

# Medical Mapping: The Revolution in Teaching—and Using—Maps for the Analysis of Medical Issues

Tom Koch and Ken Denike

## ABSTRACT

As a result of advances in GIS, students from a range of disciplines (epidemiology, public health, medical sociology, and anthropology) are seeking to learn how to apply a mapped, spatial perspective to issues of disease and health. This paper describes a pilot program whose intent has been to develop an interdisciplinary, skills-based course in medical mapping using a case-based approach. The focus of its first module is the famous nineteenth-century case of a cholera epidemic in Soho, London. Here, however, that case—critical to a range of disciplines—is an opportunity to introduce modern GIS tools and their applicability to the issues surrounding the mapping of diseases in general.

**Key words:** Cholera; Epidemics; GIS; Interdisciplinary studies; Snow, John

Tom Koch, Ph.D., is adjunct professor of geography at the University of British Columbia, Vancouver, BC, Canada, and adjunct professor of gerontology at Simon Fraser University, Vancouver, BC, Canada.

Ken Denike, Ph.D., is assistant professor of geography at the University of British Columbia and chair of the British Columbia Public School Employer's Association.

Medical mapping is undergoing a renaissance. The relative ease of use of contemporary Geographic Information Systems (GIS) and the availability of national and international, health-related data through the Internet have combined to create a new and unusually ecumenical generation of persons using maps to understand medical issues. Into the early 1990s, maps of disease and health remained laborious instruments drawn by hand. At that time Cliff and Haggett's (1988) *Atlas of Disease Distributions* represented the state of an arcane and difficult art. Maps of disease and health were laborious to draw and time-consuming to modify. Their data were hard to calculate, and by current standards the results were often visually primitive.

With the introduction of popular GIS programs, however, mapping large sets of data has been facilitated. The result has been a host of papers on medical mapping and a plethora of medical maps. Publications of medical atlases, the *Atlas of US Mortality* (USNTI, 2000) for example, have proliferated. The publication of medical maps individually and in the aggregate is now widely reported (and interpreted) both popularly (for example, Vergano 2000; Snapshot 2001) and in the medical literature (Gundersen 2000). This in turn has sparked a range of publications by health analysts (Ricketts et al. 1997), companies involved in GIS (Lang 2000), and by researchers interested in health and disease-related issues (Gattrell and Löytönen 1998).

The relative ease with which maps can be generated using GIS masks the real difficulties of the graphic exposition of health-related phenomena. In developing an interdisciplinary course on medical mapping our focus therefore has been on GIS mapping as an analytic exercise rather than as a relatively simple technical problem in mapped presentation. The goal has been to create a series of modules that can be used by GIS students in general courses, in classes specific to medical geography, and by all those (medical anthropologists, epidemiologists, public health specialists, etc.) for whom mapping is a potentially potent tool.

## BACKGROUND

While there is a growing literature on the potential of GIS as a general tool for geographical instruction (Moseley 2001), comparatively little has been published on the discursive or analytic advantages of mapping over the purely statistical approach that has long dominated not only medical geography but also public health and epidemiology. Our intention was (and remains) to present a teaching posture that argues the visual advantages of mapping as a unique form of analysis, albeit one often presenting the outcome of statistical analyses.

To this end a series of modules have been developed linking lecture materials with laboratory mapping exercises to introduce simultaneously both GIS mapping techniques and a critical perspective on the health-related data to be considered, analyzed, and mapped. The focus of the first module is Dr. John Snow's paradigmatic analysis of the 1854 cholera outbreak in Soho, London. Its

advantages as a structured teaching case are numerous. First, detailed mapped data on that epidemic are available from a range of sources (see for example CDC 2001; Tobler 1992). Second, a wealth of critical commentary exists both in the print literature (for example, Brody et al. 2000; McLeod 2000) and on the World Wide Web (Frerichs 2001) online (for example, <http://www.ph.ucla.edu/epi/snow/>). The Snow case is also used by the U.S. Centers for Disease Control as a teaching example for its software, EpiInfo, designed to facilitate the collection and analysis of data (CDC, EpiInfo 2000).

As importantly, for our purposes, epidemiologists, medical geographers, and public health experts all claim Snow's argument as a seminal event in the development of their respective disciplines (McLeod 2000). The presentation of this case is therefore a potential first step not only for medical cartographers but also for virtually all disciplines interested in the origin and diffusion of diseases.

### THE PROGRAM

Building an interdisciplinary, health-based class around GIS is difficult. The temptation is to focus on the tool rather than the ideas that propel the analysis of the data. To assure an understanding of both the issues and the problems of their analysis, therefore, the course is divided into a series of modules—lectures and labs—that can be individually assigned and used by varying instructors. Different modules present problems of health and illness at a range of scales and with attention to both social and environmental contributing factors. In each module the focus is on the difficulties of identifying causative elements at a specific scale of analysis and presenting these in a manner that is robust, inclusive, and at the same time legible.

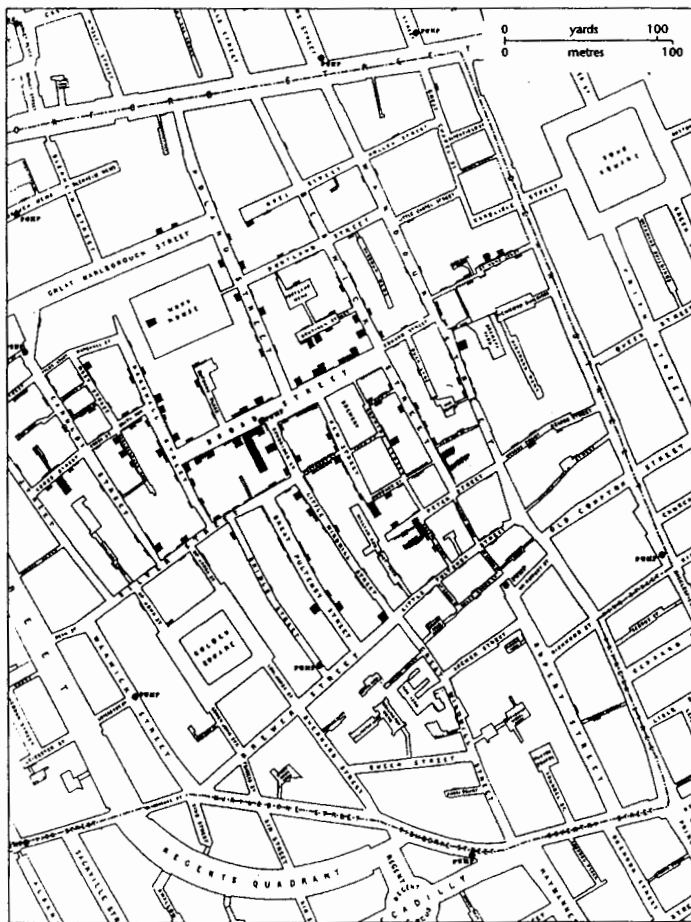
The cholera story is both the first module and the model of the work we attempt in later exercises. Here we describe in some detail our approach to the case as an introduction to the instruction of GIS mapping as a tool in the analysis of disease incidence, origin, and occurrence. Pedagogically, it permits a discussion of the critical and often complex relationship between disease, class, and public infrastructure that are Snow's singular critical intellectual legacy (Brody 2000). More narrowly it facilitates a consideration of Snow's mapped approach while permitting the introduction of a range of GIS tools. Other modules in the evolving program focus on a range of health and health policy issues. Space limitations restrict this paper's focus to our approach to teaching the Snow case. In the future we hope to present a broader assessment of the greater program and the modules and labs that follow this module in our program.

### CHOLERA: LECTURES

Cholera is introduced as a terrible and recurrent nineteenth-century pandemic that remains in modernity a recurring public health concern. The more recent pan-

demic of the 1990s is briefly discussed as the latest visitation of this hoary, bacterial disease. Because much of what we know about cholera as a water-borne disease is based on Snow's seminal work, it becomes a logical place to begin the study of this disease, and more generally, broader issues attending upon the spatial incidence of disease diffusion.

In lecture, a map of London and its urban districts is presented for orientation. The area of the 1854 cholera outbreak in Soho is identified within the greater context of the extended city and its principle water-course, the Thames River. Issues of mid-nineteenth-century industrializing society are introduced in the context of both social structure and the problems of water supply, sewage disposal, and urban infrastructure during that period (Porter 1998; Snow 1855a). Social issues surrounding urban life and issues of class, are described as they related to this disease specifically, and urban life in nineteenth-century London generally (O'Connor 2000; Porter 1998; Snow 1855a).



*Figure 1. Detail of Snow's 1854 map of Soho Cholera outbreak. A portion of Snow's 1854 map of the Soho area outbreak from his *On the Mode of Transmission of Cholera, Second Edition*. Each death is marked as a single bar on city base map. The scale of the map is 30 inches to the mile.*

Snow's map of the location of deaths from cholera in the 1854 Soho epidemic and the water pumps that were its suspected sources are then introduced (Figure 1). While the use of dot maps for the identification of the location of diseases was then still relatively new, Snow's mapping of the outbreak was not wholly original. Others had made hand maps of disease outbreaks in both Europe and North America. (Jarco 1970; Stevenson 1965). What distinguished Snow's work was both the nature and quality of the data he collected and the manner in which it was analyzed. Snow's map thus represents not the first dot map of a disease but perhaps more importantly the first analytic use of mapped data in epidemiology, one that attempts to understand a disease's origin by considering the relation between two distinct data sets (suspected sources, observed cases) in a single plane. As others have noted, his study also was among the first instances of a modern, statistical approach to epidemiology generally (Huber 1999).

During the cholera epidemic of 1854 John Snow carried out two separate studies, one of the greater epidemic in South London and another of the outbreak in the Soho area where he lived. For the latter project he gathered data on more than 518 deaths from the General Register Office—the office charged with collecting such data—from other physicians, and in a survey of affected houses that he personally conducted (Snow 1855a, 39). In his own investigations he attempted to ascertain from relatives of the deceased where they had gotten their drinking water.

Snow marked each death on an existing map of the area affected by the outbreak. Where multiple deaths occurred he stacked the symbols for those deaths—a single bar drawn with a pen—so that the density of deaths at individual locations would be evident. This stacking of deaths at a street address was an accepted convention of the day and one familiar to Snow's contemporaries (see, for example, Whitehead 1855, 96-97). Modern mapmakers—especially those using GIS—more typically show multiple occurrences at a single location through symbolizing of single points—progressively larger symbols and color ranking—to show multiple occurrences at a single location. Still, the accuracy of Snow's representational system, and his data, has been confirmed by the authors through comparison data collected by other period researchers (for example, Whitehead 1855, 96-97).

However they might be symbolized today, Snow's approach served his purpose. Famously, his map of the apparent concentration of deaths around a single water pump focused Dr. Snow's argument against a "miasmatic" (airborne) origin for the spread of cholera generally. In Soho, his map strongly argued for a single, infected water source—the Broad Street Pump (pump No. 7)—as the epicenter of the outbreak. It is

with this mapping that most cartographic reviews of Snow's analysis begin and end (Lang 2000).

It is important to remember that proving the centrality of the Broad Street Pump to this outbreak did not mean contemporary acceptance of Snow's thesis of cholera as a disease spread by contaminated water (Frost 1936). Into the 1800s, many continued to believe that cholera was somehow airborne. Thus the GIS maps in our program include the location of a nearby cemetery from which some believed foul airs generating cholera rose. The decades of resistance to Snow's theory, and his opponents' adherence to the idea of miasmatic, bad airs, emphasizes a central point of the lectures: the difference between data required to promote a general theory of disease diffusion and that required for the description of a single disease outbreak.



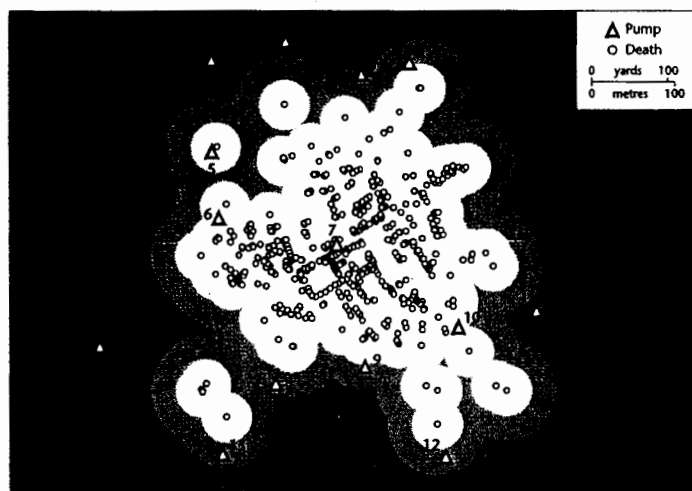
**Figure 2.** GIS version of John Snow map. A digital version of the Snow cholera exercise. Layers include: deaths, pumps, streets, and cemetery. Like the original, this map presents multiple deaths at a single house by stacking the symbol for those deaths out from the street.

**CHOLERA: LAB****Technical Issues**

The teaching program was largely developed at the University of British Columbia (UBC) for a geography course in "spatial data analysis using GIS," and secondarily, as an example of GIS approaches in epidemiology. The lab exercises were originally written for ArcView 3.2 with the spatial analyst extension but have recently been rewritten in ArcGIS 8.1. Choice of software was dictated by UBC's standardized use of first ArcView, and more recently, ArcGIS, in its GIS labs. In addition, elements of the lab have been prepared in Maptitude 4.1, a GIS program also used by one of the authors.

Digitized data on the 1854 outbreak is available from several sources. In our case, data was obtained in 1998 from Environmental Systems Research Institute (ESRI), which had prepared a demonstration model for health experts based on the Snow case. Another, similar version prepared for IDRISI users was also reviewed. The ESRI data set is eccentrically projected with coordinates that while internally consistent lack "real world" coordinates.

An accurately referenced and appropriately projected data set is available from the U.S. Centers for Disease Control as part of its EpiInfo teaching software (CDC 2000). Unfortunately, its set of pump (N = 11), street network, and the data set describing the total of cholera deaths (N = 452) are incomplete and therefore unusable in its present form (Koch, In press). The authors are currently creating a complete and appropriately referenced data set for use in this program. Future versions of this lab will include, in fact, a module using varying versions of the Snow data to teach issues of projection, coordinate system manipulation, and techniques of data checking and cartographic rectification.



**Figure 3.** Buffers around deaths from cholera in the 1854 outbreak.

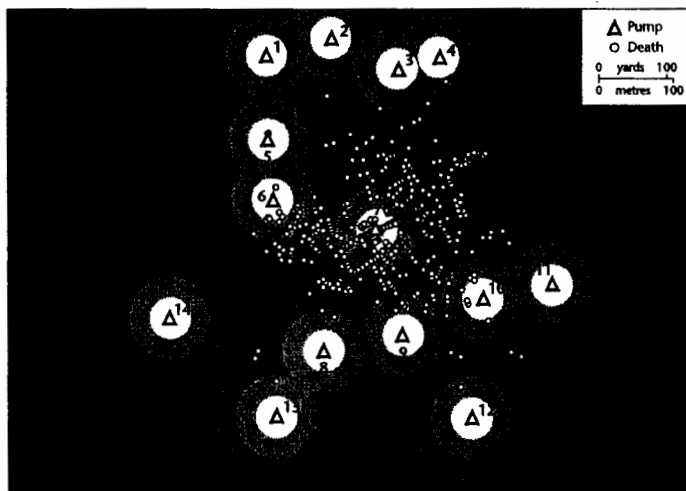
In the current lab exercise, Snow's map is introduced as a simplified schematic including only a set of streets, water pumps (N=14), cholera deaths (N = 518), and the location of the former cemetery that many believed was a potential source of the Soho outbreak. Streets and work places in the neighborhood had been built over the cemetery and an older plague burial pit more than a century earlier.

This was sufficiently proximate to the apparent epicenter of the outbreak to suggest it as a potential airborne disease source (see Resident, Times 1854) of the outbreak. Students are asked to look carefully at the data and see if, on gross inspection, any single source—water pumps or cemetery—appears to be the logical, single source of the outbreak. They are asked if the outbreak appears to be moving across the mapped surface, and if so, does that movement suggest one or another source of contagion? What other data, they are asked, would be needed to consider that hypothesis?

"Forgetting what we know," as one student said in his report, "several pumps seem to be engaged, and so does the cemetery." Most agreed that the outbreak appeared to be moving in a southeast to north, north-westerly direction, and to be spread out along an east-west axis. On the basis of disease incidence alone the centrality of the Broad Street Pump is suggested but certainly not conclusive.

**BUFFERS**

Students are then invited to investigate either the location of the deaths or the apparent centrality of the water pumps. We have found that many GIS students, and some from other disciplines, intuitively assume that any analysis should first focus on the location and incidence of the deaths themselves and only secondarily on suspected points of origin. As one student said, "After all, we're looking at deaths, not locating water sources,



**Figure 4.** Buffers around water pumps in the Soho area in 1854. Water to the Soho area in 1854.

so it's with the deaths we should start looking for clues." Pedagogically, the difference between studying the incidence of disease and the relation between disease and another factor is critical. Snow's essential contribution was to relate two distinct layers of data—pumps and deaths—to test a hypothesis about the relation between disease and a vector of its diffusion. To make this point, students are instructed in the use of contours and buffers (circles of a consistent radius drawn around a single point), perhaps the most basic GIS analytic we know. They are instructed to first create buffers centered around the individual deaths (Figure 3) and then to create a separate layer of buffers centered on each of the pumps (Figure 4). The first set represents the reflexive assumption that a study needs to focus on the incidence of the disease and the second Snow's idea that the relation between water and disease is what is required. Figures 3 and 4 present the results. Because there were so many deaths in this outbreak (518), in Figure 3 overlapping buffers are joined. But because there the set of pumps (N=14) is limited, they are left distinct even where they appear to overlap.

Both approaches suggest the centrality of the Broad Street pump, albeit in different ways. The impression of Figure 3 is of centrality and density of occurrence centered on the pump, exactly what Snow described. In Figure 4, the centrality of the Broad Street is emphasized. Next, buffers are redrawn so that the circles surrounding the central pumps (Broad Street's and those nearest to it) abut. The number of deaths within each are then counted, a simple exercise in most GIS. In this way the idea of analyzing one set of data based on the boundaries of another—an innovation Snow was perhaps the first to suggest at this scale—is introduced.

Students then manually draw a circle of equal size centered on the cemetery. The result is that fewer deaths than are counted in the Broad Street pump buffer but more deaths than occur in buffers surrounding most other pumps. For many, this presents strong evidence that the cemetery itself is not the epicenter of the outbreak and lessens its apparent involvement in the outbreak.

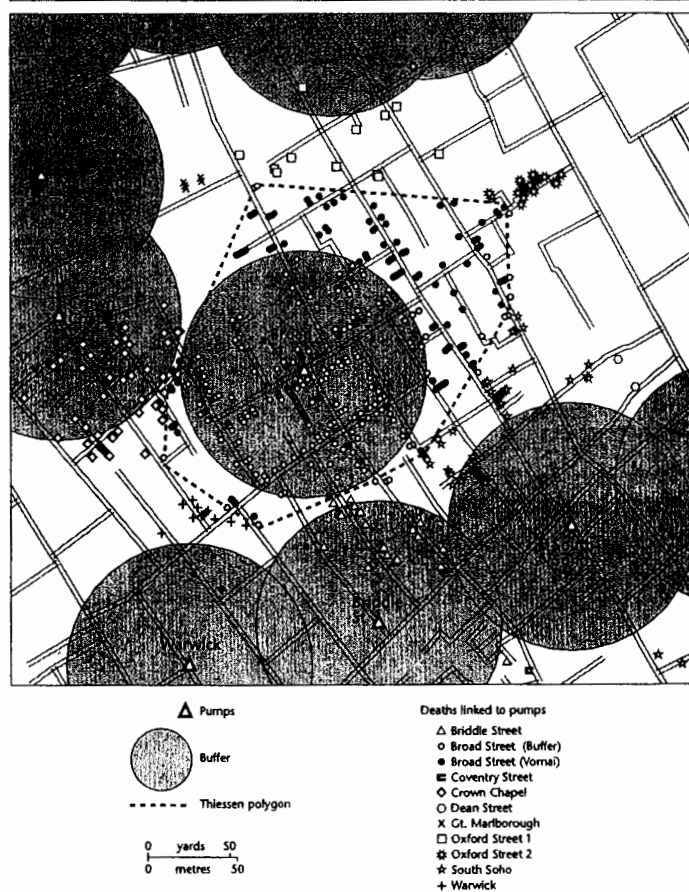
#### NEAREST NEIGHBOR ANALYSIS

In effect, buffering around individual deaths gives us Snow's perspective, one of apparent centrality. Buffering of the pumps permits a first attempt at quantifying that centrality. While using Snow's general insight as a hypothesis, GIS provides critical tools that permit the student to go beyond Snow's general approach. Nearest neighbor analysis in GIS permits the elements of two distinct data sets to be analyzed based, in this case, upon the distance between each member of the first set and the nearest member of the second set.

Students are instructed to "join" the set of pumps (N=14) to that of cholera deaths (N=578) so that each death is assigned to one and only one pump, the

one nearest to it. The GIS automatically calculates the nearest pump to each death, the distance between them, and then includes this data in its database. The results are presented in Figure 5. Different symbols are used to distinguish deaths by proximity to the nearest pump. That 361 (69 percent) of the 518 mapped deaths are nearest the Broad Street pump quantifies Snow's general statement of apparent centrality.

Students then are asked to analyze clusters of deaths that are nearest to other pumps. In this exercise the total set of deaths are symbolized not by proximity to a pump but more generally by distance from their nearest pump. In this way clusters of "near" deaths occurring relatively near other pumps—those nearest the Marlboro St. pump (Pump No. 6) northwest of Broad Street, for example—are identified. While relatively few in number compared to the concentration around the Broad Street site, the number of deaths nearer to other pumps was one reason some of Snow's contemporaries discounted his conclusions. How the apparent relation between these deaths and their pumps might be explained is the subject of classroom lecture and discussion.



*Figure 5. Detail of deaths by distance from pumps. A "nearest neighbor" analysis based on the distance relation between individual cholera deaths and the pumps. A polygon is drawn around them to show the relationship between buffer and polygon analysis.*

In a companion lecture this exercise also provides an opportunity to discuss issues of data categorization for the purposes of mapping. Thus use of “natural breaks,” “standard deviation,” and “equal interval” classification schemes are considered (Zeiler 1999, 36-38), and their results based on the Snow study are compared. Whatever method is used to categorize the data, the results show a singular concentration of deaths near the Broad Street Pump (No. 7). Concentrations nearer other pumps, especially the one at “Crown Chapel,” however, provide more precise data for detailed analysis.

### THIESSEN POLYGONS

Today, John Snow (1855b) often is credited with being the first to use what today are called Thiessen or Voronoi polygons to map one source of locational data upon another, a fundamental contribution to analytic mapping (McLeod 2000; Brody 2000). This basic form of proximity analysis creates catchment areas based on the proximity of point-based data attributes, in this case the location of the water sources (McLeod 2000, 133; Brody 2000, 4). The mid-point between locations (pumps) is marked and these are then joined to create basic catchment areas that can be used to consider the incidence of data in any other set, in this case, deaths from cholera.

In his 1855 map, Snow mapped “by the most careful calculation” a meandering line marking the median distance between the Broad street pump and others nearest to it in the district (Snow 1855b). Snow did not, however, create either a Thiessen polygon or a Voronoi network of Thiessen polygons. He created a single eccentric and highly irregular district centered on the Broad Street quite unlike the polygon returned by formal Thiessen polygon calculations.

The difference between formal Thiessen polygons and the boundary Snow drew around the Broad

Street pump is shown in Figure 6. Figure 6a is a detail of Snow’s map with the dotted line representing the boundary he drew around the Broad Street pump Figure 6b is a Voronoi network of Thiessen polygons, each polygon’s centroid a pump (Cliff and Haggett 1988). Numbers in each polygon refer to the number of deaths occurring within each polygon. The analysis was performed by Cliff and Haggett (1988) using the Snow example to describe the generation of a Voronoi network and may be the source of the belief that Dr. Snow himself created this type of polygon analysis.

While Snow intuited the potential of formal polygon analysis, the rationale of Thiessen polygons themselves is not clear to many students. As important, when a polygon network is composed of many points, its manual execution can be extremely difficult. It was in part for this reason that formal polygon analysis was rarely used until the introduction of computers in the mid-twentieth century. Using GIS, however, the rationale behind the analytic approach can be derived quite easily. One way to teach this is through the work previously described and summarized in Figure 5.

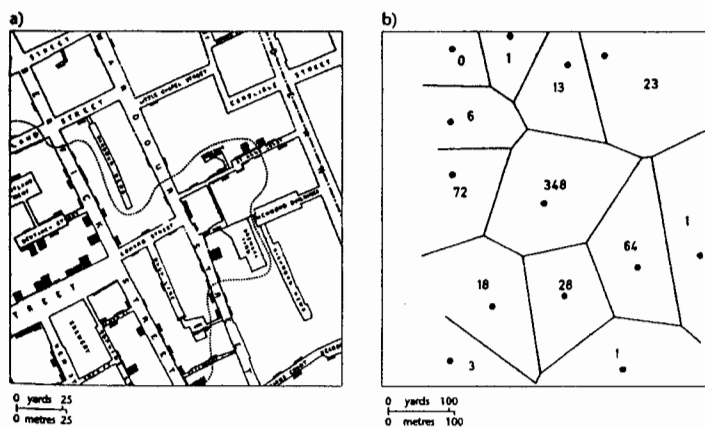
Another way is to invite students to generate a series of multiple buffers, or in some programs, contour lines, with the individual pumps as their center. As the lines converge, the shape of the resulting polygons can be seen at their juncture. Simply, the buffers—or contours—abut midway between any two locations and the result is the boundary of the Thiessen polygon.

Distance between multiple circles or contours will determine how clear the resulting shape will be. While tighter concentric circles or contour lines gives a more precise boundary of the pump polygons, a relatively coarse distancing is used in Figure 7 to facilitate clarity in reproduction. In the lab students use a finer contour or buffer and then manually add lines at the point where the circles surrounding each pump are seen to abut. The result is a Voronoi network of Thiessen polygons.

Pedagogically, this permits not only a real understanding of polygons based on median distance as an analytic tool but also naturally engages discussion of the difference between this approach and the nearest neighbor analysis (Figure 5).

While in this case the results are similar the essential difference between the two modes of analysis—one based on the relationship between two data sets and the second only on elements of a single set—is critical. Nearest neighbor analysis considers the relationship between two distinct sets of data based on proximity and distance. The polygon approach first considers only the distance relations between members of a single data set. The result may then be applied analytically to other data, in this case incidence of deaths.

In this manner more general issues of districting are introduced. Irrespective of cholera, for example, students are asked to consider the pump locations as service



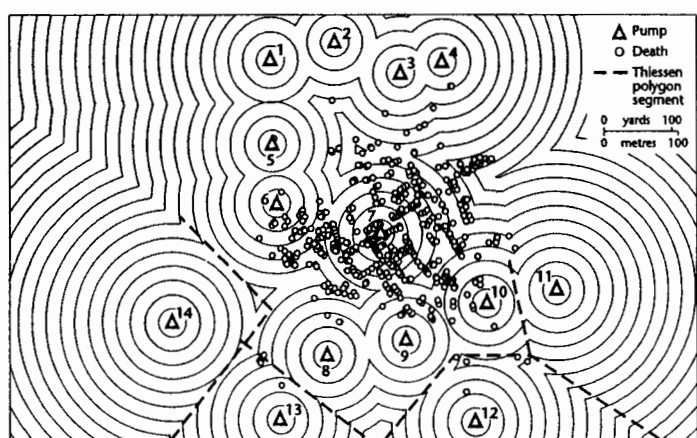
**Figure 6.** Thiessen Polygon analysis: (a) shows the irregular path Snow used to describe the “cholera area” centered on the Broad Street pump; (b) is a Voronoi network of Thiessen polygons drawn by Cliff and Haggett (1988), using the cholera data, to describe the generation of a polygon network.

areas. If they were water managers where would they introduce a second pump to equalize water service in the district? What information would they need, besides relative location, to decide upon pump location planning? In this manner the specific study is generalized to one introducing issues of districting by area, or by other factors.

When time and circumstance permit, other potentially useful analytic tools are introduced, including: spiders (directional lines joining members of one set (deaths) to the closest member of another set (pumps)), time-based histograms, spatial and weighted means, and deviational ellipses. It is useful to note in passing that Dr. Snow was among the first to use a histogram to describe the course of an epidemic outbreak by time (Snow 1855a; 1855b). Spatial and weighted means are tools Snow did not use but which are a constant of modern spatial analysis in a number of areas. Students in this program are shown how to use the GIS to calculate spatial and weighted means of both the deaths and the pumps, and to compare the data those means represent.

#### MODERN USES

Some doubt the utility of teaching this case in a course on modern issues of medical mapping. Monmonier, for example, dismisses the Snow case as irrelevant to modernity. "Real epidemiology isn't like that," he insists, "at least not in late-twentieth-century America. Cholera is rare, if not extinct" (Monmonier 1997, 283). But "real epidemiology" is often "just like that." And cholera, alas, is neither rare nor extinct. A global pandemic that began in the 1960s diffused, in the 1990s, from Africa and Asia to the Americas. The relevance of the techniques taught in the Snow module and the reality of modern epidemic outbreaks and pandemics become, in our program, modules that flow directly from the historical case.



**Figure 7.** Polygon generation. Concentric circles are drawn around individual pumps to demonstrate how Thiessen polygons are generated at the mid-point between proximate points in a data set. A relatively course circle was chosen to maximize legibility.

Following the Snow lab and lectures students are involved in a separate lab requiring the mapped analysis of a recent salmonella outbreak in Greater Vancouver in the summer of 2000. Data on the 67-case outbreak were obtained from the British Columbia Center for Disease Control (BC-CDC). Mapped with these cases are the locations of 21 restaurants and food outlets that local epidemiologists identified as possible sources of the outbreak. In lecture and in the laboratory exercise the salmonella case is used as a way of updating the cholera study and introducing the problem of field data and questionnaires. To this end students are given the exercise of creating a data query sheet that lists all the fields they think important for public health workers exploring a salmonella outbreak. This naturally leads back to a discussion of Snow and the data he collected on each case of the 1854 outbreak in Soho. The point is that mapping is only one step in the process. "What to ask" is as important as "where" those questions are asked.

#### CHOLERA IN THE 1990s: EL TOR

Finally, we wish to mention one other method by which we expand upon the Soho case study. Pandemic diseases (AIDS, cholera, poliomyelitis, influenza, etc.) by definition affect localities but diffuse globally. The Soho outbreak did not occur in isolation. To consider the diffusion of disease at an international scale, we therefore are developing a module based on the most recent cholera pandemic. In the 1960s, a new and somewhat milder form of cholera was identified in Celebes (Sulawase), Indonesia (WHO 1996). The disease spread rapidly to other countries in East Asia and reached Bangladesh in 1963, India in 1964, and the USSR, Iran, and Iraq in 1965-1966.

By 1970 the disease had invaded West Africa and spread in that decade across the African continent. In 1991, it spread to Latin America where cases of cholera *Vibrio*, El Tor biotype, were identified in Peru (WHO 1996). The source was believed to have been a Chinese freighter that emptied contaminated bilge water in the Lima, Peru, harbor near the city's water intakes (CDC 1991). Through the mid-1990s the disease spread from Peru northward through Central America and Mexico, with a few cases reported in Canada, the United States, French Guyana, and Venezuela.

Cholera, Snow noted, seemed to travel along global shipping routes (Snow 1855a, 104-105). He summarized the origin of a number of cholera outbreaks whose first cases were typically sailors, travelers, and those who lived near shipping places. That happened again in the modern pandemic.

Cliff and Haggett (1988, 5) have mapped the diffusion of the nineteenth-century cholera pandemics using traditional cartographic techniques. Data on the progression of the current pandemic, year by year, is available from the World Health organization (WHO

1999) and in a considerable literature based on the international diffusion of this modern strain (see for example, Arbona and Crum 1996). While there are advantages to mapping the global progression of this disease from 1961-1999, we are developing a module whose more limited focus is the diffusion of the El Tor biotype in Latin America in the 1990s. For us it presents an opportunity to discuss factors effecting the transmission of a disease across nations and regions. As important, perhaps, it offers the opportunity to consider the relationship between poverty, health infrastructure, and disease at an international scale.

Our laboratory follows previous researchers in tracking and mapping the disease's progression across the Americas. As important, however, we are adding national income, mortality rates, mean income, and other health data to the datasets of countries affected by the pandemic. That will be used to help explain cholera's greater incidence in poorer American nations with minimal public health infrastructures. In the Americas, the disease spread quickly through developing countries but only a limited number of cases are then reported in the United States or Canada. Because both are clearly involved in shipping and in pan-American trade, the inevitable conclusion drawn by students is that their relative immunity results from the type of public health infrastructure wealthy nations can afford but developing nations can only aspire to.

This again echoes Snow's work. In his 1855 opus he was at some pains to consider the effect of poverty on the incidence of cholera generally, and specifically in relation to the 1848 and 1854 outbreaks. Snow argued strongly that poverty and poor working conditions created the context in which cholera flourished in every outbreak. "It is amongst the poor, where a whole family live, sleep, cook, eat, and wash in a single room, that cholera has been found to spread when once introduced and still more in those places termed common lodging-houses, in which several families were crowded into a single room" (Snow 1855a, 18). He breaks down cholera incidence not only by population but also by trade (Snow 1855a, 23), finding that the house servants of the wealthy were least likely to be infected and that sailors and others working on the Thames were most likely to contract the illness.

At another scale he argued strongly, and collected data to demonstrate, that untreated river water was the probably source of cholera's diffusion in London and internationally. Snow's South London study showed that the vast majority of deaths occurred in jurisdictions whose waters were supplied by a single company whose water was drawn from the River Thames (Snow 1855a; 1855b; 1856). The location of that company's water intakes and outflows in London were so placed as to prevent the river's tides from rapidly carrying contaminated effluents to the open sea. It appears that during the epi-

demically that effluent would travel up and then back in the river, a source of continuing contamination and recontamination (Porter 1996; Snow 1855a, 95).

## DISCUSSION

The intent of this program has been to introduce the paradigmatic case of early epidemiology and public health to geography students studying data analysis using GIS, and to introduce GIS to those whose primary interest is in the health sciences. In discussions with students the question naturally arises, "What does GIS do that Snow didn't do manually?" Students with epidemiology and public health backgrounds typically ask, "Why is this better than statistical approaches to problems of disease diffusion?"

Typically, the answer to the first question is that the point of the lab is to use Snow's work to show basic mapping techniques. Even in this simple, localized study the GIS approach provides a number of advantages. First and foremost it is relatively fast. Snow took days to calculate his polygon around the Broad Street pump, something the GIS program does in seconds. Secondly, it provides an array of potent techniques—like nearest neighbor analysis and the spatial (or weighted mean)—unavailable to Snow and his contemporaries. Third, the ability of the GIS database to summarize and analyze data at a range of scales—local, regional, national and international—is an enormous boon.

Epidemiology and Public Health students sometimes believe GIS mapping is "just" an engine for the presentation of statistical analysis. For some, that alone is sufficient to recommend its use. More importantly, however, GIS presents its own uniquely graphical approach to spatially related data and its analysis. The answers it returns—and the questions it invites—may be very different from those returned, or even asked, by, other analytic approaches. As one of us has elsewhere argued (Koch; 2001a; 2001b), mapping provides a unique perspective on problems whose elements may be obscured by traditional statistical approaches.

No single case study can serve all purposes. Nor is any data set fully complete. Ideally, this data set would include total population of individual houses so density of occurrence could be calculated house-by-house, or street-by-street. Nor, alas, are income figures available for residents of the study area, a fact that inhibits a more detailed analysis of social attributes of the outbreak.

Finally, our approach permits at best only a limited understanding of the resistance to Snow's thesis encountered by scientists of his day. That, however, is perhaps best considered as a problem in the history of science rather than in introductory course in medical mapping. Still, the complexity of the problem from the perspective of the nineteenth-century scientist (Snow 1855a, 111-115) invites discussion of the epistemological differences separating data, information, and knowledge.

Critical evaluation of our module has been minimal to date. Our efforts have gone into the development of the primary modules' lectures and laboratory exercises rather than the resulting program's evaluation. Still, preliminary results have been encouraging. Since its introduction in 2000, the Snow case has been used three times in UBC geography's course on spatial data analysis and once in a lecture to students in an introductory GIS course. As important, in 2001 the basic module was introduced as a lecture in medical rounds at the University of British Columbia Department of Public Health and Epidemiology. The resulting interest has contributed to plans for a semester or year-long course on medical mapping that hopefully will be open both to geography students and to students in other health-related disciplines. At that time, a more rigorous evaluation of the approach we present here, and which we continue to develop, hopefully will be possible.

#### NOTES

The authors gratefully acknowledge the contribution of peer reviewers and Journal editor whose critical comments have assisted in the crafting of this article. We also here acknowledge the contribution of those UBC students who, through their interest and participation in our classes, have assisted in its development.

1. An earlier version of this work was presented as T. Koch and K. Denike, 2001, "Teaching Medical Geography through GIS: The Cholera Example," Association of American Geographers Annual Meeting, New York, February 2001.

#### REFERENCES

- Brody, H. M. R. Rip, P. Vinten-Johansen, et al. 2000. Map-making and myth-making in Broad Street: The London cholera epidemic, 1854, *The Lancet* 356, 64 - 68. <http://www.ph.ucla.edu/epi/snow/mapmyth/mapmyth.html>.
- CDC. 2000. EpiInfo Guided Tour. <http://www.cdc.gov/eppiinfo/EI2kdown.htm>.
- Centers for Disease Control (CDC). 2001. John Snow. <http://www.cdc.gov/ncidod/dbmd/snowinfo.htm>.
- Center for Disease Control (CDC). 1991. *Morbidity and Mortality Weekly Report* (MMWR) 40:6.
- Cliff, A. D., and P. Haggett. 1988. *Atlas of Disease Distributions: Analytic Approaches to Epidemiological Data*. London: Blackwell.
- Frerichs, R. R. 2001. UCLA Public Health site. <http://www.ph.ucla.edu/snow/>.
- Frost, W. H. 1936. Introduction, in *Snow on Cholera*. London: Oxford University Press.
- Gattrell, A., and M. Löytönen, eds. 1998. *GIS and Health*. Philadelphia, PA: Taylor & Francis.
- Gundersen, L. 2000. Mapping it out: Using atlases to detect patterns of health care, disease, and mortality, *Annals of Internal Medicine* 133 (2): 161-164.
- Huber, D. J. 1999. Statistical thinking/1854 cholera epidemic in London. <http://www.utexas.edu/depts/grg/huebner/grg312/lect9.html>.
- Jarco, S. 1970. Yellow fever, cholera, and the beginnings of medical cartography, *Journal of the History of Medicine and Allied Sciences* 25: 131-142.
- Koch, T. 1999. The organ transplant dilemma: What is fair? What is equitable? *OR/MS Today* 26:1, 22-28.
- Koch, T. 2001a. *Scarce Goods: Justice, Fairness, and Organ Transplantation*. Westport, CT: Praeger Publishing.
- Koch, T. 2001b. Organ transplants without borders, *National Post*, 30 April, A13.
- Koch, T. In Press. *The map as intent: Variations on the theme of John Snow*, *Cartographica*.
- Koch, T. and K. Denike. 2001. Teaching medical geography through GIS: The cholera example, Paper presented at the Association of American Geographers Annual Meeting, New York, February.
- Lang, L. 2000. *GIS for Health Organizations*. Redlands, CA: ESRI Press.
- McLeod, K. 2000. Our sense of Snow: The myth of John Snow in medical geography, *Social Science and Medicine* 40 (7&8): 923-936.
- Monmonier, M. 1997. *Cartographies of Danger: Mapping Hazards in America*. Chicago: University of Chicago Press.
- Mosely, W. G. 2001. Computer assisted comprehension of distant worlds: Understanding hunger dynamics in Africa, *Journal of Geography* 100; 32-45.
- O'Connor, E. 2000. *Raw Material: Producing Pathology in Victorian Culture*. Durham, NC: Duke University Press.

- Porter, D. H. 1998. *The Thames Embankment: Environment, Technology, and Society in Victorian London*. Akron, OH: University of Akron Press.
- Resident. 1854. A resident of Broad Street [Letter], *Times of London*, 11 September.
- Ricketts, T.C., L. A. Savit, W. M. Gesler, et al. 1997. Using Geographic Methods to Understand Health Issues. Rockville, MD: Agency for Health Care Policy and Research.
- Snow, J. 1853. On Continuous Molecular Changes, More Particularly in Their Relation to Epidemic Disease. London: John Churchill; Reprinted in J. Snow, 1936, *Snow on Cholera*, London: Oxford University Press.
- Snow, J. 1855a. On the Mode of Communication of Cholera, 2nd ed. London: J. Churchill; Reprinted in J. Snow, 1936, *Snow on Cholera*, London: Oxford University Press.
- Snow, J. 1855b. Snow's report, in *Report on the Cholera Outbreak in the Parish of St. James, Westminster, During the Autumn of 1854*. London: J. Churchill, 97-120.
- Snow, J. 1856. Cholera and the water supply in the South Districts of London in 1854, *Journal of Public Health* 2: 239-257.
- Stevenson, L. G. 1965. Putting disease on the map: The early use of spot maps in the study of yellow fever, *Journal of the History of Medical Science and Allied Sciences* 20: 226-261.
- Tobler, W. T. 2002. Snow's Cholera Map. <http://ncgia.ucsb.edu/pubs/snow.snow.html>.
- USNTI. 2000. *Atlas of United States Mortality*. Springfield VA: U.S. National Technical Information Service. <http://cdc.gov/nchs/products/pubs/pubd/other/atlas/atlas.htm>.
- Snapshot of Canada's Varied Medical Services. 2001. *National Post*, 8 May, A1, A8-9.
- Vergano, D. 2000. The operation you get often depends on where you live, *USA Today* 19 September, A1-2.
- Whitehead, H. 1855. Mr. Whitehead's report. In *Report on the Cholera Outbreak in the Parish of St. James, Westminster, During the Autumn of 1854*. London: J. Churchill, 121-167.
- World Health organization (WHO). 1999. Global Cholera Updated: 1999 notifications. <http://www.who.org/emc/diseases/cholera/choltb11999.html>.
- World Health organization (WHO). 1996. Cholera Fact Sheet N107. <http://www.who.org/inf-fs/en/fact107.html>.
- Zeiler, M. 1999. *Modeling Our World: The ESRI Guide to Geodatabase Design*. Redlands, CA: ESRI Press.