Choosing Geographic Units for Choropleth Rate Maps, with an Emphasis on Public Health Applications

Francis P. Boscoe and Linda W. Pickle

ABSTRACT: Choropleth maps are the most widely used map type for mapping rates, such as those involving disease, crime, and socioeconomic indicators. The essential step of choosing a geographic unit to map is often made in an ad hoc manner. Among the desirable characteristics of choropleth mapping units are high degree of resolution, homogeneity of population size, homogeneity of land area, observation of minimum population thresholds and land area thresholds, temporal stability and currency, compactness of shape, audience familiarity, data availability, and the functional relevance of the unit to the phenomena mapped. Because of the uneven distribution of human populations, no single geographic unit can meet all of these characteristics in practice, and a well designed choropleth map necessarily involves some compromise. We present guidelines for choosing geographic units that take into account the above criteria, considering 12 geographic units ranging from census blocks to states. Even allowing for differences in scale and purpose, some units confer clear advantages over others.

KEYWORDS: Choropleth maps, disease rates, areal units

Introduction

Choropleth rate maps are an effective tool for identifying spatial patterns in a dataset, and they are used widely for visualizing socioeconomic patterns, disease, crime, and other human geographic variables (Le et al. 1995; Champion et al. 1996; Pickle et al. 1996; Devesa et al. 1999; Murray and Shyy 2000; Brewer and Suchan 2001). They are simple to construct within a GIS software package and familiar to a wide audience. Among the many design decisions that must be made when constructing a choropleth rate map is which geographic unit or units to map. Unlike many other aspects of choropleth map design, such as categorization schemes (Cromley and Mrozinski 1999; Brewer and Pickle 2002) and color selection (Brewer et al. 1997), the choice of geographic unit has not received much formal scrutiny beyond the brief discussion found in Robinson et al. (1995), Dent (1999), and other cartographic texts. Frequently, this decision is made on an ad hoc basis (Walter and Birnie 1991; Rushton 2001; Kreiger et al. 2002). Aside from its direct relevance to cartography, the choice of geographic unit relates to fundamental issues regarding the delineation of space and multiscale analysis that have been identified as major research themes within the field of geography (Cutter et al. 2002).

In this paper we present some basic guidelines that can help inform the choice of geographic unit. Reflecting our public health orientation, most of our examples pertain to the mapping of disease rates. We begin by outlining the desirable characteristics of geographic units. We follow with a discussion of some of the cartographic, mathematical, and statistical approaches used to overcome the limitations imposed by the available units. We then report the results of an investigation of twelve different geographic units and their suitability for choropleth mapping, following Slocum's (1999) framework for choosing a method of data classification.

Francis Boscoe is a Research Scientist with the Geographic Analysis Section, Bureau of Environmental and Occupational Epidemiology, New York State Department of Health, 547 River St., Room 200, Troy, NY 12180. **Linda Pickle** is a Senior Mathematical Statistician with the Statistical Research and Applications Branch, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, 6116 Executive Blvd., Suite 504, Bethesda, MD 20892, USA. E-mail: <fpb01@health.state.ny.us>, <picklel@mail.nih.gov>.

Desirable Characteristics of Geographic Units

Desirable characteristics of geographic units include high degree of resolution, homogeneity of population size, homogeneity of land area, obser-

vation of minimum population thresholds and land area thresholds, temporal stability and currency, compactness of shape, audience familiarity, data availability and functional relevance. To the extent that these characteristics are not present, choropleth maps can lead to incorrect or incomplete inferences, including the failure to identify meaningful patterns and the identification of spurious patterns. It is seldom, if ever, possible to delineate space in such a way that all of these characteristics can be attained, in part because of the uneven distribution of human populations. A large share of the population typically occupies a small share of the land area, placing the characteristics of homogeneity of land area and homogeneity of population at odds. In discussing each of these characteristics below, we bear in mind three distinct map reading tasks: the use of maps as look-up tables to identify rates in a specific geographic unit; the use of maps to assess broader patterns and trends; and the comparison of patterns across maps (Pickle et al. 1994).

High Degree of Spatial Resolution

In general, the higher the spatial resolution of a map, the greater the information content. A county map of the United States contains approximately 60 times more geographic units than a state map, for example, and offers more potential for regional pattern identification. Increasing the spatial resolution does come at the direct expense of some of the other criteria, however. It is also the case that spatial patterns are sometimes more evident at a coarser scale than at a finer scale (Edsall et al. 2000), leading some authors to advocate the production of maps at multiple resolutions (Talbot et al. 2000). These caveats are not always well understood. In our experience, people interested in disease mapping but lacking scientific expertise often assume that maps at the finest possible spatial resolution, up to and including the individual address level, are the most informative.

Homogeneity of Population Size

Homogeneity of population size is important to prevent bias in the identification of spatial patterns. For a given map, the smaller the population size of a geographic unit, the more likely it will have an extreme rate, either high or low. This is because variance in rates is inversely proportional to population size, a property that has been termed variance instability (Anselin 1990; Gelman and Price 1999). The relationship is independent of geographic scale. On an international map, the

smallest-population countries will tend to have rates that are disproportionately extreme (i.e., in the highest or lowest quantiles); on a census block-group map, the smallest-population census block-groups will tend to have rates that are disproportionately extreme. The relationship holds even if every geographic unit is large, so that a unit with 200,000 people is more likely to have an extreme rate than a unit with 500,000 people. Figure 1 shows that nearly all of the counties in the highest and lowest quantiles for prostate cancer mortality are in rural parts of the country. A plot of rate against the number of cases for these data (Figure 2) also reveals this relationship, where the number of cases is closely correlated with the overall population size. The rates and the number of cases were obtained from the National Cancer Institute's web site (NCI 2002; see also Devesa et al. 1999).

To see how a map such as that in Figure 1 can be misleading, consider southern Georgia and the Florida panhandle, where one or more clusters of elevated mortality appear to be present. While these elevations are genuine, the map fails to convey that Middlesex County, New Jersey, is statistically similar in terms of both rate and number of cases and perhaps equally deserving of scrutiny. Similarly, Utah exhibits a consistent pattern of elevated mortality (24.7 cases per 100,000, 95 percent confidence interval 23.8-25.6) but is statistically indistinguishable from King County, Washington (24.0 [23.1-24.8]), or Suffolk County, New York (24.2 [23.3-25.1]). In our experience, many map users are given to conjecture on the dietary and health utilization patterns of rural southerners or Mormons, for example, before giving any consideration to the mid-Atlantic metropolitan corridor, where patterns tend to be less visually pronounced. Other users, overly sensitized to this problem, are apt to dismiss any rural patterns as likely statistical noise, which is also incorrect.

Homogeneity of Land Area

It is important to try to define units of similar land area to limit the natural tendency to associate area on a map with magnitude. Geographic units that take up the most space on the map command the most attention, but such attention is rarely warranted. This areal bias problem can seriously hamper pattern recognition on a choropleth map (Tukey 1979). As Tufte (1983, p. 20) observed, "Our visual impression of the data is entangled with the circumstances of geographic boundaries, shapes and areas." One of the few experiments conducted to assess map readers' perceptions of

clustering found a better correspondence between quantitative and perceived measures of clustering on maps with more regular geographic units (Walter 1993).

The goals of homogeneous population and land area are often conflicting. Large areas that draw the eye, such as the counties of Idaho, tend to have low population, so that the most prominent parts of a map have rates that are disproportionately extreme. But this problem is present even when population sizes are very homogeneous. Figure 3 shows the year 2000 census tract boundaries for Los Angeles County, California.[1] The 2000 census tracts were designed to contain between 1,500 and 8,000 persons. In Los Angeles County, 93 percent of the census tracts fell within this range (U.S. Census Bureau 2000). While the census tract populations are homogeneous, the areas are not, spanning nearly five orders of magnitude. The smallest tract (population 2,876) consists of a single block of high-rise apartments in downtown Los Angeles. The largest (population 685) covers the San Gabriel Mountains and falls mainly within the Angeles National Forest. Technically, since this tract contains fewer than 1,500 people, it should have been merged with one or more of its neighbors, which would have widened the range of areas even further.

Minimum Land Area and Population Thresholds

In the previous example, the problem was not only that the sparsely populated mountainous areas of northern Los Angeles County commanded most of the attention, but that information about the densely populated inner city was impossible to discern. It stands to reason that a well designed choropleth map should allow each individual geographic unit to be identified. Indeed, this is a necessary condition for the map to fulfill its function as a look-up table. But even when using a map to identify trends and patterns, the reader may also wish to be able to identify an individual unit, such as one with a rate very discrepant from the general pattern. Imagine, for example, a small town with a high rate of some disease surrounded by a large, thinly populated region with a low rate. Together, the town and its surrounding region have an average rate. A choropleth map in which these two units were combined would reveal the total area

to have an average rate, which is accurate, if not precise. A map in which the two units were not combined would imply that the total area had a low rate (because the value for the town would not be discernable), which is neither accurate nor precise.

The definition of a minimum visible mapping unit is influenced by the shape of the areal units, the manner in which the boundary lines are rendered, the color scheme, the viewing medium (paper or screen), the quality of the viewing medium, and the visual acuity of the viewer. Pickle et al. (1996) used a minimum size of 250 square miles for a national atlas map approximately 9 inches by 11 inches, which resulted in parts of Massachusetts, New York, and New Jersey being grouped for greater visibility. At this scale, units with a minimum of about one square millimeter were presented. Given the print quality of this atlas, this was probably a reasonable lower bound.

A minimum population threshold also needs to be observed. Consider, for example, a health outcome map involving a census tract that has a population of just one person (there were 21 such tracts in the nation in the 2000 census, including two in New York City). The rate for such a tract will either be 1, if the person experienced the outcome being measured, or 0, if the person did not (or in the usual scaling of disease rates, 100,000 or 0 per 100,000). The tract must either fall into the highest or the lowest category, but this information is not especially useful. Alternatively, consider a disease so rare that there are only 10 incident cases per year nationwide. At the tract level, rates would either be 0, for the vast majority of the tracts, or some non-zero value in the tracts where the cases occurred. In such an instance, a simple dot map is more suitable than a choropleth rate map.

Epidemiologists, in particular, have confronted this problem extensively, yet there is no single rule of thumb that is in widespread use. In one experiment, the target population for aggregation was defined as the sample size required to detect a doubling of the average disease rate with a power of 90 percent and a significance level of 95 percent (Morris and Munasinghe 1993). This meant, for example, that the target population for mapping bacterial pneumonia among white males was 1,350 but was 6,240 for rarer central nervous system cancers. Other sources state, more simply, that rates based on fewer than 20 events should be

---

[1] Census tracts, created by the Census Bureau specifically to facilitate demographic analysis, are "small, relatively permanent geographic entities" that, at the time of their creation, are as "homogeneous as possible with respect to population characteristics, economic status, and living conditions" (U.S. Census Bureau 1994).
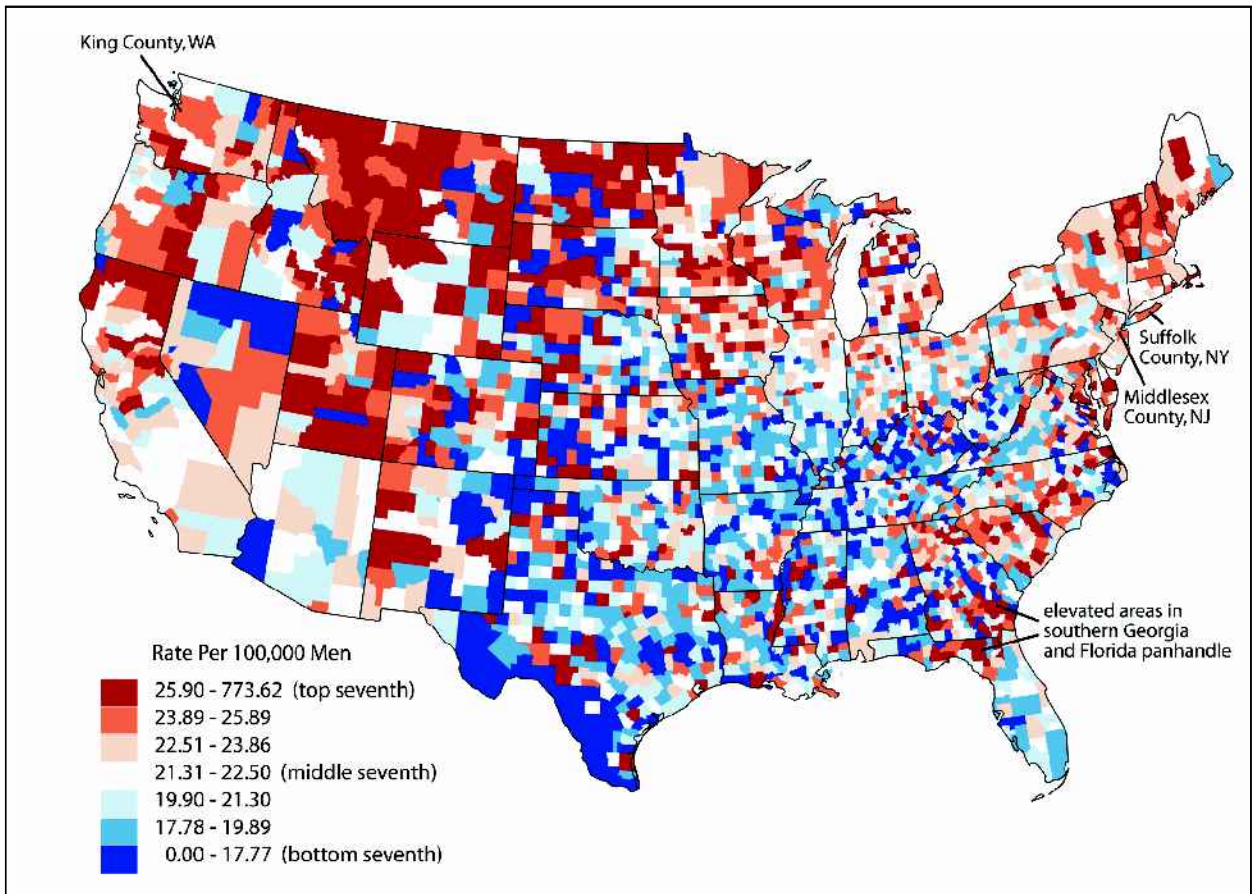
Figure 1. Prostate cancer mortality rate among white men, continental United States, by county, 1970-1994. The diverging color scheme is taken from Brewer et al. 1997. Counties mentioned in the text are highlighted.

considered unstable (National Center for Health Statistics 1994; New York State Department of Health 2001). Perhaps most commonly, if unwisely, this issue is simply ignored (Diehr 1984).

Temporal Stability and Currency

When investigating a lengthy time series or comparing patterns over time, it is important that the geographic units refer to the same areas. For the United States as a whole, only states have been entirely stable in recent decades. A handful of Western counties have been added, and towns and cities have been incorporated and unincorporated nationwide. Census tracts, block groups, and blocks are prone to change dramatically from decade to decade, particularly in growing areas, making comparisons between decades difficult. U.S. Postal Service (USPS) ZIP codes are subject to change to facilitate mail delivery, and they are most dynamic in growing areas. Custom-defined units, such as Hospital Service Areas, that are derived from these other units are necessarily impacted by any changes to them. Temporal

currency refers to whether the original defining characteristics of a geographic unit have changed over time. For example, aggregations of census tracts based on 1970 median income levels would not be an appropriate base layer on which to map 2000 crime data, quite apart from the difficulty in rectifying the different tract definitions.

Compactness, Familiarity, Data Availability, and Functional Relevance

Other desirable characteristics of geographic units include compactness of shape, familiarity, data availability, and functional relevance. Compactness of shape refers, in general, to units that have a low perimeter-to-area ratio, though there is no universally agreed upon measure (MacEachren 1985; Young 1988). Given this criterion, the square and rectangular counties of Iowa lend themselves to mapping and pattern identification more readily than the irregular counties of Idaho. Familiarity refers to geographic units that relate to "pre-existing knowledge structures" within the map reader (MacEachren 1995, p. 167).
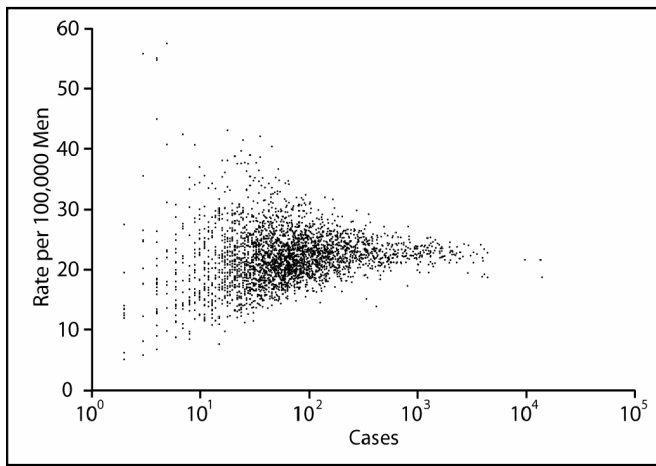
Figure 2. Relationship between the number of cases and rate for the data in Figure 1.
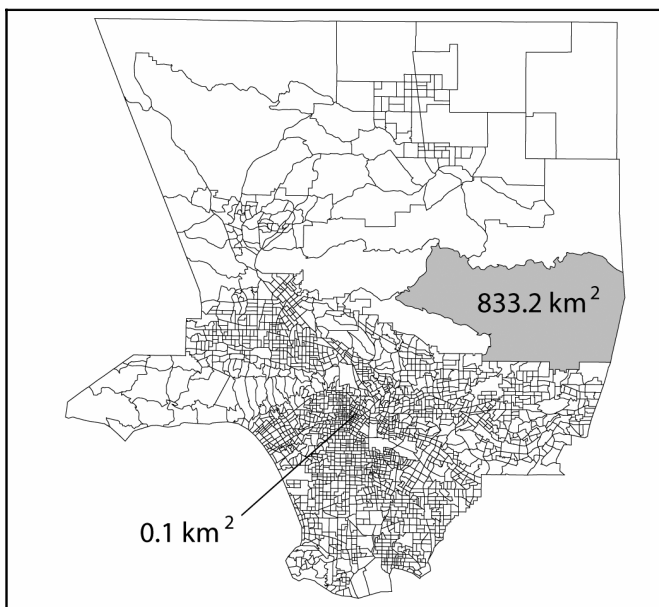

833.2 km^2

0.1 km^2

Figure 3. Census tract boundaries, Los Angeles County, showing nearly five orders of magnitude difference between the smallest and largest tracts.

According to this criterion, a state-level map is of greater value than its coarse resolution would suggest, because map readers can readily link the map content to their own pre-existing knowledge of the states. Familiarity is not limited to political divisions; "folk categorization" schemes, such as regional or neighborhood identifiers, also fit this criterion.

Data availability refers to the breadth of data that have been collected for the geographic unit in question. There is a great breadth of data available for census-defined geographic units compared with, say, ZIP codes. Custom-defined units derived from political or census-defined units have the same data availability as their source units. Related

to data availability is the idea of comparability with previously published maps. If, for example, one is interested in assessing the ecological relationship between an environmental pollutant and a disease outcome, and maps of the disease outcome have already been published at the ZIP code level, then maps of the environmental pollutant might also logically be published at the ZIP code level. This need not be a limiting factor, however, as there exist both cartographic and statistical approaches for comparing data compiled at different resolutions (Mugglin and Carlin 1998; Langford et al. 1999; Eicher and Brewer 2001).

Functional relevance refers to whether the map units are linked topically or conceptually to the theme being mapped. In some circumstances, this can be the most important criterion for choosing a mapping unit. For example, a state health department may wish to map late-stage breast cancer rates using geographic units that best match service areas for cancer screening programs in that state. Many states have such predefined "service planning areas" for purposes of policy and service decisions. Similarly, legislative districts, with their often convoluted shapes, do not lend themselves to mapping many kinds of data, but are an obvious choice for mapping election returns.

Techniques for Adding Value to Choropleth Maps

A variety of cartographic, mathematical, and statistical approaches have been employed to add value to choropleth maps and minimize the limitations described above, though none have proven entirely satisfactory (Gelman and Price 1999). Perhaps the most common are to aggregate data spatially or temporally in order to improve rate stability, meet minimum population thresholds, and limit the risk of disclosure of confidential information. Spatial aggregation can be used to help meet minimum land area and population thresholds and can result in more homogeneous populations or land area. The literature on how to aggregate geographic units optimally is extensive, and it is driven in part by the political redistricting problem, in which districts need to be defined that are equal in population, compact, and that follow logical and natural boundaries (Huel et al. 1986; Morris and Munasinghe 1993; Openshaw and Rao 1995; Carvalho et al. 1996; George et al. 1997; Mehrotra et al. 1998; D'Amico et al. 2002). Temporal aggregation, while useful

for meeting minimum population thresholds, does not impact population homogeneity, since the proportional relationship between the smallest and largest population units does not depend on the number of years of data involved.

Another commonly used technique is to suppress the data in the smallest-population units through the definition of a "sparse data" category (Le et al. 1995; Devesa et al. 1999; New York State Department of Health 2000). This technique may, however, interfere with pattern recognition, while only displacing the problem (Lewandowsky et al. 1995). If United States counties with populations below 5,000 are suppressed, for example, then the highest and lowest rates would be most likely to be found among counties with populations just slightly over 5,000. Other approaches involve highlighting only those geographic units with statistically significant high or low rates, mapping the p-values of each geographic unit, and various permutations thereof (Mason et al. 1975; Walter and Birnie 1991; Schlattmann et al. 1996). Since statistical power is a function of population size, these approaches are biased toward highlighting rates in large-population units. Most commonly used data-smoothing algorithms, including Bayesian smoothing techniques, suffer from this same problem (Pringle 1996; Gelman et al. 2000). Because small-population geographic units tend to be smoothed to a greater extent than large-population units by most smoothing methods, extreme rates end up being disproportionately found in the large-population units. Moving-average smoothing techniques are also prone to edge effects, where the probability of a geographic unit being highlighted is influenced by whether it happens to be in the center or on the edge of the area being mapped (Kafadar 1996).

Another option is to transform the data through the application of a mathematical function designed to minimize some confounding influence, such as the relationship between population size and rate (Cressie 1993). A drawback is that such functions tend to generate variables that are in different, not easily interpretable, metrics. Such functions may also be unique to each data set. Still another alternative is to generate multiple maps of possible realizations of the true underlying distribution, given the measured rates (Ehlschlaeger et al. 1997; Gelman and Price 1999; Nandram et al 2000). Examination of the consistency of rate estimates across these maps provides a visualization of rate stability. While these maps are unbiased with respect to population size, the complexity of the resulting map-reading task introduces other sources of error and bias. One method proposed to simplify the visualization of rate stability across multiple maps is to present a map showing the proportion of the total number of maps for which the rate falls in a different category than on the average map (Nandram et al. 2000). Even so, multiple map generation tends to work against the conceptual straightforwardness that makes mapping attractive in the first place.

Because rate maps tend to highlight low-population geographic units and maps showing statistical significance tend to highlight high-population units there have been numerous attempts to conflate the two onto a single map. Hatched lines have been superimposed over a choropleth map to indicate areas of uncertainty (MacEachren et al. 1998), as well as areas of statistically significant clustering (New York State Department of Health 2000). Statistically significant clusters have been stratified by rate, producing a contouring effect that resembles a smoothed map (Boscoe et al. 2003). While these maps offer improved information content, they may still be misleading. For example, in Figure 1 each individual county in North Dakota has an uncertain rate; taken together, however, the rates are significantly elevated. Although the use of hatching to indicate unreliability has been shown not to interfere with pattern recognition (MacEachren et al. 1998), a reader might be uncertain about the significance and importance of the North Dakota rates if the entire state was hatched.

In terms of making geographic areas appear more equal, the simplest technique is the inclusion of one or more inset maps at a larger scale than the primary map. On a county map of the United States, for example, it is reasonable to include inset maps for the Washington, D.C., and New York City areas. Areas can also be represented symbolically, such as by placing a color-coded or proportional point symbol at each area centroid or by placing in each area a number of point symbols proportional to the mapped statistic (a dot density map). This approach tends to result in widely spaced points in low-population areas, making pattern detection difficult (Lewandowsky et al. 1993). Geographic areas can also be distorted so that the areas themselves relate to a variable such as population, but so-called cartograms or density-equalizing map projections can be problematic to construct and interpret (Dorling 1993; Merrill et al. 1996). There are other map types, such as dasymetric maps and isarithmic maps, which potentially offer ways of reducing visual distortion arising from heterogeneous populations and areas. All of these

non-choropleth map types, however, fall beyond the scope of this paper.

Finally, it has been suggested that the best way to add value to choropleth maps might be to eschew their use altogether (Rushton 2001). The division of space into arbitrary units is seen as artificially limiting the scope of potential pattern identification, particularly when variables such as cases of infectious disease or air pollution plumes have little reason to follow any artificial boundaries. This view is motivated by the promise of the availability of data geocoded to the point level. It is true that from a data collection and analysis standpoint, all data on human populations would ideally be geocoded to the point level so that they could be aggregated to all conceivable geographic units of interest, whether political or census units, 80 meter squares, or "natural" units suggested by the data themselves. Although data collection organizations are beginning to call for this level of geographic specificity (Wiggins 2002) and GPS technology has become affordable and portable enough to permit common use, concerns of confidentiality and privacy may limit implementation of point-level geocoding of health-related data. New Health Insurance Portability and Accountability (HIPAA) regulations prohibit disclosure of information that would permit identification of any individual patient; compliance requires data aggregation or masking. Quite apart from these privacy concerns there will always remain data sets, particularly historical data sets, for which the acquisition of point-level information is impractical. Thus, while we are sensitive to the concern that choosing a geographic unit can, in some situations, mean restricting the benefits of exploratory spatial data analysis, we feel that our mapping-unit selection criteria will remain applicable for many projects for the foreseeable future.

Methods

Using the continental United States as an example, we evaluated 12 different geographic units for which data are readily available or to which data can be aggregated, including several custom units particular to public health. Each unit was then evaluated based on the desirable criteria described in this paper. Following the format of Slocum's (1999) criteria for selecting a method of classification, each unit was rated as "good," "fair," or "poor" for each criterion (Table 1). Ratings were either objective or subjective, depending on the criterion. The rating process was then repeated for the island of Manhattan, using the

four geographic units into which Manhattan can be subdivided (Table 2).

The geographic units that were evaluated included states, State Economic Areas (SEAs), counties, 3-digit ZIP Code Tabulation Areas (ZCTAs), 5-digit ZCTAs, county subdivisions, Census tracts, Census block groups, Census blocks, Hospital Referral Regions (HRRs), Health Service Areas, and Hospital Service Areas. For the first nine units, definitions follow those of the Census Bureau (U.S. Census Bureau 1994; 2000; 2002a; 2002b). The latter three units are aggregations of counties or ZIP codes that were constructed by health researchers for specific purposes. Since not all of these are likely to be familiar to all readers, a brief summary follows.

Counties, known as parishes in Louisiana, are governmental units found in each state; they also include independent cities, most of which are in Virginia, and the District of Columbia. County subdivisions are mainly comprised of cities, towns, and townships. The Census Bureau has created artificial county subdivisions in states lacking these political units. Five-digit ZIP code tabulation areas are Census Bureau-defined areas that correlate as closely as possible with five-digit ZIP codes defined by the United States Postal Service. Three-digit ZCTAs are aggregations of five-digit ZCTAs, aggregated by the first three digits. Census tracts are "small, relatively permanent geographic entities" which, at the time of their creation, are as "homogeneous as possible with respect to population characteristics, economic status, and living conditions" (U.S. Census Bureau 1994). Census block groups are subdivisions of tracts, with each tract averaging about four block groups. Census blocks are the smallest units tabulated by the Census Bureau, consisting of areas enclosed by streets, roads, railroads, other physical features, and political boundaries. On average there are about 100 populated blocks per tract, but the actual number varies enormously depending on local population density. Roughly one-third of all blocks have a population of zero; these, along with a small number of other zero-population units, were excluded from the analysis.

The remaining units are aggregations of counties or ZIP codes constructed by other agencies or institutions for specific health-related purposes. The Dartmouth Atlas of Health Care Project has defined two geographic units, both of which are aggregations of five-digit ZIP codes as delineated by a private vendor (Dartmouth Medical School 1999). Hospital Service Areas are collections of ZIP codes whose residents are typically served by

a specific hospital or hospitals. Hospital referral regions (HRRs) represent health care markets for tertiary care, with each HRR containing at least one hospital that performs major cardiovascular procedures and neurosurgery. Note that we developed populations and areas for the Dartmouth units using ZCTAs rather than their ZIP code data. State economic areas (SEAs) are counties or groups of counties that were homogeneous with respect to various demographic factors. First defined by the Census Bureau after the 1950 census and revised after the 1960 census, they are still in limited use (Devesa et al. 1999). Health Service Areas are counties or groups of counties designed by the National Center for Health Statistics to be self-contained with respect to hospital care (Makuc et al. 1991). We use the amended version that was implemented in the Atlas of United States Mortality (Pickle et al. 1996).

Homogeneity of population size and homogeneity of land area were examined using data from the 2000 census (U.S. Census Bureau 2001). These were rated quantitatively by constructing a histogram of population and area values on a logarithmic scale (Figure 4). The position on the x-axis shows the relative coarseness or fineness of scale, while the spread of each curve indicates the homogeneity of the spatial unit. The spread, as quantified by the geometric standard deviation, was used to assign the good, fair, and poor ratings. Note the homogeneity of census tract and block group populations, and the homogeneity of county and Health Service

Table 1. Evaluation of mapping unit criteria for a map of the continental United States.

| Geographic Unit	Spatial Resolution (# Units)	Homogeneous Population Size (geom. s.d.)	Homogeneous Land Area (geom. s.d.)	Minimum Population Threshold	Minimum Land Area Threshold	Temporal Stability	Temporal Currency	Compactness	Familiarity	Data Availability/ Compatibility	Functional Relationship
Geographic Unit	Poor (49)	Fair (0.44)	Fair (0.62)	Good	Good	Good	Good	Fair	Good	Good	Fair
Hospital Referral Region	Fair (304)	Good (0.37)	Fair (0.57)	Good	Good	Good	Good	Fair	Poor	Poor	Good
State Economic Area	Fair (506)	Good (0.39)	Fair (0.55)	Good	Good	Poor	Poor	Poor	Poor	Poor	Poor
Health Service Area	Fair (798)	Poor (0.60)	Good (0.38)	Good	Good	Good	Fair	Poor	Poor	Poor	Good
Three-digit ZCTA	Fair (877)	Fair (0.44)	Poor (0.75)	Good	Good	Good	Fair	Fair	Poor	Good	Poor
County	Good (3,109)	Poor (0.61)	Good (0.38)	Good	Good	Good	Good	Good	Good	Good	Good
Hospital Service Area	Good (3,408)	Poor (0.58)	Fair (0.55)	Fair	Fair	Poor	Fair	Fair	Poor	Poor	Good
Five-digit ZCTA	Good (31,783)	Poor (0.78)	Poor (0.80)	Poor	Poor	Poor	Good	Poor	Poor	Poor	Poor
County Subdivision	Good (35,157)	Poor (0.80)	Fair (0.68)	Poor	Poor	Good	Good	Good	Fair	Fair	Fair
Census Tract	Good (64,735)	Good (0.28)	Poor (0.97)	Fair	Poor	Poor	Good	Fair	Poor	Good	Fair
Census Block Group	Good (206,920)	Good (0.27)	Poor (0.96)	Fair	Poor	Poor	Good	Fair	Poor	Fair	Fair
Census Block	Good (5,280,212)	Poor (0.59)	Poor (0.93)	Poor	Poor	Poor	Good	Poor	Poor	Poor	Poor
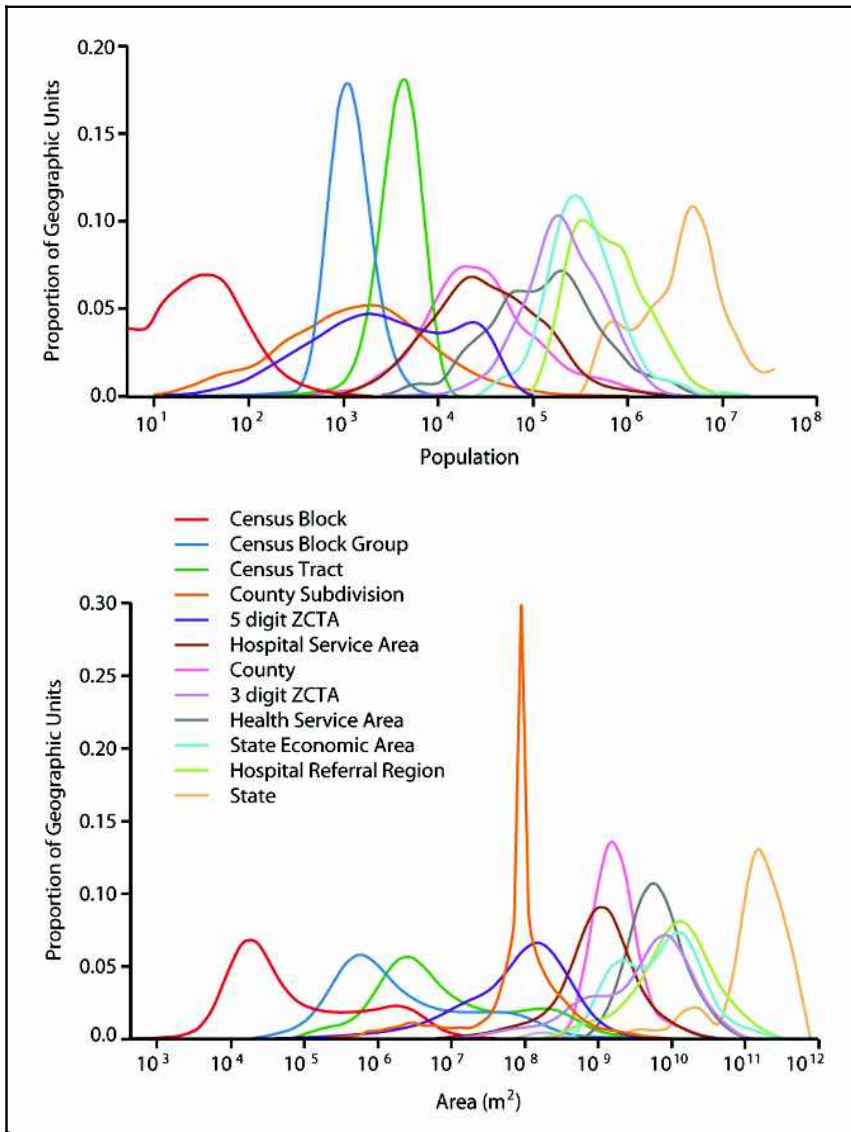
Figure 4. Distribution of geographic units by population and area, continental United States. For clarity of presentation, the histograms were smoothed using a locally weighted regression smoothing algorithm (Cleveland 1979, *S-Plus 2000 User's Guide* 2000). A bandwidth (smoothing span) of 12.5% of the x-axis range was used for all geographic units except for the area of county subdivisions. For county subdivisions a bandwidth of 4% was used so as not to obscure the sharp peak of the distribution of area.

Area land areas. The spike at 93 million square meters in the county subdivision curve is a legacy of the Government Land Office, which, beginning in 1812 divided lands west of Ohio into 36 square mile townships (Meinig 1995). Spatial resolution, minimum population thresholds, and minimum land areas also made use of objective measures in determining the ratings. Because the latter two criteria are sensitive to the rarity of the phenomenon being mapped, we assumed a map of a relatively common disease outcome (lung cancer mortality) over a 10-year time period. Other cri-

teria were evaluated in qualitative fashion. In determining our ratings, we also presumed a map author with some GIS expertise who is familiar with the Census web site and other basic means of acquiring data.

Discussion

Tables 1 and 2 suggest that while all units have their merits there are many situations where one unit may be clearly favored over another. For the continental U.S., states, counties and Health Service Areas offer consistently better ratings than Hospital Referral Regions, Hospital Service Areas, State Economic Areas, or 3-digit ZCTAs. For Manhattan, census tracts and census block groups confer more advantages than five-digit ZCTAs or census blocks, with all four units having better ratings for Manhattan than for the nation as a whole. Census blocks rate consistently poorly despite promising the greatest resolution; again, in our experience, spatial resolution is sometimes inappropriately valued above all other criteria. The comparison between counties and Hospital Service Areas is of particular interest because the two units have similar spatial resolutions. Counties fare better on six measures, while Hospital Service Areas fare better on only one (functional relationship). A map author would have to give serious thought as to whether the use of this proxy measurement of hospital utilization was important enough to offset its several drawbacks.

Because Tables 1 and 2 are sensitive to particular geographic locations, particular data sets, and particular output formats, it is not possible to draw universal generalizations about which mapping units are ideal. We suspect that ZIP codes and ZIP-code-derived measures have too many factors working against them to ever rate well, but even

here there could be exceptional circumstances that dictate the use of this mapping unit, such as a telephone survey where the results were tabulated only by five-digit ZIP code. We do recommend that the construction of a table with advantages and disadvantages of each available unit, such as Table 1, be an early and prominent part of the planning process for any substantive choropleth mapping project. We also recommend that the largest number of mapping units should be shown, so long as each unit is clearly distinguishable from its neighbors and each unit has a sufficiently large enough population so that the calculated rates are stable. Where units are either too small in area or population, they should be aggregated with neighboring units. Overly large units should be split if the data required for mapping the sub-units are available. When resources permit, researchers may consider creating custom units designed to maximize the desirable characteristics of choropleth maps; ideally, these custom units should be defined in terms of other units for which the data were collected.

Although this paper focused on the choice of geographic unit for mapping, many of the issues discussed are also relevant for GIS and spatial analysis more generally. The problem of variance instability, for instance, is pervasive in many data sets and study designs that lack a specific mapping component. Aggregation of disease rate data is often necessary in order to permit analysis at units for which covariate data exist. Availability of disease rate data is also limited to certain minimum-sized units by legal restrictions or confidentiality concerns. The expectation that finer-resolution data yield more robust analytical results is generally well founded, but the costs and benefits of collecting finer-resolution data need to be weighed against overall project goals.

Despite their shortcomings, choropleth maps are simple to construct and widely understood. Paying careful attention to the units of analysis at the design stage can help to mitigate some of their potential shortcomings and perhaps allay the need for more complex cartographic presentation styles. The 11 criteria presented here offer a systematic means of informing this choice.

Table 2. Evaluation of mapping unit criteria for a map of Manhattan.

	Spatial Resolution (# of Units)	Homogenous Population Size	Homogenous Land Area	Minimum Population Threshold	Minimum Land Area Threshold	Temporal Stability	Temporal Currency	Compactness	Familiarity	Data Availability/ Compatibility	Functional Relationship
Five-digit ZCTA	Fair (48)	Poor (1.32)	Fair (0.63)	Poor	Poor	Good	Good	Good	Poor	Poor	Poor
Census Tract	Good (294)	Fair (0.56)	Good (0.25)	Fair	Good	Fair	Good	Good	Poor	Good	Fair
Census Block Group	Good (866)	Fair (0.47)	Good (0.27)	Fair	Fair	Fair	Good	Good	Poor	Fair	Fair
Census Block	Good (2,541)	Poor (0.72)	Good (0.30)	Poor	Poor	Poor	Good	Good	Poor	Poor	Poor

REFERENCES

Anselin, L. 1990. Spatial dependence and spatial structural instability in applied regression analysis. Journal of Regional Science 30(2): 185-207.

Boscoe, F.P., C.C. McLaughlin, M.J. Schymura, and C.L. Kielb. 2003. Visualization of the spatial scan statistic using nested circles. Health and Place 9(3): 273-277.

Brewer, C.A., and T.A. Suchan. 2001. Mapping Census 2000: The geography of U.S. diversity. Washington, D.C.: U.S. Government Printing Office.

Brewer, C.A., and L.W. Pickle. 2002. Evaluation of methods for classifying epidemiological data on choropleth maps in series. Annals of the Association of American Geographers 82(4): 662-81.

Brewer, C.A., A.M. MacEachren, L.W. Pickle, and D. Herrmann. 1997. Mapping mortality: Evaluating color schemes for choropleth maps. Annals of the Association of American Geographers 87(3): 411-38.

Carvalho, M.S., O.G. Cruz, and F.F. Nobre. 1996. Spatial partitioning using multivariate cluster analysis and a contiguity algorithm. Statistics in Medicine 15(17-18): 1885-94.

Champion, T., C. Wong, A. Rooke, D. Dorling, M. Coombes, and C. Brundson. 1996. The population of Britain in the 1990s: A social and economic atlas. Oxford, U.K.: Clarendon Press.

Cleveland W.S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74(368): 829-36.

Cressie, N.A.C. 1993. Statistics for spatial data. New York, New York: Wiley.

Cromley, R.G., and R.D. Mrozinski. 1999. The classification of ordinal data for choropleth mapping. Cartographic Journal 36(2): 101-109.

Cutter, S.L., R. Golledge, and W.L. Graf. 2002. The big questions in geography. Professional Geographer 54(3): 305-17.

D'Amico, S.J., S.J. Wang, R. Batta, and C.M. Rump. 2002. A simulated annealing approach to police district design. Computers and Operations Research 29(6): 667-84.

Dartmouth Medical School. Center for the Evaluative Clinical Sciences. 1999. The Dartmouth atlas of health care 1999. Chicago, Illinois: American Hospital Publishing.

Dent, B.D. 1999. Cartography: Thematic Map Design, 5th edition. Boston, Massachusetts: WCB/Mc-Graw-Hill.

Devesa, S.S., D.J. Grauman, W.J. Blot, G.A. Pennello, R.N. Hoover, and J.F. Fraumeni, Jr. 1999. Atlas of cancer mortality in the United States, 1950-94. NIH Publication Number 99-4564. National Institutes of Health, Bethesda, Maryland.

Diehr, P. 1984. Small area statistics: Large statistical problems. American Journal of Public Health 74(4): 313-14.

Dorling, D. 1993. Map design for Census mapping. Cartographic Journal 30(2): 167-83.

Edsall, R.M., M. Harrower, and J.L. Mennis. 2000. Tools for visualizing properties of spatial and temporal periodicity in geographic data. Computers & Geosciences 26(1): 109-18.

Ehlschlaeger C.R., A.M. Shortridge, and M.F. Goodchild. 1997. Visualizing spatial data uncertainty using animation. Computers & Geosciences 23(4): 387-95.

Eicher, C.L., and C.A. Brewer. 2001. Dasymetric mapping and areal interpolation: Implementation and evaluation. Cartography and Geographic Information Science 28(2): 125-38.

George, J.A., B.W. Lamar, and C.A. Wallace. 1997. Political district determination using large-scale network optimization. Socio-Economic Planning Sciences 31(1): 11-28.

Gelman, A., and P.N. Price. 1999. All maps of parameter estimates are misleading. Statistics in Medicine 18(23): 3221-34.

Gelman, A., P.N. Price, and C. Lin. 2000. A method for quantifying artefacts in mapping methods illustrated by application to headbanging. Statistics in Medicine 19(17-18): 2309-20.

Huel, G., J.F. Petiot, and P. Lazar. 1986. Algorithm for the grouping of contiguous geographic zones. Statistics in Medicine 5(2): 171-81.

Kafadar, K. 1996. Smoothing geographical data, particularly rates of disease. Statistics in Medicine 15(23): 2539-60.

Kreiger, N., J.T. Chen, P.D. Waterman, M. Soobader, S.V. Subramanian, and R. Carson. 2002. Geocoding and monitoring of U.S. socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? American Journal of Epidemiology 156(5): 471-82.

Langford, I.H, A.H. Leyland, J. Rasbash, H. Goldstein, R.J. Day, and A. McDonald. 1999. Multilevel modelling of area-based health data. In: A. Lawson, A. Biggeri, D. Böning, E. Lesaffre, J.-F. Viel, and R. Bertollini (eds), Disease mapping and risk assessment for public health. Chichester, U.K.: John Wiley & Sons. pp. 217-28.

Le, N., L.D. Marrett, D.L. Robinson, R.M. Semenciw, D. Turner, and S.D. Walter. 1995. Canadian cancer incidence atlas. Ministry of Supply and Services, Ottawa, Canada.

Lewandowsky, S., D.J. Herrmann, J.T. Behrens, S.-C. Li, L. Pickle, and J. Jobe. 1993. Perception of clusters in statistical maps. Applied Cognitive Psychology 7: 533-51.

Lewandowsky, S., J.T. Behrens, L.W. Pickle, D.J. Herrmann, and A.A. White. 1995. Perception of clusters in mortality maps: Representing magnitude and statistical reliability. In: L.W. Pickle and D.J. Herrmann (eds), Cognitive aspects of statistical mapping. Centers for Disease Control, National Center for Health Statistics, Washington, D.C. pp. 107-32.

MacEachren, A.M. 1985. Compactness of geographic shape: Comparison and evaluation of measures. Geografiska Annaler B 67(1): 53-67.

MacEachren, A.M. 1995. How maps work: Representation, visualization, and design. New York, New York: Guilford.

MacEachren, A.M., C.A. Brewer, and L.W. Pickle. 1998. Visualizing georeferenced data: Representing reliability of health statistics. Environment and Planning A, 30(9): 1547-61.

Makuc, D.M., B. Haglund, D.D. Ingram, J.C. Kleinman, and J.J. Feldman. 1991. Health service areas for the United States. Vital and Health Statistics 2(112).

Mason, T.J., F.W. McKay, R. Hoover, W.J. Blot, and J.F. Fraumeni. 1975. Atlas of cancer mortality for U.S. counties: 1950-1969. National Institutes of Health, Washington, D.C.

Mehrotra, A., E.L. Johnson, and G.L. Nemhauser. 1998. An optimization based heuristic for political districting. Management Science 44(8): 1100-14.

Meinig, D.W. 1995. The shaping of America, a geographic perspective on 500 years of history: Continental America, 1800-1867. New Haven, Connecticut: Yale University Press.

Merrill, D.W., S. Selvin, E.R. Close, and H.H. Holmes. 1996. Use of density equalizing map projections (DEMP) in the analysis of childhood cancer in four California counties. Statistics in Medicine 15(17-18): 1837-48.

Morris, R.D., and R.L. Munasinghe. 1993. Aggregation of existing geographic regions to diminish spurious variability of disease rates. Statistics in Medicine 12(19-20): 1915-29.

Mugglin, A.S., and B.P. Carlin. 1998. Hierarchical modeling in geographic information systems: Population interpolation over incompatible zones. Journal of Agricultural, Biological, and Environmental Statistics 3: 111-30.

Murray, A.T., and T.K. Shyy. 2000. Integrating attribute and space characteristics in choropleth display and spatial data mining. International Journal of Geographic Information Science 14(7): 649-67.

Nandram, B., J. Sedransk, and L.W. Pickle. 2000. Bayesian analysis and mapping of mortality rates for chronic obstructive pulmonary disease. Journal of the American Statistical Association 95(452): 1110-18.

National Center for Health Statistics. 1994. Vital statistics of the United States, 1990, mortality technical appendix. Public Health Service, Washington, D.C.

NCI (National Cancer Institute). 2002. Cancer mortality maps & graph web site. [http://www3.cancer.gov/atlasplus/. Accessed August 2, 2002].

New York State Department of Health. 2000. New York State Cancer Surveillance Improvement Initiative. [http://www.health.state.ny.us/nysdoh/cancer/csii/nyscsii.htm].

New York State Department of Health. New York State Cancer Registry. 2001. Cancer incidence and mortality in New York State. Volume 1: Cancer incidence and mortality by county, 1994-1998. New York State Department of Health, Albany, New York.

Openshaw, S., and L. Rao. 1995. Algorithms for reengineering 1991 census geography. Environment and Planning A 27(3): 425-46.

Pickle, L.W., D. Herrmann, J. Kerwin, C.M. Croner, and A.A. White. 1994. The impact of statistical graphic design on interpretation of disease rate maps. In: Proceedings of the 1993 Annual Meeting of the American Statistical Association, Statistical Graphics Section. pp. 111-16.

Pickle, L.W., M. Mungiole, G.K. Jones, A.A. White. 1996. Atlas of United States mortality. National Center for Health Statistics, Hyattsville, Maryland.

Pringle, D.G. 1996. Mapping disease risk estimates based on small numbers: An assessment of empirical Bayes techniques. Economic and social review 27(4): 341-63.

Robinson, A.H, J.L. Morrison, P.C. Muehrcke, A.J. Kimerling, and S.C. Guptill. 1995. Elements of cartography, 6th edition. New York, New York: John Wiley and Sons.

Rushton, G. 2001. Selection of areal units for mapping cancer incidence and mortality rates by state cancer registries. Abstracts, Association of American Geographers 97th Annual Meeting, New York City. Washington, DC: Association of American Geographers.

Schlattmann, P., E. Dietz, and D. Bohning. 1996. Covariate adjusted mixture models and disease mapping with the program DismapWin. Statistics in Medicine 15(7-9): 919-29.

S-Plus 2000 User's Guide. 2000. MathSoft, Data Analysis Products Division, Seattle, Washington.

Slocum, T.A. 1999. Thematic cartography and visualization. Upper Saddle River, New Jersey: Prentice Hall.

Talbot, T.O., M. Kulldorff, S.P. Forand, and V.B. Haley. 2000. Evaluation of spatial filters to create smoothed maps of health data. Statistics in Medicine 19(17-18): 2399-2408.

Tufte, E.R. 1983. The visual display of quantitative information. Cheshire, Connecticut: Graphics Press.

Tukey, J.W. 1979. Statistical mapping: What should not be plotted. In: Proceedings of the 1976 Workshop on Automated Cartography and Epidemiology. DHEW Publication No. (PHS) 79-1254, Arlington, Virginia. pp. 18-33.

United States Census Bureau. 1994. Geographic areas reference manual. Washington, D.C.: U.S. Government Printing Office.

United States Census Bureau. 2000. Census tracts and block numbering areas. [http://www.census.gov/geo/www/cen_tract.html. Accessed August 1, 2002].

United States Census Bureau. 2001. Census 2000 summary file 1 [United States]. [http://www2.census.gov/census_2000/datasets/Summary_File_1.Accessed August 2, 2002].

United States Census Bureau. 2002a. ZIP Code Tabulation Areas (ZCTAs). [http://www.census.gov/geo/ZCTA/zcta.html. Accessed August 1, 2002].

United States Census Bureau. 2002b. Reference resources for understanding Census Bureau geography. [http://www.census.gov/geo/www/reference.html. Accessed August 1, 2002].

Walter, S.D., and S.E. Birnie. 1991. Mapping mortality and morbidity patterns: An international comparison. International Journal of Epidemiology 20(3): 678-89.

Walter, S.D. 1993. Visual and statistical assessment of spatial clustering in mapped data. Statistics in Medicine 12(14): 1275-91.

Wiggins, L. (ed.). 2002. Using geographic information systems technology in the collection, analyses and presentation of cancer registry data: A handbook to basic practices. North American Association of Central Cancer Registries, Springfield, Illinois.

Young, H.P. 1988. Measuring the compactness of legislative districts. Legislative Studies Quarterly 13(1): 105-15.