typhoid's impact on the city following the implementation of water filtration in 1905 (Lumsden and Anderson 1911).

Lumsden and Anderson (1911) elaborate upon this idea of other possible sources of transmission in their report by stating that too much emphasis had been placed on finding a single source and means of distribution for typhoid, and that possibly the continued presence of typhoid in Washington was the combination of small contributions from a number of potential spreading mechanisms. For example, the 1909 report mentions how flies and soil pollution were not considered important contributors to the continued presence of typhoid as suggested in earlier reports because the removal of privies did not lead to an end of typhoid in those areas. Yet, in the discussion of flies and privies in the 1909 report, a possible connection was made between an improperly

cleaned privy and a nearby contaminated well[7] that caused a few cases of typhoid.

Lumsden and Anderson suggest:

> … from a careful consideration of all the facts now at hand it
> seems not only possible, but highly probable that all the things
> mentioned, viz, Potomac River water, polluted well water,
> shellfish, flies and soil pollution, milk (and other foods), and
> personal contact, have operated as very considerable factors and
> every one of them perhaps has operated on different occasions
> and for longer or shorter periods as the chief factor in the
> distribution of typhoid-fever in Washington (Lumsden and
> Anderson 1911, 45).

This admission of multiple sources of typhoid helps explain the geography of the disease

in Washington not only in 1909, but in the three other years investigated by the PHS. At

the same time, this idea of multiple origins and spreading mechanisms raises questions

about how to eradicate typhoid in twenty-first century cities that are undergoing rapid

urbanization after those cities build water treatment facilities.

Diphtheria and scarlet fever were only briefly discussed in the 1909 PHS report

regarding their geographic pattern in the city. The annual health officer's report for the

District of Columbia offers little to supplement the information provided by the PHS

concerning these diseases. From the health officer's report we learn the racial breakdown

of each disease, their morbidity and mortality rates, that there were a significant number

of cases in certain institutions, and that scarlet fever was epidemic in 1909-10 beginning

in October 1909. From the exploratory histograms insight can be gained into the disease

load of each region. These histograms alone indicate a lack of uniformity in the point

distributions of each disease. What is particularly interesting is the lack of typhoid fever

in Northeast (Figure 7.1, 7.7). This disease had a minimal presence in that region

---

[7] All shallow public wells and most deep wells were closed before 1909, but an unknown, albeit small,
number of private wells remained in use in 1909 (Lumsden and Anderson 1911)

throughout this study, excepting the 1895 clusters (Figure 3.4), and was mostly attributed to its smaller population size. Yet, the histograms and Table 7.3 clearly show a strong scarlet fever presence in that region with 209 cases. So why was this part of the city, which clearly is not free from disease, only a minor to contributor to the lingering typhoid problem? Future research may help to shed light upon the barriers to typhoid in Northeast.

Southwest poses a similar question, but for diphtheria. Why would there be a large group of 1000m scarlet fever clusters and a few small typhoid clusters, but no clusters and relatively few cases of diphtheria in this region, especially as the reverse should be expected based upon population density alone? Since surviving a case of diphtheria and scarlet fever provides a person with immunity to that disease, could a previous diphtheria epidemic in Southwest have reduced the size of the susceptible population to an extent that few new cases could develop in this area? It would be fascinating to investigate the patterns of scarlet fever and diphtheria cases for the other three years studied in this dissertation, but the data are unavailable for such an investigation. Despite the limitation in terms of a lack of other diphtheria and scarlet fever datasets for the period studied by the PHS and a lack of temporal information for the two contact-based diseases, the results presented here display unexpected clustering patterns. These clustering patterns will act as a foundation upon which future historical disease surface research can be built.

CHAPTER 8
CONCLUSIONS:
INTREGRATING HISTORICAL DISEASE GEOGRAPHY
INTO URBAN ENVIRONMENTAL RESEARCH

This dissertation sought to explore the spatial distribution of typhoid fever at a

highly localized urban scale in order to gain more insight into the geography of a water-

borne disease at the turn of the twentieth century. This study adds to the growing

literature using GIS and spatial analysis to investigate historical datasets, and in particular

historical medical datasets. The G_i* statistic proved to be a useful and effective tool for

identifying local hotspots of typhoid that otherwise are not visible to the un-aided eye. In

identifying these local clusters and interpreting their occurrence, new questions arose,

some of which were explored in the preceding pages (especially chapter 6) and others

will form the starting point for a future research agenda focused on urban health at the

turn of the twentieth century and its connections to changes in the urban environment.

Five separate studies presented in this dissertation related by the themes of typhoid,

Washington, D.C., and the opportunity to investigate data at a rarely attempted localized

scale are the center of this project. A number of conclusions were drawn throughout these

five chapters and are summarized below.

First, historical datasets are most certainly a useful resource for understanding the

geography of disease. One of the most important functions that historical data serve when

considering questions of medical geography is to allow researchers to work with

information from an entire outbreak or epidemic. As the results of chapters 5 and 6

illustrate, disease clusters and patterns develop over the span of an outbreak, and a study

of the final disease pattern, though revealing in itself, may miss several key social and

spatial components of the outbreak. For example, what pathways and barriers were integral to the creation of the final disease surface? Although traditional diffusion studies approach these topics, the spatial resolution of the data analyzed in this dissertation is relatively rare. In addition, the morbidity and mortality surfaces from 1895 (chapter 3) provided an exceptional opportunity to investigate epidemic data spatially. Usually only mortality *or* morbidity data for a given epidemic event are available during this time period, and more particularly at the residence-level scale recorded by Dr. Kober. While, local clusters could be identified in both types of health data, Figures 3.3 and 3.4 clearly illustrate that, when available, morbidity data produce more telling results of disease clustering and disease patterning even given some of the limitations to morbidity assessments. This chapter should also be considered a contribution to the medical history literature and the debate between morbidity and mortality data. More of these types of data are available, though in general they have only been analyzed for larger geographic spaces (Elman and Myers 1997, 1999; Wilson 1993), therefore the block-level investigation of the 1895 epidemic is rare.

In addition to enabling the study of epidemics and outbreaks as complete events, the analyses used in this dissertation illustrate that the G_i^* statistic is an effective means of identifying local-scale clusters in historical datasets that visually do not display clustering patterns. The authors of the five reports (1895, 1906, 1907, 1908, 1909) could not identify specific spatial clusters of typhoid fever (or diphtheria or scarlet fever) using their traditional visual analysis of dot maps. Through a re-examination of the same data, modern epidemiological approaches in a GIS environment identified the presence of local clusters. It would have been virtually impossible for the original PHS report authors to

identify such local patterning within a larger general distribution of disease. These local

area clusters varied in terms of their size and location for the years studied. Both of these

features are interesting in many of the ways described in the previous chapters, but to

repeat the most important of these, the lack of size and temporal stability is a strong

indication that the origins of typhoid varied across the city and over time, which is also

suggested in the 1909 report (Lumsden and Anderson 1911). With a water filtration

system in place, the water distributed through the city of Washington, D.C. was almost

certainly free of typhoid bacteria by 1906, but this does not mean that the remainder of

the Capitol City's environment was free of typhoid. Milk contamination clearly continued

to play a significant role in the continuation of typhoid fever's presence in Washington.

Some regulations were in place to prevent dairies from selling contaminated milk, but

these were not enforced until after 1910 (Lumsden and Anderson 1911). Milk

pasteurization, the licensing of dairies, and regulation enforcement helped eliminate

diseases that could be spread through milk contamination (Whipple 1908; Atkins 1992).

Minor originating sources of typhoid, such as oysters and house flies, which would

sometimes facilitate the development of a small localized cluster for individual years

were eliminated as problems through greater sanitization of the local environment. This

included the abolition of privies and the adjustments of sewerage systems so that raw

sewage was not deposited into local rivers too close to the city from where it came

(Boone 2003).

Finally, probably the most interesting results generated through the use of the G_i*

statistic in this dissertation are the spatially distinct clustering patterns of typhoid,

diphtheria, and scarlet fever in 1909. This chapter not only identified that diphtheria and

scarlet fever also clustered locally, again a finding not described in the original reports, but their spatial separation from one another raises questions needing further study. As the literature review in chapter 1 described, disease and health vary over space, and traditionally socio-economic status plays a large role in these spatial patterns. Therefore, chapter 7 (disease comparison) hypothesized that typhoid, diphtheria, and scarlet fever would display hotspots in approximately the same parts of the city, thus highlighting the geography of poverty as much as of disease. Instead, the three diseases displayed clusters in different parts of the city. This raises questions about the influences of the socio-cultural geography of the city and the social interactions/pathways influencing these three diseases in early twentieth century cities. Additionally, while it is entirely impossible to prove at this time, these results may suggest some element of spatial competition between typhoid, diphtheria, and scarlet fever, and this is most certainly an avenue for future epidemiological research.

The findings of this doctoral work would have provided even more insight if population data had been available. This would have allowed for the study of typhoid fever rates rather than creating surfaces of typhoid presence. One of the key reasons why many historical disease studies do not include this type of highly localized scale is the lack of available population numbers at a city-block level. However, the results of the cluster analysis can be used to improve upon this situation by acting as a means to focus further inquiry into a geographic space of several city blocks rather than for the entire city. At this geographic scale, population information from sources such as the manuscript census can be used to re-construct neighborhood structure. This approach of

neighborhood "hotspot" investigation and control could also provide a more suitable analytical frame.

Another limitation in this study is the small number of cases on many blocks. As typhoid fever morbidity decreased so too did the available data for this study. To avoid violating the statistical assumptions of the G_i^* statistic, adjustments were made to the proposed analysis in chapter 5, which limited the temporal analysis of clustering to 3-month intervals. Due to these small numbers, the results presented in this dissertation must be interpreted with caution. Chapter 6 introduced neighborhood-level investigations of the original points, taking a visualization-based approach to display how the small number of cases diffused through a neighborhood and developed into hotspots.

Despite the limitations in this study, when interpreted with caution, the results of the statistical and visual analyses presented here provide views of typhoid fever in an early twentieth century city previously not considered. These results fit in with the types of localized urban environmental research being conducted for other cities, like Baltimore, Maryland (Boone 2003; Hinman 2002). The Baltimore studies have focused on sewerage system development, the political decisions behind the building of infrastructure, the influence of this infrastructure on residential geography, and urban land use in general to gain greater insight into how past environmental conditions and mitigation decisions impacted (and can still impact) human health. These environment-based and health-based decisions of the past continue to influence the present urban landscape.

Since typhoid fever and health inequalities were not unique to Washington, D.C. in the late nineteenth and early twentieth centuries, other research avenues to explore

186

might include a comparison between Washington and other cities during the same time period. This might be particularly interesting if the comparison of spatial disease patterns was to a typhoid epidemic of known origin, such as the water-borne typhoid epidemic in Lowell, Massachusetts (Sedgwick 1893). Similarly, the results of the 1909 comparison of typhoid, diphtheria, and scarlet fever were particularly intriguing and while similar data do not exist for another year in Washington, D.C. they may exist for other cities. It would be interesting to further investigate the possibility of spatial competition between diseases.

The use of GIS in historical studies and localized historical medical studies is currently in its infancy. The importance of these studies, however, is not restricted to the disciplines of history and historical geography. Insights can be gained into how urban epidemics form and spread. Current epidemiology can benefit from the addition of macro space, neighborhood structure and cultural diffusion patterns. Not only may these findings aid cities in parts of the developing world, where diseases such as typhoid remain a problem, but the approaches may also prove to be useful for North American urban environments and the re-emergence of pathogens from the nineteenth century. For, example, a debate surrounds the development of strategies for addressing treatment and prevention should another influenza pandemic develop. Researchers are relying upon records, data, and mitigation techniques employed during the 1918 influenza pandemic to inform current decision-making and planning for the next potential flu pandemic (NPR 2007). The type of data, techniques used in, and results from this dissertation can also be utilized in a similar fashion for decision-making concerning potential typhoid fever or other water-borne disease outbreaks in the present.

It has already been suggested that more urban detail can be included in future investigations, such as sewers, well locations, the location of doctor's offices and hospitals, garbage collection routes, and even property value. Additionally, population data from the 1910 manuscript census can help to provide individual level population characteristics that are missing from the current investigation. Detailed block-level urban reconstructions are time consuming to create in a GIS, but the potential results to be gained from such a study would be tremendous. Possibly the most exciting avenue of investigation in terms of these localized analyses is the ability to render the historical city in three-dimensions, which in addition to interactive animations will allow for researchers to walk through these historic epidemics. This dissertation provides the first step on the path towards that type of urban epidemic reconstruction by identifying neighborhoods of particular disease interest.

References

Alter, G., and J. C. Riley. 1989. Frailty, sickness, and death: Models of morbidity and mortality in historical populations. *Population Studies* 43 (1):25-45.

Atkins, P. J. 1992. White poison?: The social consequences of milk consumption in London, 1850-1939. *Social History of Medicine* 5 (2): 207-227.

Bailey, N. 1975. *The Mathematical Theory of Infectious Diseases and its Applications*. New York: Hafner.

Boone, C. G. 2003. Obstacles to infrastructure provision: The struggle to build comprehensive sewer works in Baltimore. *Historical Geography* 31:151-168.

Boone, C. G., and A. Modarres. 2006. *City and Environment*. Philadelphia: Temple University Press.

Borchert, J. 1980. *Alley Life in Washington: Family, Community, Religion, and Folklife in the City, 1850-1970*. Chicago: University of Illinois Press.

Bourdain, A. 2001. *Typhoid Mary: An Urban Historical*. New York: Bloomsbury.

Braithwaite, C., P. Keating, and S. Viger. 1996. The problem of diphtheria in the Province of Quebec: 1894-1909. *Histoire Sociale/Social History* 29 (57):71-95.

Brock, T. D. 1988. *Robert Koch: A Life in Medicine and Bacteriology*. New York: Springer-Verlag.

Brody, H., M. R. Rip, P. Vinten-Johansen, N. Paneth, and S. Rachman. 2000. Map-making and myth-makinng in Broad Street: The London cholera epidemic, 1854. *The Lancet* 356:64-68.

Casner, N. 2001. "Do it now!" Yakima, Washington and the campaign against rural typhoid. *American Journal of Public Health* 91 (11):1768-1775.

Center for Disease Control. 2005a. Disease Information: Typhoid Fever. http://www.cdc.gov/ncidod/dbmd/diseaseinfo/typhoidfever_g.htm. Last accessed 24 June 2005.

_____. 2005b. Disease Information: Cholera. http://www.cdc.gov/ncidod/dbmd/diseaseinfo/cholera_g.htm. Last accessed 30 September 2005.

Cliff, A., P. Haggett, and M. Smallman-Raynor. 1998. *Deciphering Global Epidemics: Analytical Approaches to the Disease Records of World Cities, 1888-1912, Cambridge Studies in Historical Geography*. New York: Cambridge University Press.

Cliff, A. D., P. Haggett, and J. K. Ord. 1986. *Spatial Aspects of Influenza Epidemics*. London, England: Pion Limited.

Cliff, A. D., P. Haggett, J. K. Ord, and G. R. Versey. 1981. *Spatial Diffusion: An Historical Geography of Epidemics in an Island Community*. New York: Cambridge University Press.

Colten, C. E. 2002. Basin Street blues: drainage and environmental equity in New Orleans, 1890-1930. *Journal of Historical Geography* 28 (2):237-257.

Condran, G. A., and E. Crimmins-Gardner. 1978. Public health measures and mortality in U.S. cities in the late nineteenth century. *Human Ecology* 6 (10):27-54.

Condran, G. A., and E. Crimmins. 1980. Mortality differentials between rural and urban areas of states in the northeastern United States, 1890 - 1900. *Journal of Historical Geography* 6 (2):179-202.

Condran, G. A., H. Williams, and R. A. Cheney. 1984. The decline in mortality in Philadelphia from 1870 to 1930: The role of municipal services. *Pennsylvania Magazine of History and Biography* 108 (2):153-177.

Craddock, S. 1998. Tuberculosis, tenements and the epistemology of neglect: San Francisco in the nineteenth century. *Ecumene* 5 (1):53-80.

Crane, B. D. 2000. Filth, garbage, and rubbish: Refuse disposal, sanitary reform, and nineteenth-century yard deposits in Washington, D.C. *Historical Archaeology* 34 (1):20-38.

Cromley, E. K., and S. L. McLafferty. 2002. *GIS and Public Health*. New York: The Guilford Press.

Curson, P. H. 1985. *Times of Crisis: Epidemics in Sydney, 1788-1900*. Sydney, Australia: Sydney University Press.

Duffy, J. 1968. *A History of Public Health in New York City, 1625-1866*. Volume 1. New York: Russell Sage Foundation.

———. 1974. *A History of Public Health in New York City, 1866-1965*. Volume 2. New York: Russell Sage Foundation.

Durbeck, H., D. Greiling, L. Estberg, A. Long, and G. Jacquez. 2002. ClusterSeer: Software for Identifying Event Clusters User's Guide: TerraSeer, Inc.

Edling, C., and F. Liljeros. 2004. Social contact patterns and disease dynamics. *Real Instituto Elcano de Estudios Internacionales y Estrantegicos* Working Paper (No. 19):1-17.

Elman, C., and G. Myers. 1997. Age- and sex- differentials in morbidity at the start of an epidemiological transition: Returns from the 1880 U.S. Census. *Social Science and Medicine* 45 (6):943-956.

Elman, C., and G. C. Myers. 1999. Geography morbidity differentials in the late nineteenth-century United States. *Demography* 36 (4):429-443.

Emch, M. 1999. Diarrheal disease risk in Matlab, Bangladesh. *Social Science and Medicine* 49:519-530.

Frenkel, S., and J. Western. 1988. Pretext of prophylaxis? Racial segregation and malarial mosquitoes in a British tropical colony: Sierra Leone. *Annals of the Association of American Geographers* 78 (2):211-228.

Gandy, M. 2002. *Concrete and Clay: Reworking Nature in New York City*. Cambridge, Massachusetts: The MIT Press.

Gatrell, A. C. 2002. *Geographies of Health*. Malden, Massachusetts: Blackwell Publishers, Ltd.

Getis, A., and J. Aldstadt. 2004. Constructing the spatial weights matrix using a local statistic. *Geographical Analysis* 32:90-104.

Getis, A., A. C. Morrison, K. Gray, and T. W. Scott. 2003. Characteristics of the spatial patterns of the dengue vector, *Ades aegypti*, in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene* 69 (5):494-505.

Getis, A., and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24 (3):189-206.

———. 1996. Local spatial statistics: an overview. In *Spatial Analysis: Modeling in a GIS Environment*, eds. P. Longley and M. Batty, 261-277. Cambridge, UK: GeoInformation International.

Goodchild, M. 1986. *Spatial Autocorrelation, COMTOG, 47*. Norwich, England: Geo Books.

Green, C. M. 1963. *Washington: Capital City, 1979-1950*. Princeton, NJ: Princeton University Press.

Gregory, I. N., and P. S. Ell. 2005. Analyzing spatiotemporal change by use of National Historical Geographical Information Systems: Population change during and after the Great Irish Famine. *Historical Methods* 38 (4):149-167.

Groves, P. A. 1973-74. The Development of a black residential community in Southwest Washington: 1860-1897. *Records of the Columbia Historical Society* 49:260-275.

———. 1974. The "Hidden" Population: Washington alley dwellers in the late nineteenth century. *The Professional Geographer* 26 (3):270-278.

Groves, P. A., and E. K. Muller. 1975. The evolution of black residential areas in late nineteenth-century cities. *Journal of Historical Geography* 1 (2):169-191.

Gumprecht, B. 1999. *The Los Angeles River: Its Life, Death, and Possible Rebirth*. Baltimore: The Johns Hopkins University Press.

Hägerstrand, T. 1953. *Innovation Diffusion as a Spatial Process*. Translated by A. Pred, 1967. Chicago: University of Chicago Press.

Haines, M. R. 1995. Socio-economic differentials in infant and child mortality during mortality decline: England and Wales, 1890-1911. *Population Studies* 49 (2):297-315.

Hammonds, E. M. 1993. *Childhood's Deadly Scourge: The Campaign to Control Diphtheria in New York City, 1880-1930*. Baltimore: The Johns Hopkins University Press.

Harris, R. 1996. *Unplanned Suburbs: Toronto's American Tragedy 1900 to 1950*. Baltimore: The Johns Hopkins University Press.

Higgs, R., and D. Booth. 1979. Mortality differentials within large American cities in 1890. *Human Ecology* 7 (4):353-370.

Hinman, S. E. 2002. *Urbanization and Public Health: A Study of the Spatial Distribution of Infant Mortality in Baltimore, Maryland, 1880.* Masters Thesis, Geography, Ohio University, Athens, Ohio.

Hinman, S. E., J. K. Blackburn, and A. Curtis. 2006. Spatial and temporal structure of typhoid outbreaks in Washington, D.C., 1906-1909: evaluating local clustering with the Gi* statistic. *International Journal of Health Geographics* 5 (13).

Jordan, E. O. 1909. *The Text-book of General Bacteriology*. Philadelphia: W. B. Saunders Company.

Kellogg, J. 1977. Negro urban clusters in the Postbellum South. *Geographical Review* 67 (3):310-321.

Kenny, J. T. 1995. Climate, race, and imperial authority: The symbolic landscape of the British Hill Station in India. *Annals of the Association of American Geographers* 85 (4):694-714.

Knowles, A. K. 2000. Introduction. *Social Science History* 24 (3):451-470.

———, ed. 2002. *Past Time, Past Place: GIS for History*. Redlands, California: ESRI Press.

Kober, G. M. 1895. Report on the Prevalence of Typhoid Fever in the District of Columbia. In *Report of the Health Officer to the Commissioners of the District of Columbia for the Year Ended June 30, 1895*. Washington, D.C.: Government Printing Office.

Layton, L. 2006. What to do with what lies beneath? Catacombs from old D.C. filtration system pose unique development challenges. *Washington Post*, Saturday April 22, B01.

Leavitt, J. W. 1992. "Typhoid Mary" strikes back bacteriological theory and practice in early twentieth-century public health. *Isis* 83 (4):608-629.

———. 1996. *Typhoid Mary: Captive to the Public's Health*. Boston: Beacon Press.

Lee, C. H. 1991. Regional inequalities in infant mortality in Britain, 1861-1971. *Population Studies* 45 (1):55-65.

Lessoff, A. 1994. *The Nation and its City: Politics, "Corruption," and Progress in Washington, D.C., 1861-1902*. Baltimore: The Johns Hopkins University Press.

Levine, N. 2004. *CrimeStat: Manual* A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 3.0). Houston, Texas: Ned Levine & Associates and the National Institute of Justice, Washington, D.C.

Liebenau, J. M. 1987. Public health and the production and use of diphtheria antitoxin in Philadelphia. *Bulletin of the History of Medicine* 61:216-236.

Lovell, W. G. 1992. "Heavy shadows and black night": Disease and depopulation in colonial Spanish America. *Annals of the Association of American Geographers* 82 (3):426-443.

Lumsden, L. L., and J. F. Anderson. 1911. *Report No. 4 on the Origin and Prevalence of Typhoid Fever in the District of Columbia (1909-1910)*. Washington, D.C.: Government Printing Office.

McKeown, T. 1976. *The Rise of Modern Population*. New York: Academic Press.

McKeown, T., and R. G. Record. 1962. Reasons for the decline of mortality in England and Wales during the nineteenth century. *Population Studies* 16 (2):94-122.

Meade, M. S., and R. J. Earickson. 2000. *Medical Geography*. Second edition. New York: The Guilford Press.

Medical Society of the District of Columbia. 1894. *Report on Typhoid Fever in the District of Columbia*. Washington, D.C.: Government Printing Office.

Meeker, E. 1972. The improving health of the United States, 1850-1915. *Explorations in Economic History* 9 (4):353-373.

Melosi, M. V. 1994. Sanitary services and decision making in Houston, 1876-1945. *Journal of Urban History* 20 (3):365-406.

———. 2000. *The Sanitary City*. Baltimore: The Johns Hopkins University Press.

Mercier, M. E., and C. G. Boone. 2002. Infant mortality in Ottawa, Canada, 1901: assessing cultural, economic and environmental factors. *Journal of Historical Geography* 28 (4):486-507.

Moldow, G. 1980. "For Women, By Women": Women's dispensaries and clinics in Washington, 1882-1900. *Records of the Columbia Historical Society*:290-308.

Mooney, G. 1994. Did London pass the "sanitary test"? Seasonal infant mortality in London, 1870-1914. *Journal of Historical Geography* 20 (2):158-174.

Mortimer, P. P. 1999. Mr. N the milker and Dr. Koch's concept of the health carrier. *The Lancet* 353 (9161):1354-1356.

Myers, S. H. 1973-74. Capitol Hill, 1870-1900: The People and Their Homes. *Records of the Columbia Historical Society* 49:276-299.

National Institutes of Health. 2004. Medical Encyclopedia: Diphtheria. 16 July. http://www.nlm.nih.gov/medlineplus/ency/article/001608.htm. Last accessed 28 February 2007.

_____. 2005. Medical Encyclopedia: Scarlet Fever. 27 October. http://www.nlm.nih.gov/medlineplus/ency/article/000974.htm. Last accessed 28 February 2007.

National Public Radio. 2007. Flu Studies Point to Isolation for Prevention. Reported by Richard Knox. *Morning Edition.* April 3, 2007.

Olson, S. 1997. *Baltimore: The Building of an American City*. Baltimore: The Johns Hopkins University Press.

Ord, J. K., and A. Getis. 1995. Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis* 27 (4):286-306.

Parascandola, J. 1997. Privies and the Public Health Service. *Public Health Reports* 112:168-169.

Parker, H. N., B. Willis, R. H. Bolster, W. W. Ashe, and M. C. Marsh. 1907. *The Potomac River Basin*. Washington, D.C.: Government Printing Office.

Parker, J. N., and P. M. Parker, eds. 2002. *The Official Patient's Sourcebook on Typhoid Fever*. San Diego, California: Icon Health Publications.

Patterson, K. D. 1992. Yellow fever epidemics and mortality in the United States, 1693-1905. *Social Science and Medicine* 34 (8):855-865.

———. 1994. Cholera diffusion in Russia, 1823-1923. *Social Science and Medicine* 39 (9):1171-1191.

Patz, J. A., P. R. Epstein, T. A. Burke, and J. M. Balbus. 1996. Global climate change and emerging infectious diseases. *JAMA* 275 (3):217-223.

Paul, B. K. 1985. Approaches to medical geography: An historical perspective. *Social Science and Medicine* 20 (4):399-409.

Preston, S. H., and M. R. Haines. 1991. *Fatal Years: Child Mortality in Late Nineteenth-Century America*. Princeton, New Jersey: Princeton University Press.

Preston, S. H., and E. van de Walle. 1978. Urban French mortality in the nineteenth century. *Population Studies* 32 (2):275-297.

Provincial Board of Health Ontario. 1912. Report of the Special Investigation into the Cause of Typhoid Fever in the Town of Sarnia Jan. 29th and 30th. Printed by order of The Legislative Assembly of Ontario. Toronto: L. K. Cameron.

Pyle, G. F. 1969. The diffusion of cholera in the United States in the nineteenth century. *Geographical Analysis* 1:59-75.

———. 1973. Measles as an urban health problem: The Akron example. *Economic Geography* 49 (4):344-356.

Pyle, G. F., and P. H. Rees. 1971. Modeling patterns of death and disease in Chicago. *Economic Geography* 47 (4):475-488.

Report of the Commissioners of the District of Columbia for the Year Ended June 30, 1898. Volume III: Health Department. 1898. Washington, D.C.: Government Printing Office.

Report of the Commissioners of the District of Columbia for the Year Ended June 30, 1908. Volume III: Health Department. 1910. Washington, D.C.: Government Printing Office.

Report of the Commissioners of the District of Columbia for the Year Ended June 30, 1909. Volume III: Health Department. 1911. Washington, D.C.: Government Printing Office.

Report of the Health Officer of the District of Columbia for the Year Ending June 30, 1880. 1881. Washington, D.C.: Government Printing Office.

Report of the Health Officer to the Commissioners of the District of Columbia for the Year Ended June 30, 1895. 1895. Washington, D.C.: Government Printing Office.

Rosenau, M. J., L. L. Lumsden, and J. H. Kastle. 1907. *Report on the Origin and Prevalence of Typhoid Fever in the District of Columbia*. Washington, D.C.: Government Printing Office.

———. 1908. *Report No.2 on the Origin and Prevalence of Typhoid Fever in the District of Columbia (1907)*. Washington, D.C.: Government Printing Office.

———. 1909. *Report No.3 on the Origin and Prevalence of Typhoid Fever in the District of Columbia (1908)*. Washington, D.C.: Government Printing Office.

Rosenkrantz, B. G. 1972. *Public Health and the State: Changing Views in Massachusetts, 1842-1936*. Cambridge, Massachusetts: Harvard University Press.

Schultz, S. K., and C. McShane. 1978. To engineer the metropolis: Sewers, sanitation, and city planning in late-nineteenth century America. *The Journal of American History* 65 (2):389-411.

Sedgwick, W. T. 1893. Investigation of recent epidemics of typhoid fever in Massachusetts. In *Twenty-Fourth Annual Report of the State Board of Health of Massachusetts*. Boston: Wright and Potter Printing Company, State Printers.

Smallman-Raynor, M., and A. D. Cliff. 1998a. The Philippines insurrection and the 1902-4 cholera epidemic: Part I – Epidemiological diffusion processes in war. *Journal of Historical Geography* 24 (1):69-89.

———. 1998b. The Philippines insurrection and the 1902-4 cholera epidemic: Part II – Diffusion patterns in war and peace. *Journal of Historical Geography* 24 (2):188-210.

———. 1999. The spatial dynamics of epidemic diseases in war and peace: Cuba and the insurrection against Spain, 1895-98. *Transactions of the Institute of British Geographers* 24 (3):331-352.

———. 2001. Epidemic diffusion processes in a system of U.S. military camps: Transfer diffusion and the spread of typhoid fever in the Spanish-American War 1898. *Annals of the Association of American Geographers* 91 (1):71-91.

Smith, S. J. and D. Easterlow. 2005. The strange geography of health inequalities. *Transactions of the Institute of British Geographers* 30 (2): 173-190.

Snow, S. 2002. Commentary: Sutherland, Snow, and water: The transmission of cholera in the nineteenth century. *International Journal of Epidemiology* 31:908-911.

Statistical Maps: Map No. 2 Street Grades, City of Washington. 1880. In *Annual Report to the Commissioners of the District of Columbia for the Year Ending June 30th, 1880*. Washington, D.C.: Government Printing Office.

Tarr, J. A. 1979. The separate vs. combined sewer problem: A case study in urban technology design choice. *Journal of Urban History* 5 (3):308-339.

Thornton, P. A., and S. Olson. 1991. Family contexts of fertility and infant survival in nineteenth-century Montreal. *Journal of Family History* 16 (4):401-417.

———. 1997. Infant vulnerability in three cultural settings in Montreal, 1880. In *Infant and Child Mortality in the Past*, eds. A. Bideau, B. Desjadins and H. P. Brignoli. Oxford, England: Clarendon Press.

———. 2001. A deadly discrimination among Montreal infants, 1860-1900. *Continuity and Change* 16 (1):95-135.

Trevelyan, B., M. Smallman-Raynor, and A. D. Cliff. 2005a. The spatial dynamics of poliomyelitis in the United States: From epidemic emergence to vaccine-induced retreat, 1910-1971. *Annals of the Association of American Geographers* 95 (2):269-293.

———. 2005b. The spatial structure of epidemic emergence: Geographical aspects of poliomyelitis in north-eastern USA, July-October 1916. *Journal of the Royal Statistical Society* 168 (4):701-722.

Watterson, P. A. 1986. Role of the environment in the decline of infant mortality: An analysis of the 1911 census of England and Wales. *Journal of Biosocial Science* 18:457-470.

Watts, D. J., R. Muhamad, D. C. Medina, and P. S. Dodds. 2005. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proceedings of the National Academy of Sciences* 102 (32):11158-11162.

Whipple, G. C. 1908. *Typhoid Fever: Its Causation, Transmission and Prevention*. New York: John Wiley & Sons.

Whitmore, T. M. 1991. A simulation of the sixteenth-century population collapse in the basin of Mexico. *Annals of the Association of American Geographers* 81 (3):464-487.

Williams, N. 1992. Death in its season: Class environment and the mortality of infants in Nineteenth-century Sheffield. *Social History of Medicine* 5:71-94.

Williams, N., and C. Galley. 1995. Urban-rural differentials in infant mortality in Victorian England. *Population Studies* 49:401-420.

Wills, C. 1996. *Yellow Fever Black Goddess: The Coevolution of People and Plagues*. Cambridge, MA: Perseus Publishing.

Wilson, J. L. 1993. Mapping the geographical diffusion of a Finnish smallpox epidemic from historical population records. *The Professional Geographer* 45 (3):276-286.

Woods, R. I., P. A. Watterson, and J. H. Woodward. 1988. The causes of rapid infant mortality decline in England and Wales, 1861-1921, Part I. *Population Studies* 42 (3):343-366.

———. 1989. The causes of rapid infant mortality decline in England and Wales, 1861-1921, Part I. *Population Studies* 43 (1):113-132.

Woolhouse, M. E. J., D. T. Haydon, and R. Anita. 2005. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends in Ecology and Evolution* 20 (5):238-244.

World Health Organization. 2001. Cholera in India: Disease Outbreak Reported. 14 August. http://www.who.int/csr/don/2001_08_14/en/. Last accessed 30 September 2005.

_____. 2003. Typhoid Fever in Haiti: Disease Outbreak Reported. 17 June. http://www.who.int/csr/don/2003_06_17a/en/index.html. Last accessed 30 September 2005.

_____. 2004. Typhoid Fever in Democratic Republic of the Congo. 15 December. http://www.who.int/csr/don/2004_12_15/en/index.html. Last accessed 30 September 2005.

_____. 2005a. Typhoid Fever in Democratic Republic of the Congo – Update. 19 January. http://www.who.int/csr/don/2005_01_19/en/index. html. Last accessed 30 September 2005.

_____. 2005b. Cholera in West Africa – Update. 23 September. http://www.who.int/csr/don/2005_09_23/en/. Last accessed 30 September 2005.

_____. 2006. Cholera in Sudan – Update 4. 21 June. http://www.who.int/csr/don/2006_06_21a/en/index.html. Last accessed 28 February 2007.

Wu, J., J. Wang, B. Meng, G. Chen, L. Pang, X. Song, K. Zhang, T. Zhang, and X. Zheng. 2004. Exploratory spatial data analysis for identification of risk factors to birth defects. *BMC Public Health* 4:Article No. 23.

Yu, H.-L., and G. Christakos. 2006. Spatiotemporal modelling and mapping of the bubonic plagues epidemic in India. *International Journal of Health Geographics* 5 (12).

APPENDIX A
GLOSSARY OF TERMS

Attribute Table – The data table linked to the map; this contains all of the map's information. For example dates of cases, nativity of cases, for each point, line or polygon in the GIS layer.

Buffer – A distance around a mapped feature(s), usually continuous, used to identify an area of proximity. For example, a buffer around a well might contain all those residences that used it as a water source.

Centroid – The center point of a mapped polygon feature.

Cluster or Clustering (hotspot) – The presence of positive spatial auto-correlation for the events being tested, clustering can occur at a global (the entire study area) or local (small portions of the study area) level.

Disease surface – This term refers to the mapped distribution of a disease across a defined area. For example, the map of all typhoid cases in 1895 for Washington, D.C. (chapter 3) is considered a disease surface.

Field Calculator – A function within the attribute table that allows mathematical functions to be performed upon the attributes (columns of data) contained in the table.

Georeferencing – The process of attaching latitude and longitude coordinates to a digital image in a GIS so that the image will have real world geographic coordinates for digitizing and analysis purposes.

Global spatial auto-correlation – When an entire study area is tested for spatial auto-correlation, using all of the data points, to identify possible clustering across a large area.

Heads-up digitize – The process of creating digital data layers from a scanned map image with geographic coordinates attached in a GIS. The name is coined from the act of physically looking "up" at the computer screen rather than having ones head down using a digitizing tablet.

Line – One of the three types of geographic objects used in a GIS. These features represent objects with one dimension, length. For example, a road is usually represented as a line in a GIS.

Local spatial auto-correlation – When small portions of a study area are tested for spatial auto-correlation so that the small areas are compared to one another or a global mean in order to determine if one small area contains significant clusters of an event in comparison to those other areas or the global mean.

Point – One of three types geographic objects utilized in a GIS. These represent locations as a dot on a map with latitude and longitude for that specific place. For example, a point might represent a well location.

Polygon – One of the three types of geographic objects utilized in a GIS. These simple or complex shapes represent two-dimensional areas. An example of such a feature is a city block or census tract.

Positive Spatial auto-correlation – When mapped phenomena in closer proximity are more likely to be similar than phenomena that are farther apart, thus resulting in clustering, or grouping, of like events near one another.

Boundary Area1906 May and June
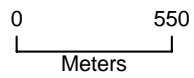
Boundary Area 1906 July and August

1st through the 15th of the month

16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

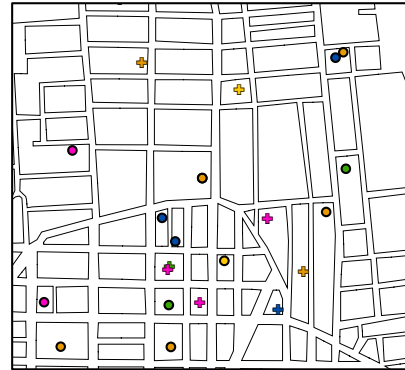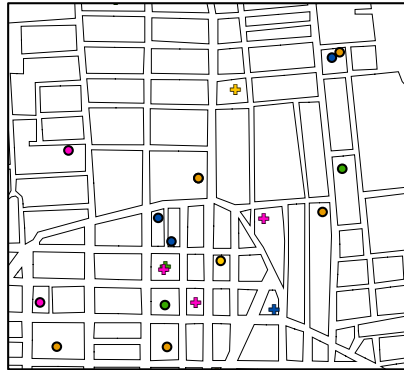| May	June	July	August	
●	●	●	●	1 - 15
✚	✚	✚	✚	16 - 31

0   500

Meters

202

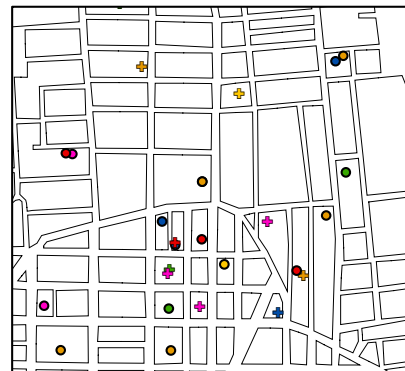Boundary Area 1906 September and October
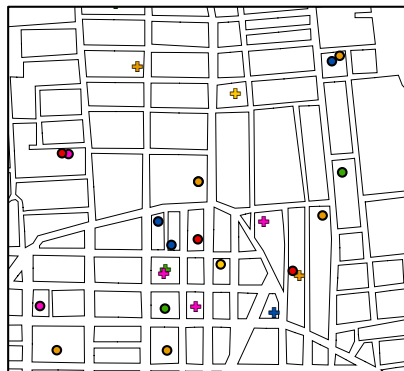
1st through the 15th of the month

16th through the end of the month

September

October

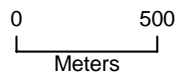Two-Week Period of Case Symptom Onset

May   June   July   August   September   October

1 - 15

16 - 31

0   500

Meters

203
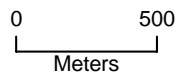
Boundary Area1907 May and June

1st through the 15th of the month

16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   **June**

1 - 15

16 - 31

0   550

Meters

204
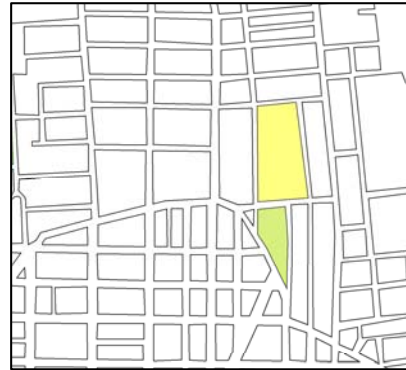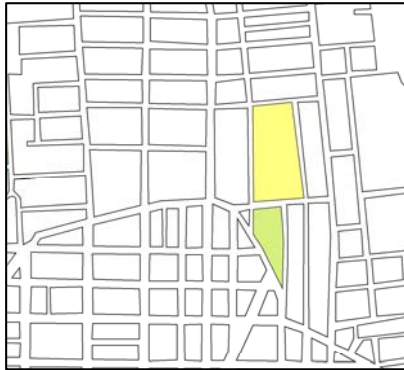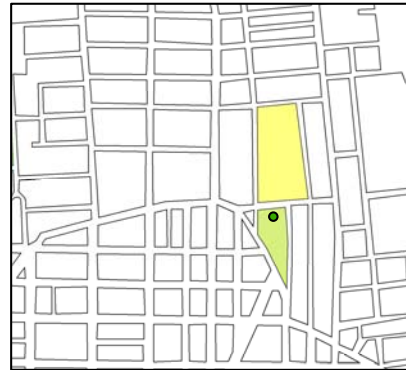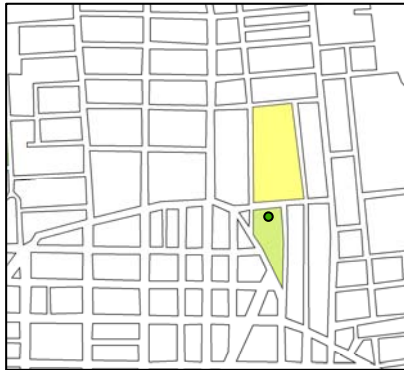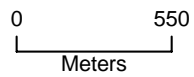
Boundary Area 1907 July and August

1st through the 15th of the month

16th through the end of the month
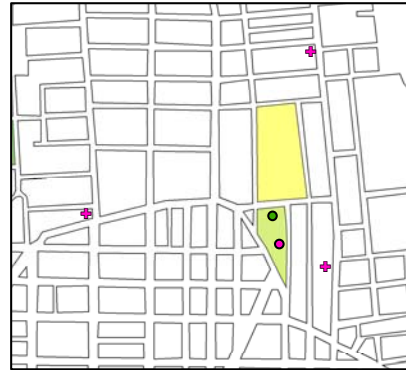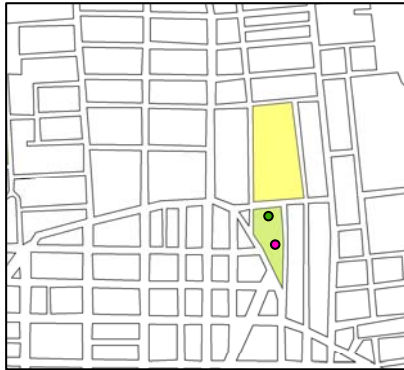
July

August

Two-Week Period of Case Symptom Onset

May	June	July	August	
○	●	●	●	1 - 15
✚	✚	✚	✚	16 - 31

0   500
Meters

Boundary Area 1907 September and October

1st through the 15th of the month

16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
●	●	●	●	●	●	1 - 15
✚	✚	✚	✚	✚	✚	16 - 31

0   500

Meters

206
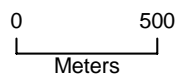
Boundary Area1908 May and June

1st through the 15th of the month   16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

○   ●   1 - 15

✚   ✚   16 - 31

0   500

Meters

207
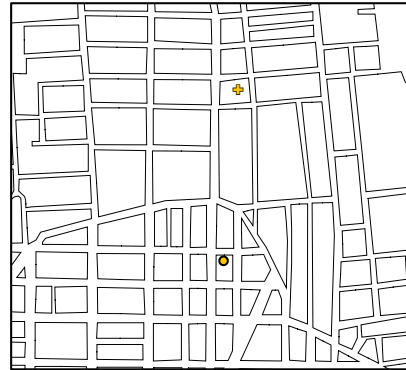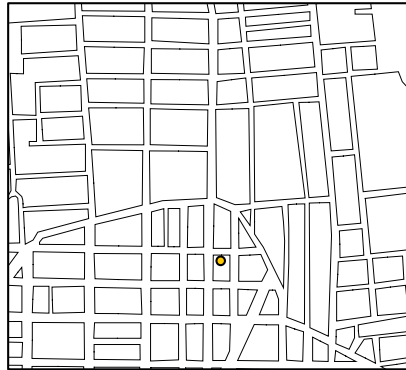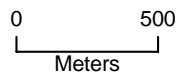
Boundary Area 1908 July and August
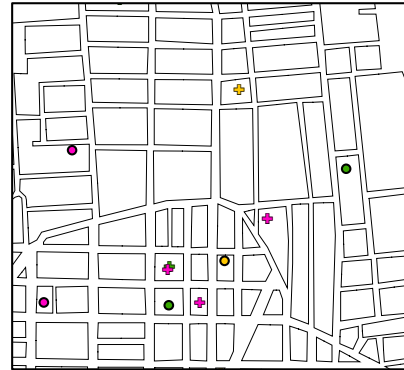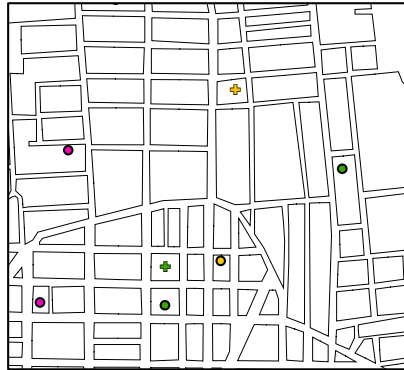
1st through the 15th of the month   16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

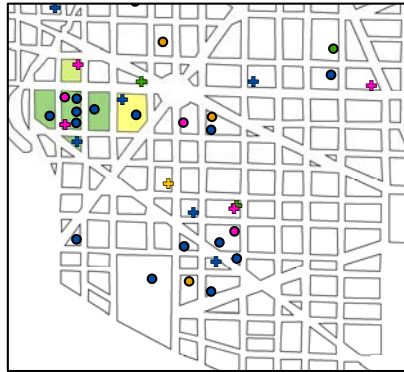| May	June	July	August	
○	●	●	●	1 - 15
✚	✚	✚	✚	16 - 31

0   500

Meters

208
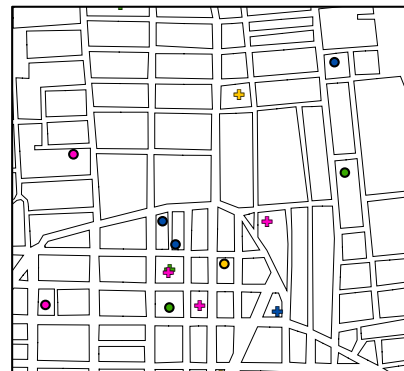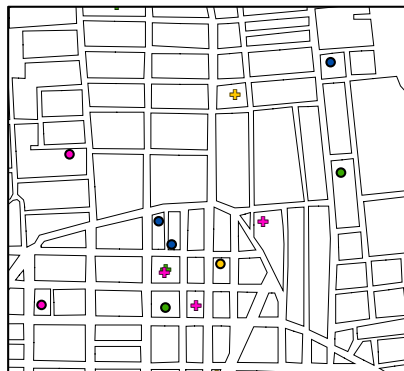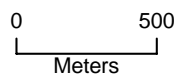
Boundary Area 1908 September and October

1st through the 15th of the month

16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
●	●	●	●	●	●	1 - 15
✚	✚	✚	✚	✚	✚	16 - 31

0   500

Meters

Foggy Bottom 1906 May and June

1st through the 15th of the month

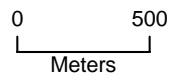16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

1 - 15
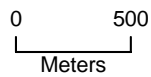
16 - 31

0   500

Meters

Foggy Bottom 1906 July and August

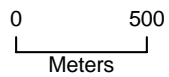1st through the 15th of the month   16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

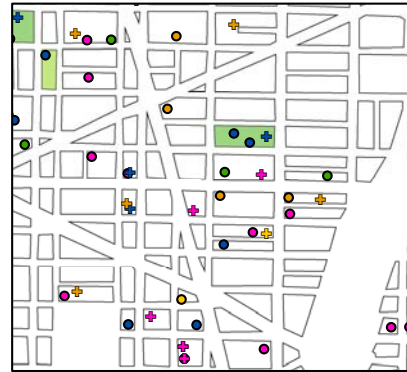|	May	June	July	August	
	●	●	●	●	1 - 15
	✚	✚	✚	✚	16 - 31

0   500

Meters

211

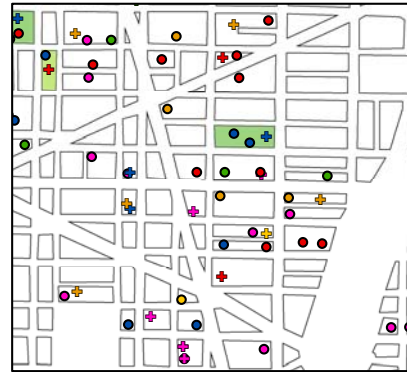Foggy Bottom 1906 September and October

1st through the 15th of the month

16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
○	●	●	●	●	●	1 - 15
✚	✚	✚	✚	✚	✚	16 - 31
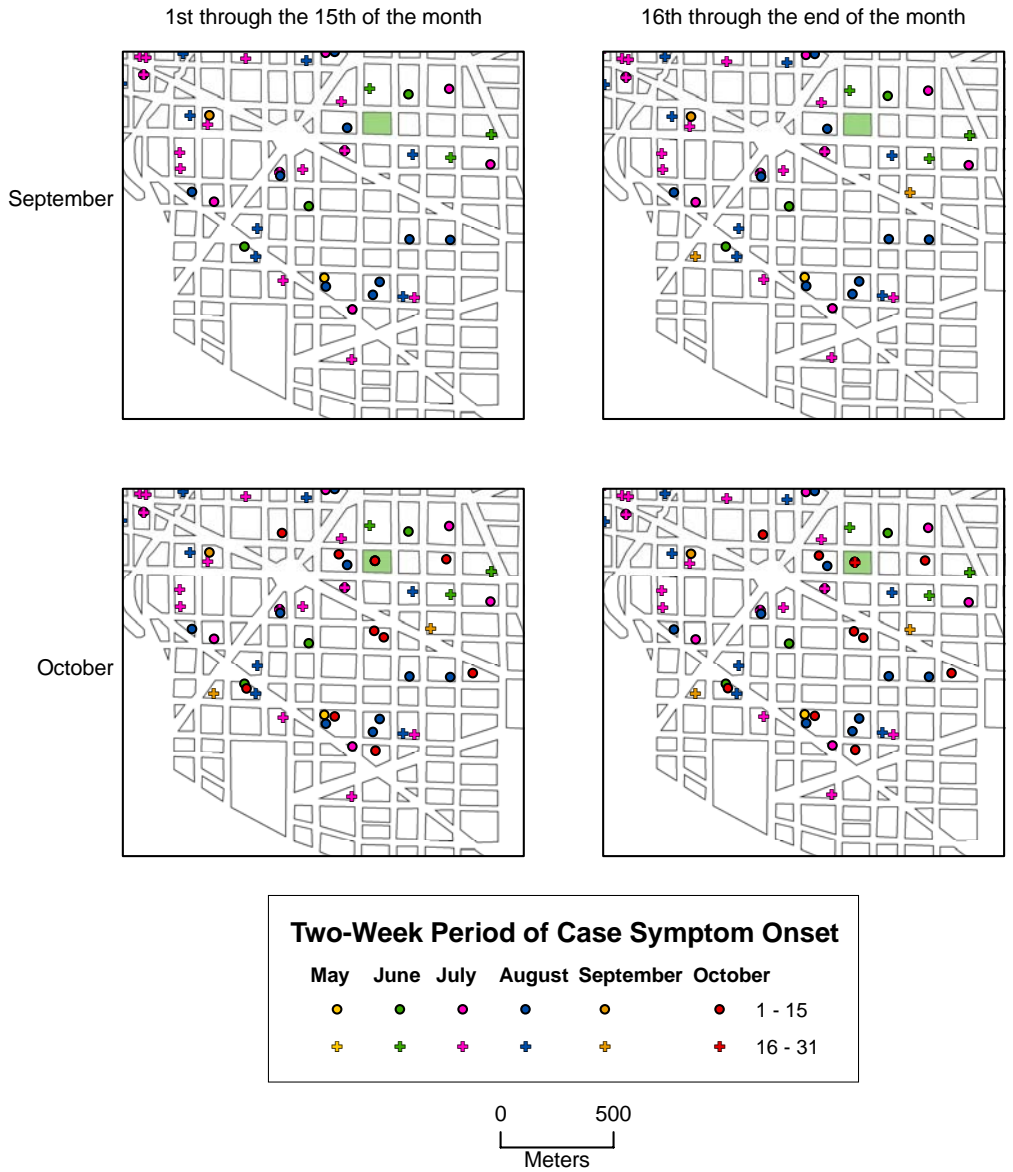
0   500

Meters

Foggy Bottom 1907 May and June

1st through the 15th of the month

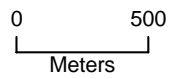16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

1 - 15

16 - 31

0   500

Meters

Foggy Bottom 1907 July and August

1st through the 15th of the month

16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

May	June	July	August	
○	●	●	●	1 - 15
✚	✚	✚	✚	16 - 31

0   500

Meters

214
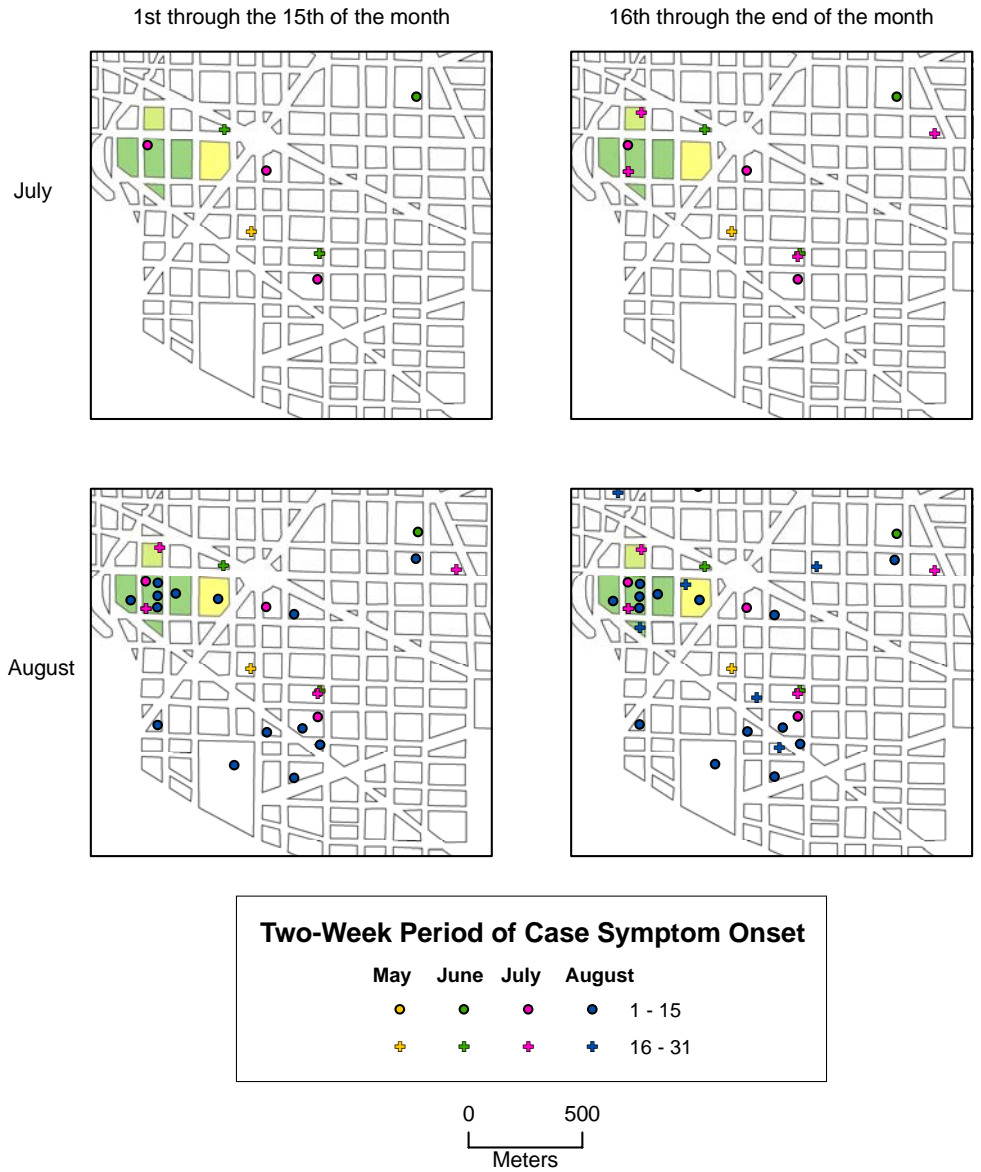
Foggy Bottom 1907 September and October

| 1st through the 15th of the month | 16th through the end of the month |

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
●	●	●	●	●	●	1 - 15
✚	✚	✚	✚	✚	✚	16 - 31

0   500

Meters

215

Northwest 1906 May and June

1st through the 15th of the month

16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

1 - 15

16 - 31

0   500

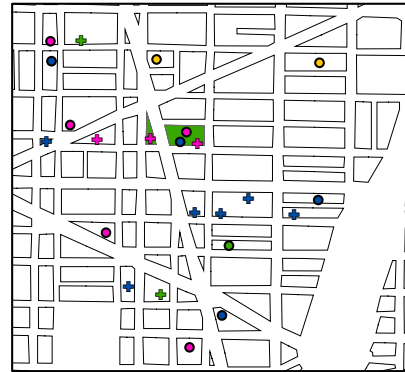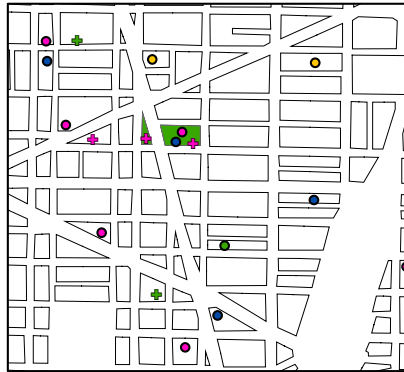Meters

216

Northwest 1906 July and August

1st through the 15th of the month

16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

	May	June	July	August	
1 - 15	●	●	●	●	
16 - 31	✚	✚	✚	✚	

0   500

Meters

217

Northwest 1906 September and October

1st through the 15th of the month — 16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

	May	June	July	August	September	October	
●							1 - 15
✚							16 - 31

0   500

Meters

Northwest 1908 May and June

| 1st through the 15th of the month | 16th through the end of the month |

May

June

Two-Week Period of Case Symptom Onset

May	June	
○	●	1 - 15
✚	✚	16 - 31

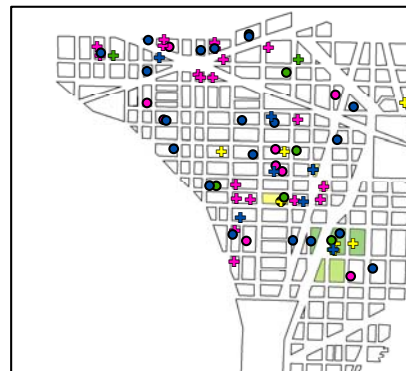0   500

Meters

Northwest 1908 July and August

1st through the 15th of the month

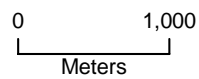16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

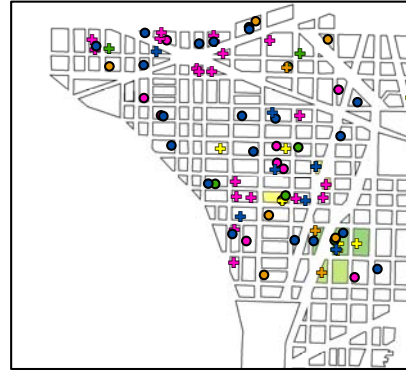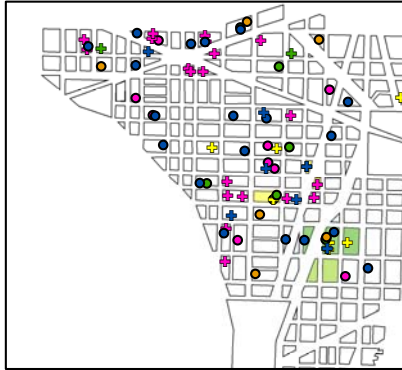|	May	June	July	August	
	●	●	●	●	1 - 15
	✚	✚	✚	✚	16 - 31

0   500

Meters
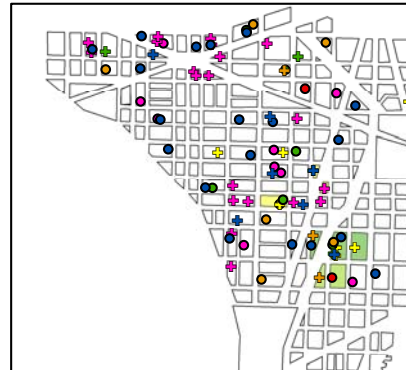
220

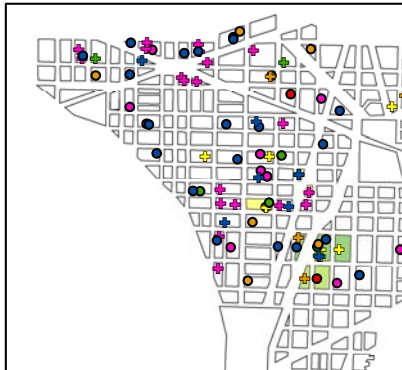Northwest 1908 September and October

1st through the 15th of the month   16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
						1 - 15
						16 - 31

500

Meters

Southwest 1906 May and June

1st through the 15th of the month

16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

1 - 15

16 - 31

0   1,000

Meters

Southwest 1906 July and August

1st through the 15th of the month

16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

	May	June	July	August	
	○	●	●	●	1 - 15
	✚	✚	✚	✚	16 - 31

0   1,000

Meters

223

Southwest 1906 September and October

1st through the 15th of the month

16th through the end of the month

September

October

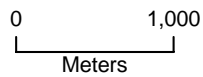## Two-Week Period of Case Symptom Onset

	May	June	July	August	September	October	
●							1 - 15
✚							16 - 31

0   1,000

Meters

224

Southwest 1907 May and June

1st through the 15th of the month   16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

○ ●   1 - 15

✚ ✚   16 - 31

0   1,000
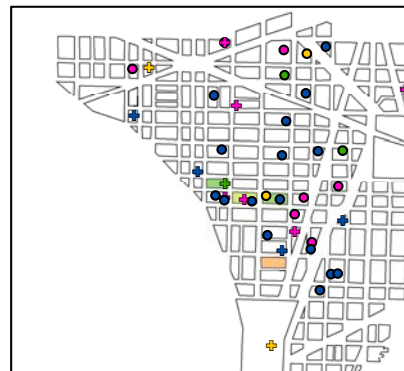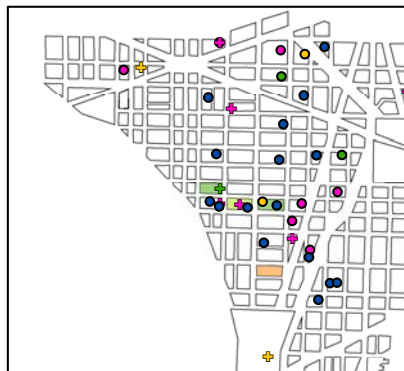
Meters

225

Southwest 1907 July and August

1st through the 15th of the month
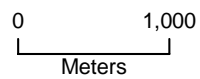
16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

	May	June	July	August	
	⬤	⬤	⬤	⬤	1 - 15
	✛	✛	✛	✛	16 - 31

0   1,000
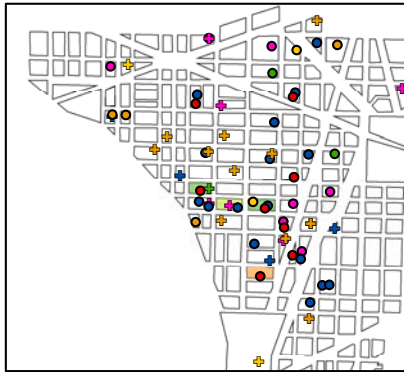
Meters

Southwest 1907 September and October

1st through the 15th of the month   16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
●	●	●	●	●	●	1 - 15
✚	✚	✚	✚	✚	✚	16 - 31

0   1,000

Meters

227

Southwest 1908 May and June

1st through the 15th of the month   16th through the end of the month

May

June

Two-Week Period of Case Symptom Onset

May   June

○   ●   1 - 15

✚   ✚   16 - 31

0   1,000

Meters

Southwest 1908 July and August

1st through the 15th of the month
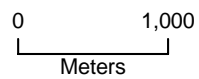
16th through the end of the month

July

August

Two-Week Period of Case Symptom Onset

	May	June	July	August	
	●	●	●	●	1 - 15
	✚	✚	✚	✚	16 - 31

0   1,000

Meters

229

Southwest 1908 September and October

1st through the 15th of the month

16th through the end of the month

September

October

Two-Week Period of Case Symptom Onset

May	June	July	August	September	October	
-----	------	------	--------	-----------	---------	
●	●	●	●	●	●	1 - 15
✚	✚	✚	✚	✚	✚	16 - 31

0   1,000

Meters

VITA

Sarah Elizabeth Hinman was born on October 6, 1977, in New Haven, Connecticut. She earned her Bachelor of Arts degree, which included a junior year abroad at St. Andrews University in St. Andrews, Scotland, from Mary Washington College in May 2000. In June 2002, she completed her Master of Arts degree in geography from Ohio University and a few months later moved to Baton Rouge to begin her doctoral studies. Sarah will complete her Doctor of Philosophy in geography from Louisiana State University in May 2007.