

Review

Open Access

## A brief visual primer for the mapping of mortality trend data

Wesley L James\*<sup>1</sup>, Ronald E Cossman<sup>1</sup>, Jeralynn S Cossman<sup>2</sup>,  
Carol Campbell<sup>1</sup> and Troy Blanchard<sup>2</sup>

Address: <sup>1</sup>Social Science Research Center, 103 Research Park, Mississippi State, MS, 39762-5287, Mail Stop #9628, USA and <sup>2</sup>Department of Sociology, Anthropology and Social Work, Mississippi State University, 103 Research Park, Mississippi State, MS, 39762-5287, Mail Stop #9628, USA

Email: Wesley L James\* - wes.james@ssrc.msstate.edu; Ronald E Cossman - rcossman@ssrc.msstate.edu;  
Jeralynn S Cossman - lcossman@ssrc.msstate.edu; Carol Campbell - carol.campbell@ssrc.msstate.edu;  
Troy Blanchard - tblanchard@ssrc.msstate.edu

\* Corresponding author

Published: 08 April 2004

Received: 30 January 2004

*International Journal of Health Geographics* 2004, **3**:7

Accepted: 08 April 2004

This article is available from: <http://www.ij-healthgeographics.com/content/3/1/7>

© 2004 James et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

Maps are increasingly used to visualize and analyze data, yet the spatial ramifications of data structure are rarely considered. Data are subject to transformations made throughout the research process and then used to map, visualize and conduct spatial analysis. We used mortality data to answer three research questions: Are there spatial patterns to mortality, are these patterns statistically significant, and are they persistent across time? This paper provides differential spatial patterns by implementing six data transformations: standardization, cut-points, class size, color scheme, spatial significance and temporal mapping. We use numerous maps and graphics to illustrate the iterative nature of mortality mapping, and exploit the visual nature of the *International Journal of Health Geographics* journal on the World Wide Web to present researchers with a series of maps.

### Review

#### Introduction

Important and substantial differing conclusions will result from statistical analysis, based on the manner in which variables are defined, operationalized, calculated and standardized. Variation across these transformations of data must be addressed, as changes in operationalization result in different outcomes [1]. Increasingly for social scientists, maps function as both spatial representations of data, as well as tools for exploratory spatial data analysis [2]. To explore the spatial patterns of the ultimate health outcome, death, we calculated and mapped United States mortality rates at the county level. In one study, age-adjusted, five-year-averaged, all-cause mortality rates at the county level were calculated using data from the Compressed Mortality File [3]. In this research [4] we found

visual and statistical evidence of spatial clustering of relatively high and low mortality rates in several regions across the United States. This paper recounts the data transformations and resulting spatial patterns that were evident in this inquiry.

A wide range of studies has examined different determinants of health and clustering of health outcomes, either within demographic and socioeconomic classifications or spatial location [5-7]. In this paper we visually demonstrate how variations in methodology result in different spatial patterns of mortality, some striking and others less dramatic. We make full use of the visual nature of publishing on the Internet to provide examples from different methodologies with the use of maps and charts. These graphics are a product of the many steps taken throughout

the research process in order to reach valid and reliable mortality data calculations. Reframing our empirical exercise as a research question, after examining variation in these results, we ask: do ecological level measures of mortality cluster differently based on how mortality is standardized, measured and operationalized? The answer is unequivocally yes and conditionally no. We explain these results more fully in this paper with an eye toward providing guidance for other investigators.

### **Popularity and importance of mapping**

The wide variety of health atlases, categorized by population or disease, is testimony to the popularity and value of mapping health outcomes in both print form [8-10] and on the Web [11-14]. This popularity is due in large part to its effectiveness in data analysis. Beginning with Dr. John Snow's London cholera maps [15], health researchers have exploited the advantages of data visualization. Detecting spatial patterns in social, economic, and health variables influences research by showing where these phenomena exist, their intensity and their spatial anchoring over time. Investigators should be reminded that initial decisions regarding research approach and assumptions fundamentally affect the resulting spatial patterns. All subsequent decisions (e.g., operationalization, transformation, classification and standardization) necessarily affect the resulting spatial pattern as well.

Variable distributions and data methodology have long been issues related to mapping, as Jenks and Caspall [16] have outlined. According to Monmonier [17], "Social scientists need maps to explore and understand their data and to confirm and refine their hypotheses." These researchers have acknowledged the importance between underlying data and its relationship to the visual output it produces in a choropleth map. This relationship is further explored throughout the article using county-level mortality data.

### **Statement of the empirical research**

Three research questions are derived from our empirical research project of healthy and unhealthy places in America: Are there spatial patterns to mortality, are these patterns statistically significant, and are they persistent across time? A series of mapped examples is used demonstrating the variety of geographic patterns that emerge across each methodological technique. These maps include high and low mortality standardizations, cut points (standard deviations, natural breaks, quantiles), class size and color scheme, spatial significance of high and low mortality clusters, and temporal persistence of these clusters. Summary maps and charts are provided, where necessary, to highlight the levels of change that exist across methodological techniques.

### **United States mortality 1993–1997**

Using five-year-average age-adjusted mortality rates for 1993–1997, distinct high and low mortality clusters appear around the United States. Five-year averages are calculated to provide rate stability for counties with small populations, where a minor increase or decrease in deaths for a single year may cause a dramatic change in the county's mortality rate for that particular year [4].

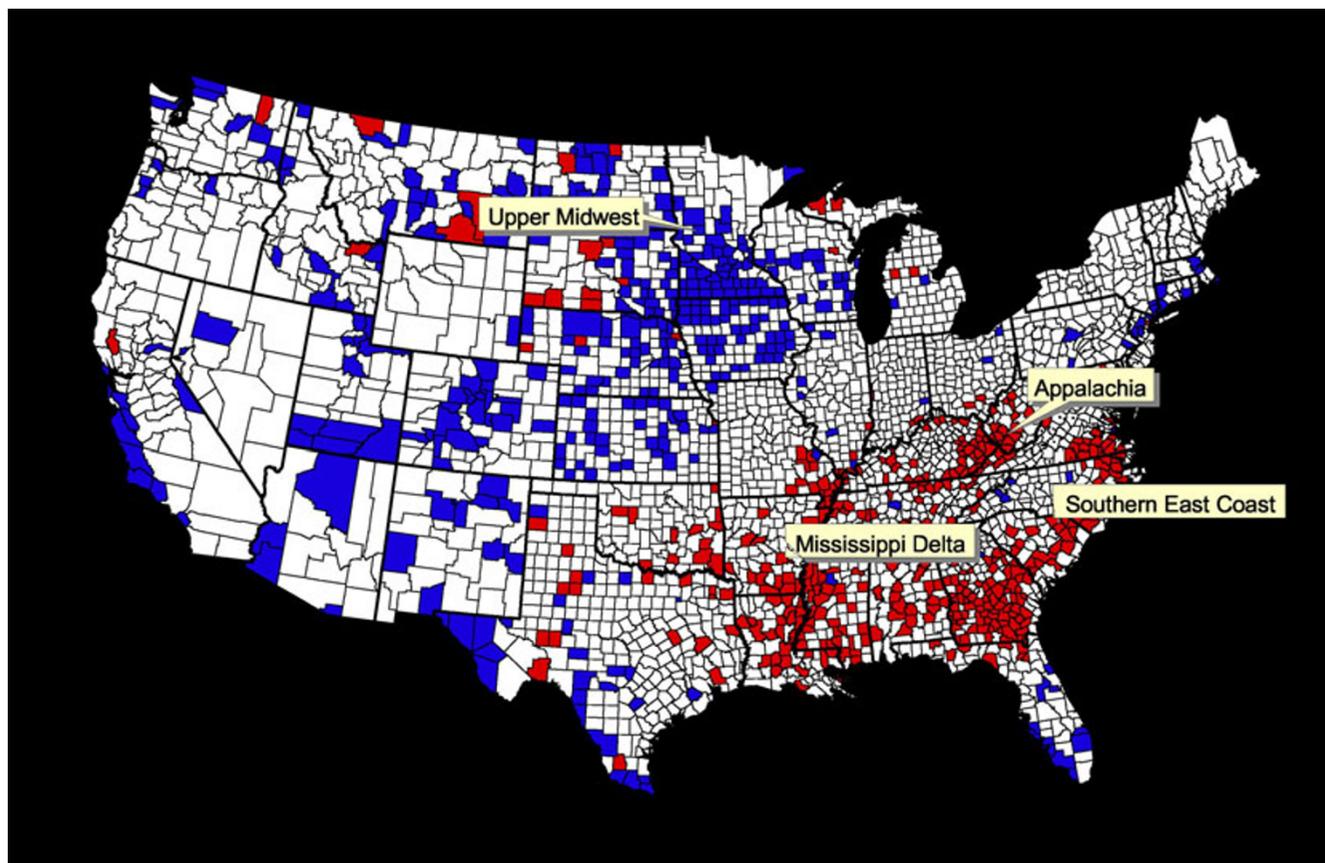
These high and low clusters beg the question, how are high and low mortality measured, and what differences exist when rates are not standardized or are done so in another fashion? In an attempt to address these concerns, illustrations are used throughout to acknowledge the importance of methodology when conducting spatial data analysis.

Regarding the U.S. mortality map for 1993–1997 (Figure 1), high mortality is defined as any county greater than one standard deviation above the mean mortality rate for the five-year-average. Also, any county greater than one standard deviation below the mean is defined as low mortality. Every county within one standard deviation of the mean is classified together with the national mean for this particular time period. This methodology results in high mortality clusters located in the southern East Coast (parts of Virginia, North Carolina, South Carolina, and Georgia), Appalachia (parts of Tennessee, Kentucky, Virginia, and West Virginia), and the Mississippi Delta (parts of Arkansas, Mississippi, and Louisiana). Low mortality clusters are predominantly found in the Upper Midwest (parts of North Dakota, South Dakota, Nebraska, Minnesota, Wisconsin, Iowa, and Kansas) and are scattered throughout the remaining sections of the western half of the country. These spatial outcomes are derived from a three-classification system, high mortality, average mortality, and low mortality.

The subsequent sections of this article follow the line of research in the "Healthy and Unhealthy Clusters in America" [4] project of this research team, detail the important differences that emerge in spatial outcomes as various methodological techniques are applied to the data and measurement of variables and classes, and outline potential impacts on research.

### **Standardization methods/rate adjustments**

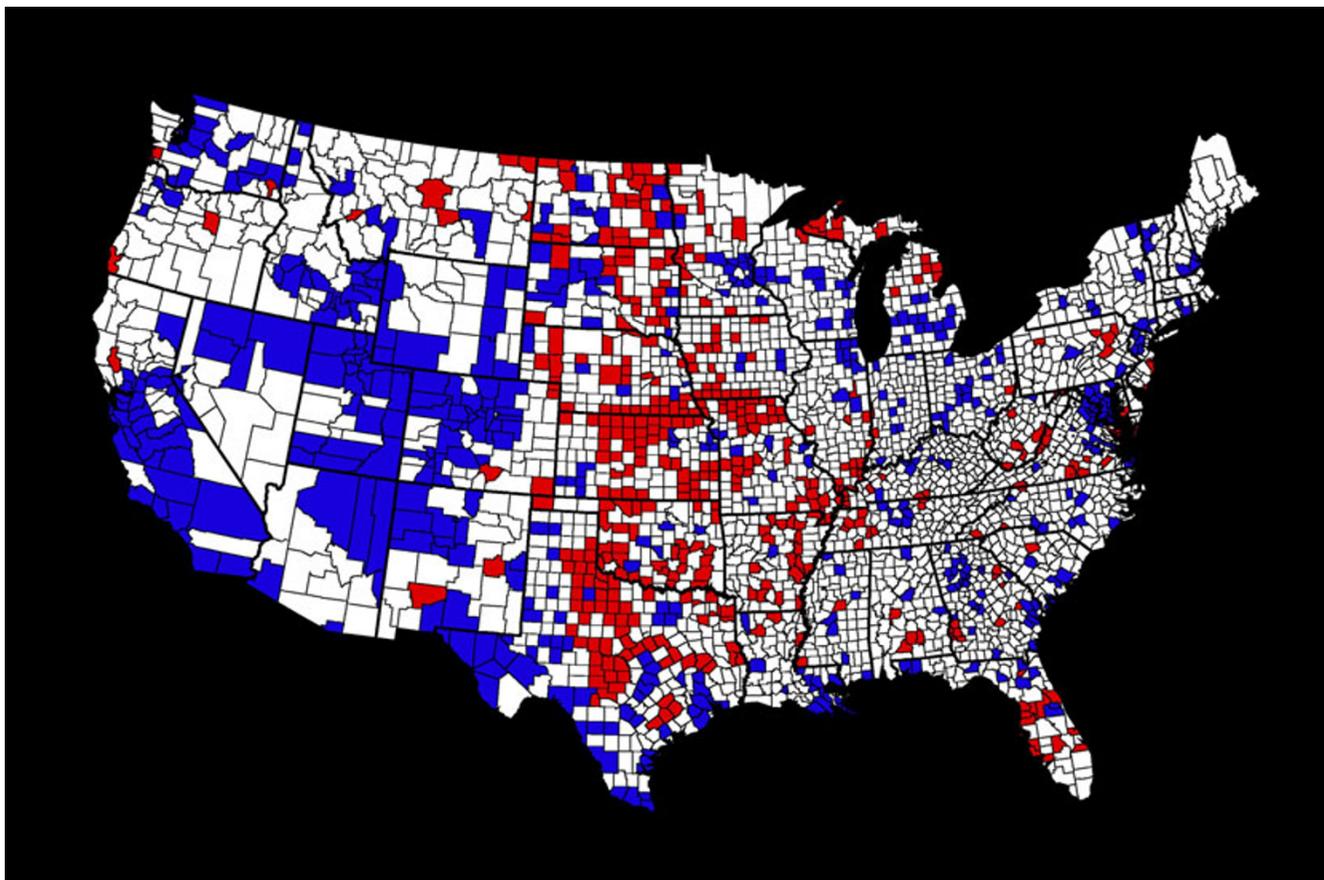
The way in which the dependent variable, mortality rate, is adjusted/standardized must be acknowledged when mapping or analyzing data of any sort. The importance of this acknowledgement lies in the substantial differences in spatial outcomes and analysis across a variety of adjustment techniques. Several techniques of standardization are possible. In this research crude rates, age, age-sex, and age-sex-race rate adjustments are used. Crude death rates



**Figure 1**  
Age-Adjusted Mortality 1993–1997 Red = High Mortality White = Normal Mortality Blue = Low Mortality

are simply the number of deaths per 100,000 of the population. While simple to calculate, the disadvantage is that demographic differences in populations are not reflected in the rate and thus counties cannot be compared directly. Age-adjusted rates are standardized by age, which adjusts each age group of each county, or unit of analysis, to represent the proportion of the total population of that specific age group. The same is done for age and race, as well as for age, race, and sex, standardized rates. The principle behind this standardization technique is that it makes the appropriate adjustments to the county rate so that the demographic profile of the county mirrors the demographics of the entire country, thus facilitating direct comparisons. If a county has an above average minority or elderly population, it is adjusted accordingly, therefore eliminating the possibility that this particular population contributes to even higher death rates. The spatial distribution of mortality rates changes after each adjustment, or combination of adjustments, consistent with the major research question of this article.

Each of the three rate adjustments differs considerably from unadjusted mortality rates. Standardizing mortality provides a dramatic shift in spatial outcomes in the United States, as is shown in the following maps. Age, age-sex, and age-sex-race standardizations of the dependent variable mortality from 1993–97 have similar results at first glance, yet very important differences exist within each adjustment. Regarding the unadjusted rates (Figure 2), high mortality counties are dispersed throughout the middle section of the country and the extreme Southeastern corner (Florida), with many contiguous low mortality counties in the West and others scattered throughout the East. With no adjustment, the Southeastern tip of the U.S. (Florida) has high mortality rates because of the high numbers of elderly who retire in the area, which inflates the death rate because of an average population with a shorter life expectancy than other parts of the country. After adjusting for age (Figure 3), the high mortality cluster shifts away from the Southeastern tip of the U.S. and concentrates across the general Southeast region as a whole, whereas the low mortality concentrations are in



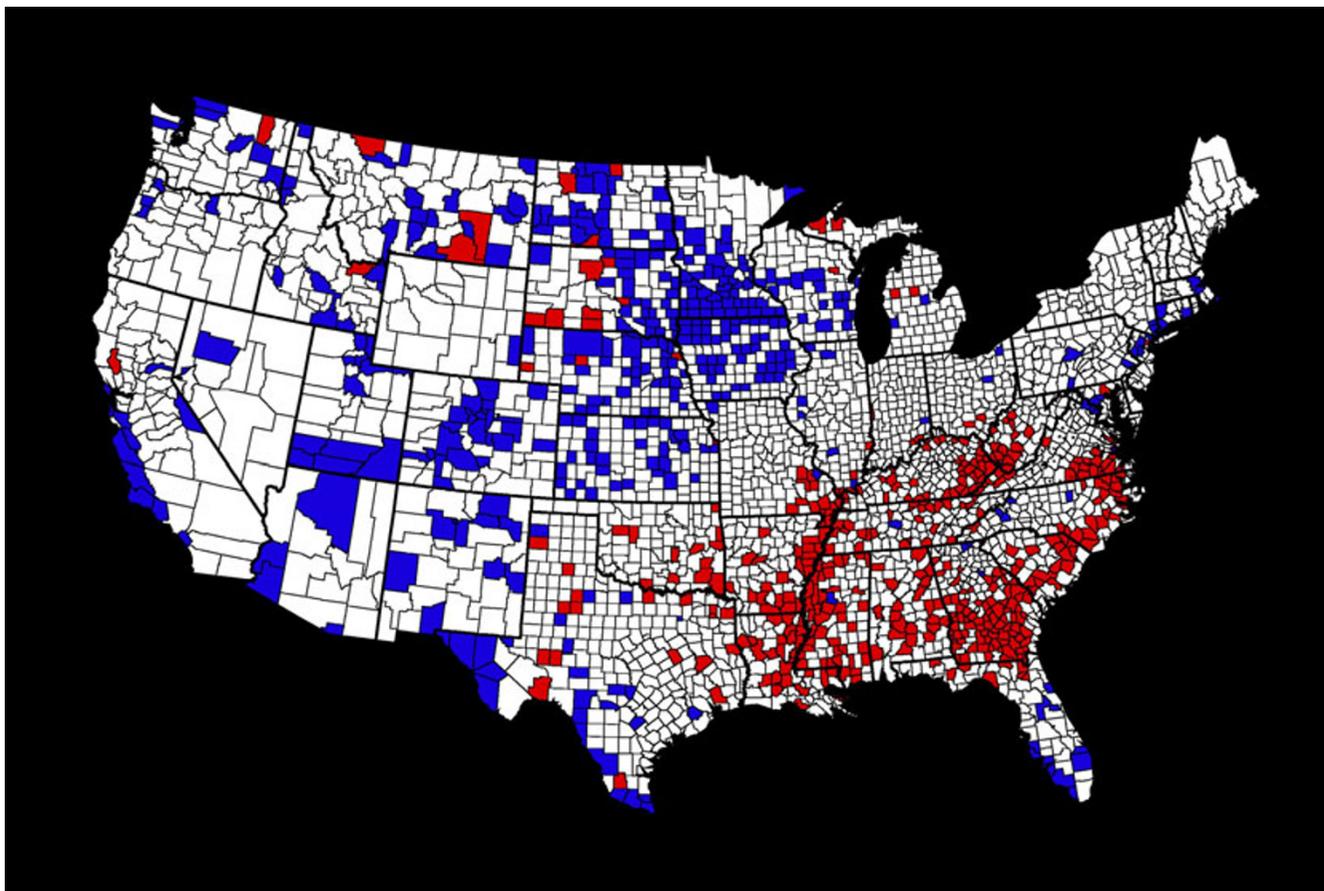
**Figure 2**  
Unadjusted Mortality Rates 1993–1997 Red = High Mortality White = Normal Mortality Blue = Low Mortality

the Midwest and Central Great Plains. This substantial change in spatial patterns of the data characterizes the importance in methodological change and its impact.

As we move from crude mortality rates through age-adjusted, age-sex adjusted (Figure 4), and age-sex-race adjusted rates (Figure 5), we see a subtle geographic shift. Specific to both age-sex and age-sex-race adjustments, low mortality counties move from a concentration in the West and a wide distribution in the East to a concentrated cluster in the Upper Great Plains of the Central United States. High mortality clusters are again located in the Southeast with a slight westward expansion, but are more sparsely concentrated in the age-sex-race adjustment (Figure 5). Here, high mortality rates across the Southeast have been reduced, based on adjustment of the high proportions of African-Americans, who have a higher risk of death than other races in the United States. Essentially, moving from crude rates to any form of standardized mortality rates (age, age-sex, age-sex-race) results in a dramatic change,

with less variation across each particular adjustment method. The value of these maps is that the researcher can look beyond the concentration of blacks, elderly, or any other demographic measure as the root of health disparities throughout the United States, or specific geographic area, and focus on other social or economic factors that may significantly influence poor health outcomes.

Examination of differing spatial results is available across this series of maps, with emphasis on figures 6 and 7, detailing the distinct shift that occurs between unadjusted mortality rates and age-sex adjusted mortality rates. This incredible change in spatial outcomes, as a result of simple standardization/adjustment procedures, supports the necessity of choosing the correct procedure with which to transform data, as well as the impacts that occur based on these transformations. The graph of mean change across each adjustment procedure (Figure 15) highlights variation within the data to accompany this series of maps. Change in the mean across methodological technique



**Figure 3**  
Age Adjusted Mortality Rates 1993–1997 Red = High Mortality White = Normal Mortality Blue = Low Mortality

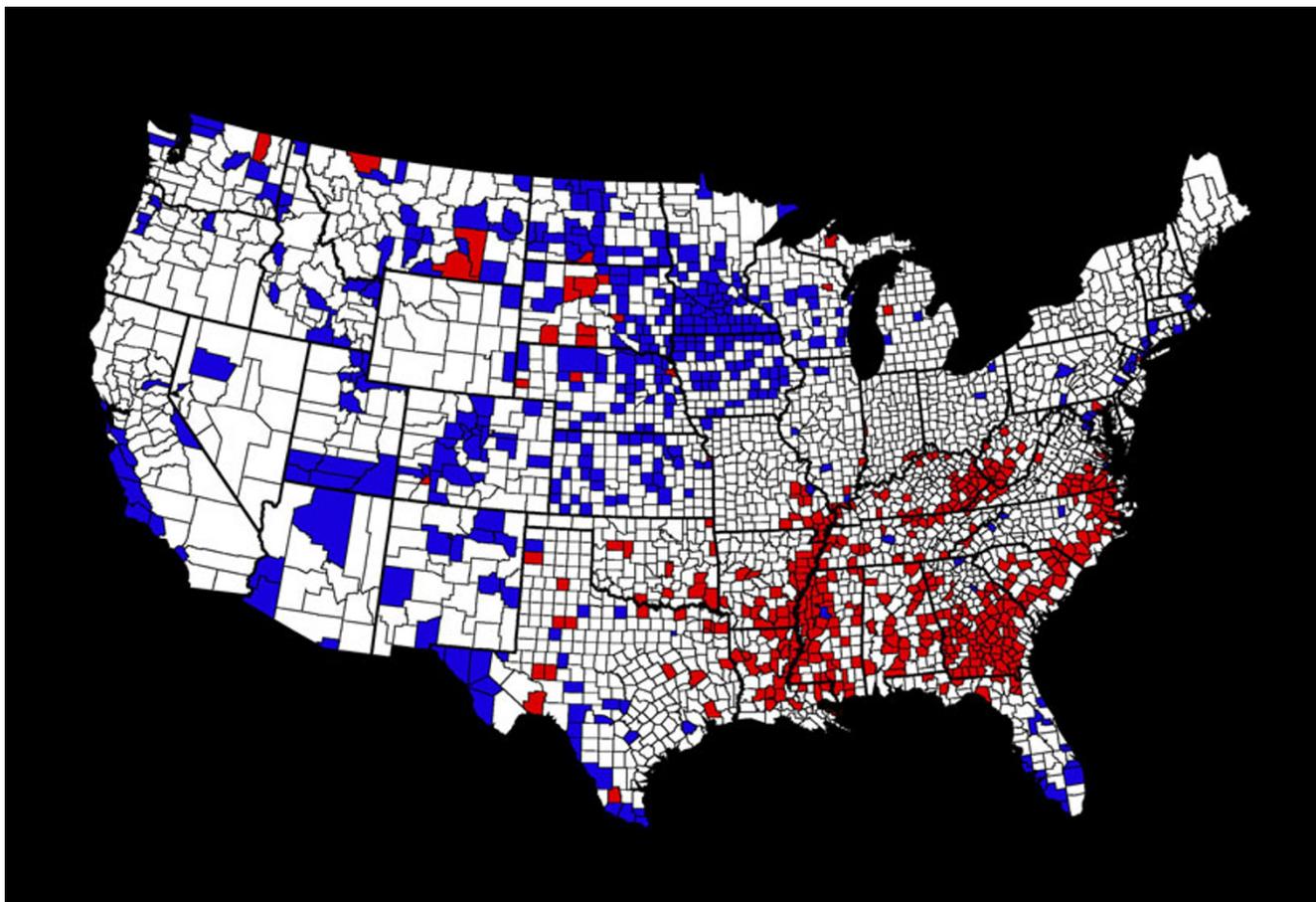
shows the mean for unadjusted 5-year averaged mortality rates at 1,034, while the age-adjusted rates average 934, the age-sex adjusted rates average 919, and the age-sex-race adjusted mortality has a mean of 937. Variation exists in the mean across these four rates, and is much smaller across the three adjustment techniques, with much larger variation between the crude rates and any of the three adjustments.

#### **Cut-points/operationalization**

Operationalizing the mortality variable is another necessary step in reaching the appropriate spatial outcome in a map. As is the case with standardization procedures, differences in these operationalizations lead to different outcomes, which are mapped in this section. Referring back to figure 1, 5-year average, age-adjusted mortality rates are used to assess change across cut-point procedures.

Once the dependent variable is standardized, an important aspect of mapping data that must be considered is the

definition of cut-points. Each series of cut-points is based on a particular mathematical and statistical formula, therefore leading to this variation in results. There are three standard cut-points used in these mortality maps: 1) standard deviation, 2) natural breaks, and 3) quantiles. Statistically speaking, standard deviation is the positive square root of the variance, measuring a designated area above and below the mean. "Natural breaks is based on an algorithm produced by Jenks that is an optimization procedure which minimizes within class variance and maximizes between class variance in an iterative series of calculations" [18]. In other words, it identifies natural cut-points in the data, rather than imposing classification boundaries with set widths. Quantiles are another commonly used classification method, simply placing an equal number of enumeration units into each class. For instance, in a five-class group, each class holds twenty percent (p. 670). This procedure is used in the Geographic Information System (GIS) software ArcView 3.2, which formulates the categorizations for data separation.



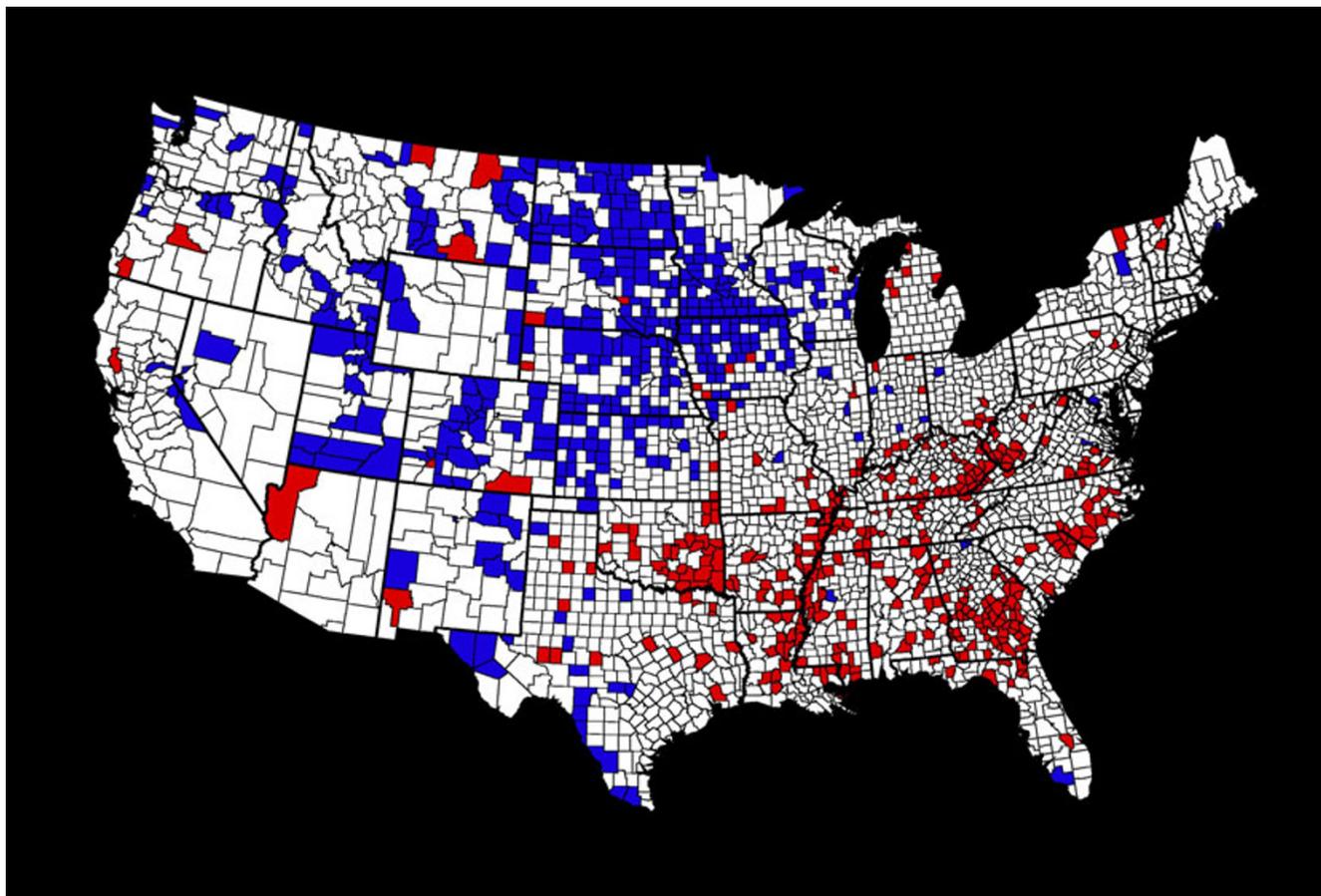
**Figure 4**  
Age/Sex Adjusted Mortality Rates 1993–1997 Red = High Mortality White = Normal Mortality Blue = Low Mortality

Examining the series of United States *age-adjusted* mortality maps for 1968–1997, these three cut-point techniques display different and interesting spatial results (Figures 8, 9, 10). In a very general sense, the same spatial outcomes occur in each of the three maps. Although the broad high and low mortality patterns or clusters are in the same regions of the United States across each map, their magnitudes or concentrations differ greatly. Each of the three maps displays high mortality in the Southeast region of the United States. Low mortality is concentrated in the Midwest and Plains States in the middle section of the country, as seen in figure 1, using standard deviation cut points.

Quantiles (Figure 8) show much larger clusters of high and low mortality and is the most concentrated of any classification method. Far fewer counties are considered on par with the national average than is the case with figure 1. Natural breaks (Figure 9), broken into three classes,

show the same general clusters, but are not quite as filled out as the quantiles. The Midwest is less concentrated, as well as parts of the high mortality clusters in the South. The standard deviation technique results in the most sparsely clustered maps (Figure 10). Figure 10 distinguishes a sparse number of counties in both high and low mortality clusters, with the Midwest being far less concentrated than in previous figures, once again the case with the Southern unhealthy clusters as well. While the patterns shown using the standard deviation method are consistent with quantiles and natural break methods, it is comprised of far fewer counties.

Analyzing spatial outcomes in mortality data across standard deviations, natural breaks, and quantiles provides interesting results. Although outcomes across each cut-point technique may not always vary dramatically, differences definitely exist. These differences are enough to distinguish among each methodology the reasons why the



**Figure 5**  
Age/Sex/Race Adjusted Mortality Rates 1993–1997 Red = High Mortality White = Normal Mortality Blue = Low Mortality

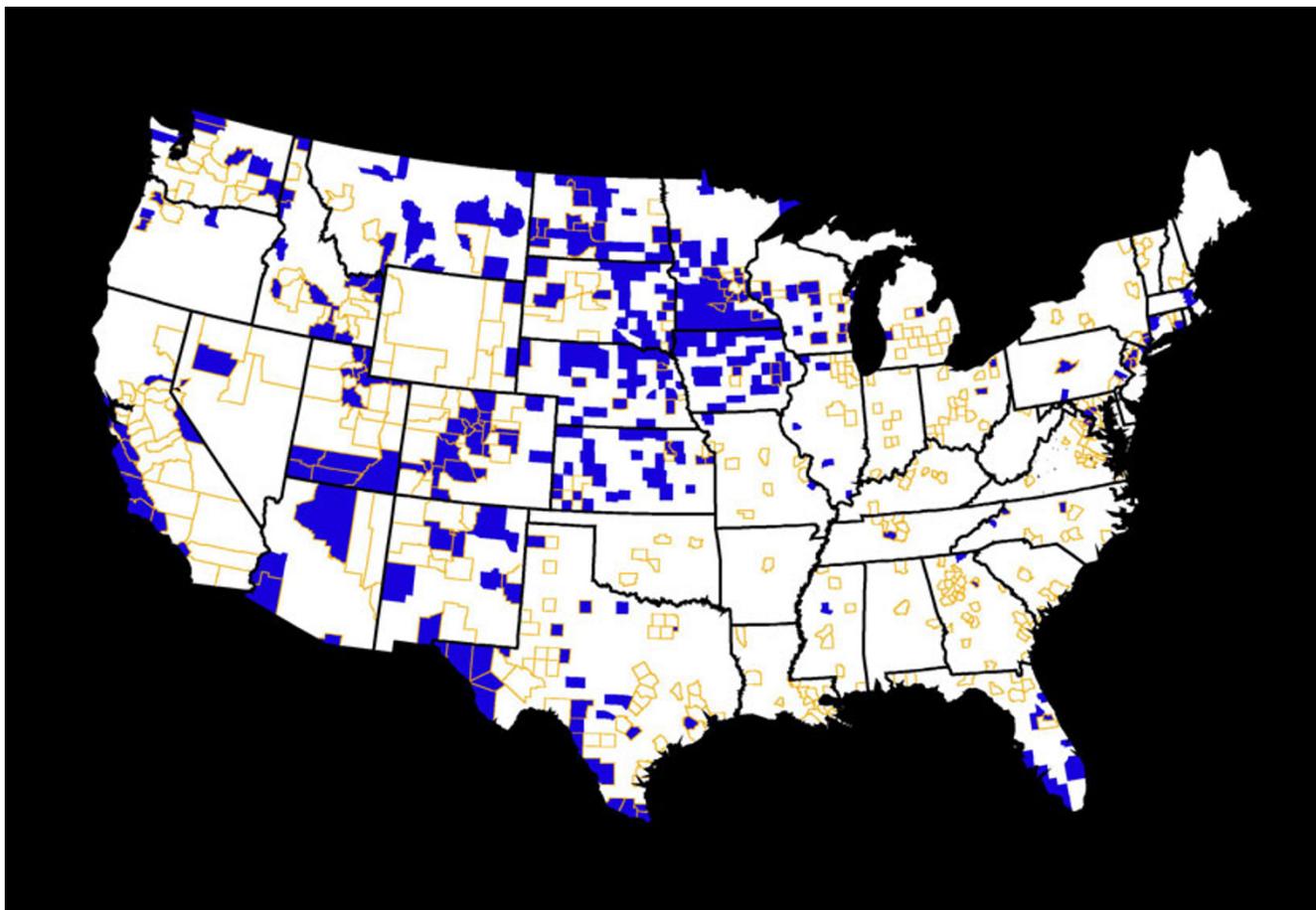
spatial patterns occur in a variety of ways based on how the dependent variable is measured within counties. Making comparisons across method, while focusing at the county level, greatly contributes to the researchers knowledge and understanding of the underlying data. While each of these techniques is unique, we chose to use standard deviation cut points because it is the most statistically sound estimate of the three. Using standard deviations, we test the spatial significance and temporal stability of our mortality clusters, but first we analyze classification size and color scheme of mortality.

#### **Class size and color scheme**

Class size brings a new dimension to data visualization in maps. Figure 9 emphasizes simplicity by presenting only three mortality classes (high, average, and low) using natural breaks. By using a greater number of classes for instance, an intermediate class located within the high and low classifications yields interesting results. Calculated

by dividing the rates within each class into two or more separate classes for many techniques, or with the use of a new algorithm in the case of natural breaks, intensity of mortality can be detected. Extremely high or low rates can be distinguished from moderate rates using this technique, and corresponding shades of color with each classification are used to differentiate among the classes. Figures 11 and 12 show the differences in mortality mapping when using five and seven classes, respectively, as opposed to just three.

MacEachren [19] acknowledges the relationship between simplicity versus complexity regarding class size in maps and recommends producing maps in the most simple form possible, as we have done in our research. This allows for an easily readable and understandable map. But, not all maps have this luxury; some are of a more complex nature showing a greater amount of information. The spatial patterns in the 5-class natural breaks map

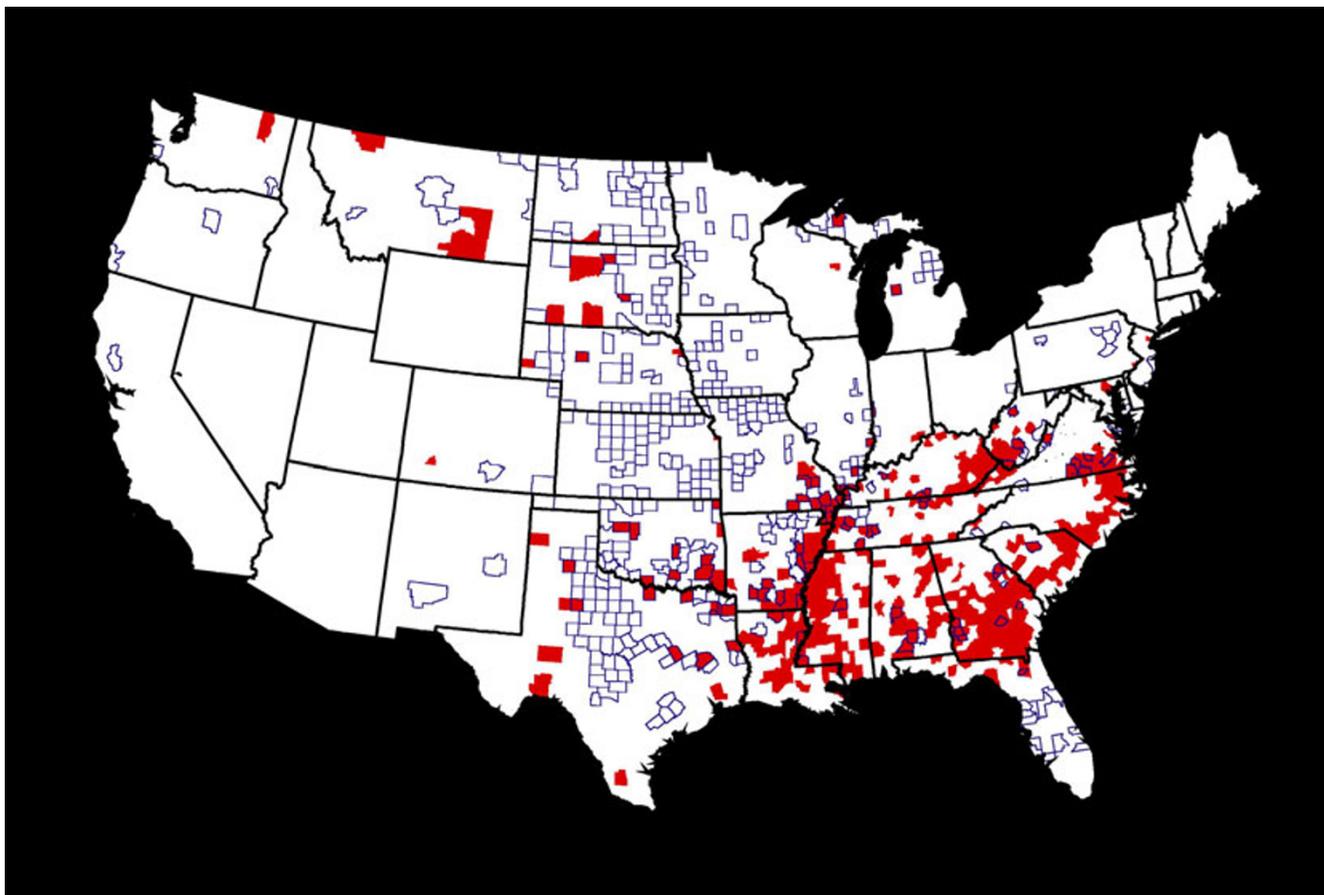


**Figure 6**  
Age/Sex Adjusted vs. Unadjusted Low Mortality Rates 1993–1997 Blue = Age/Sex Adjusted White = Normal and Low Mortality Orange Outline = Unadjusted Rates

(Figure 11) do not possess solid blankets of high and low mortality, as is the case with many previous maps. This is due simply to the additional colors that accompany new classes of mortality. Similar patterns of high and low mortality continues to exist, with the added dimension of intensity included in the map. The new algorithm with which these class calculations are produced causes the size of each class to change, as well as the counties that each encompasses. With the same general high and low mortality patterns as most figures presented in this article, different shades of red are dispersed throughout the Southeast, while the Midwest displays the majority of each low mortality intensity level. Average mortality counties are less prevalent throughout the country, due to the redistribution of counties into new classifications. This map has added displays of intensity in both red and blue clusters of counties. White counties are completely absent here, again due to the recalculation of high and low mortality,

hence, redistribution of these counties. The major difference between the natural breaks maps (Figures 9, 11, and 12) is the number of counties classified in the high and low groups, where figure 12 distinguishes higher concentrations in both high and low mortality than the previous two figures. The embedded clusters of high mortality in the Southeast and low mortality in the Midwest are still apparent, but new clusters of lower intensity appear for both categories throughout the United States.

As briefly mentioned in the previous paragraph, another cartographic issue that fundamentally relates to the visual appearance of intensity is color scheme. The use of shades across a particular color (red, blue) is effective in showing the hierarchy within one category (high or low mortality), when class size is greater than three. Using a high intensity color is effective in portraying unhealthy areas, whereas a softer color may be used to indicate healthy areas. When

**Figure 7**

Age/Sex Adjusted vs. Unadjusted High Mortality Rates 1993–1997 Red = Age/Sex Adjusted White = Normal and Low Mortality Purple Outline = Unadjusted Rates

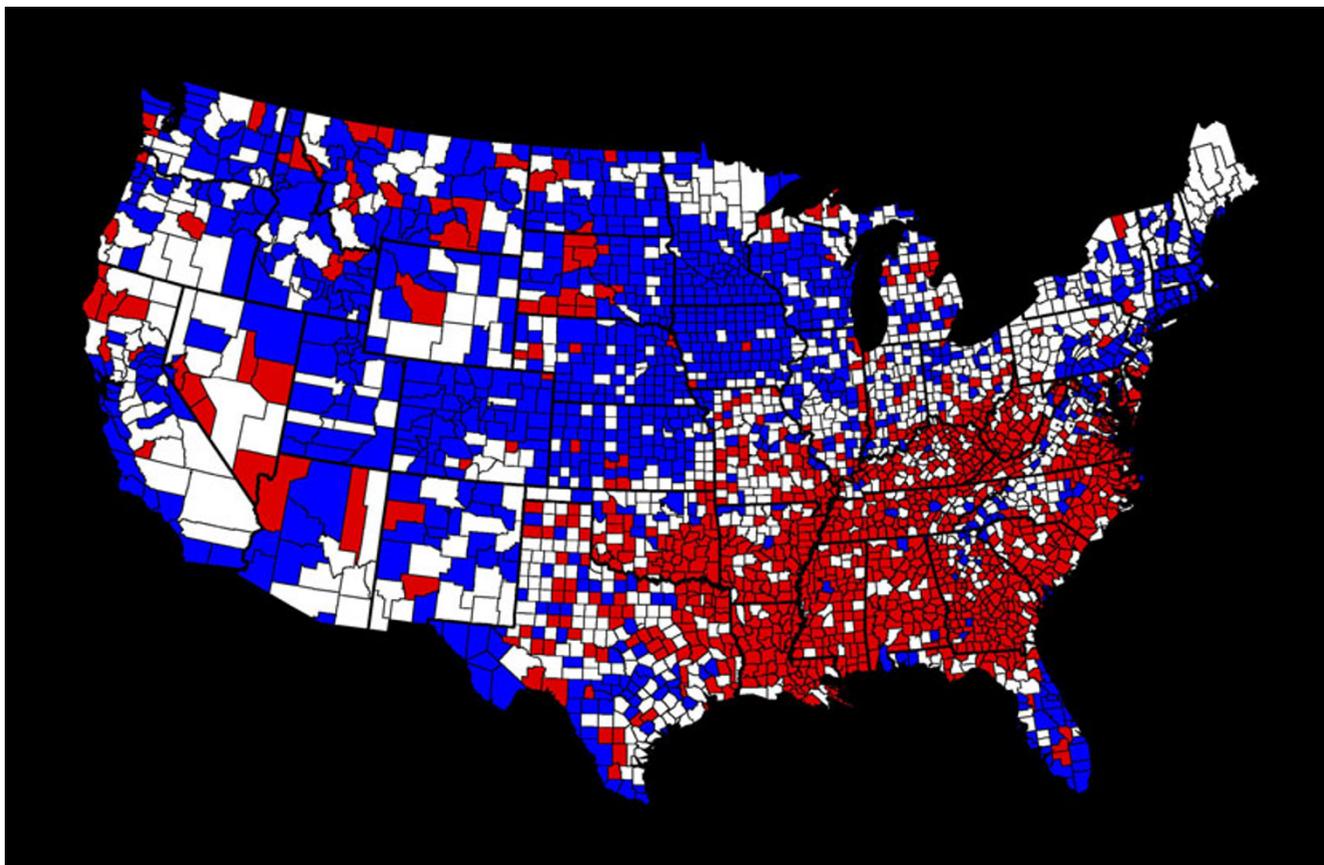
constructing a simple map, typically three classifications (Figure 1), the use of red and blue easily distinguishes high and low mortality or healthy and unhealthy clusters, the focus of this map. The designation of white for average mortality counties visually influences the reader away from these counties and allows the high and low counties to stand out more clearly. It is important to remember what medium is being used to display a choropleth map when experimenting with color schemes. Some are more appropriate to use for maps in print form, others may be fitting in web-based maps, or those shown electronically in presentations. These factors contribute heavily to variable colors, as well as background colors of maps.

#### **Statistical significance of spatial pattern**

Thus far, it is clear that standardizing and operationalizing the mortality variable is critically important when determining healthy and unhealthy county clusters. Now that a determination has been made using the particular trans-

formations of data in this research, testing the statistical significance of these clusters is the next logical step in the process of validating our findings. This spatial statistic test is done via the Local Moran's I. Using SpaceStat, a spatial statistics program operating as an extension to ArcView 3.2, we plotted a Local Moran's I scatter plot. The Local Moran's I is calculated by taking the standardized rate in County A and comparing it to the rates in adjacent Counties B1 through B4 (Figure 16). Assuming a normal distribution of surrounding rates, researchers may quantify whether statistically significant spatial clustering is present. This technique illustrates first order adjacency or in other words, its focus is on contiguous counties.

The basis of clustering in the Local Moran's I is dependent on the type of counties surrounding a target county. For example, a high mortality county surrounded by other high mortality counties is classified as "high-high", indicated by the color red. Blue counties refer to relatively low



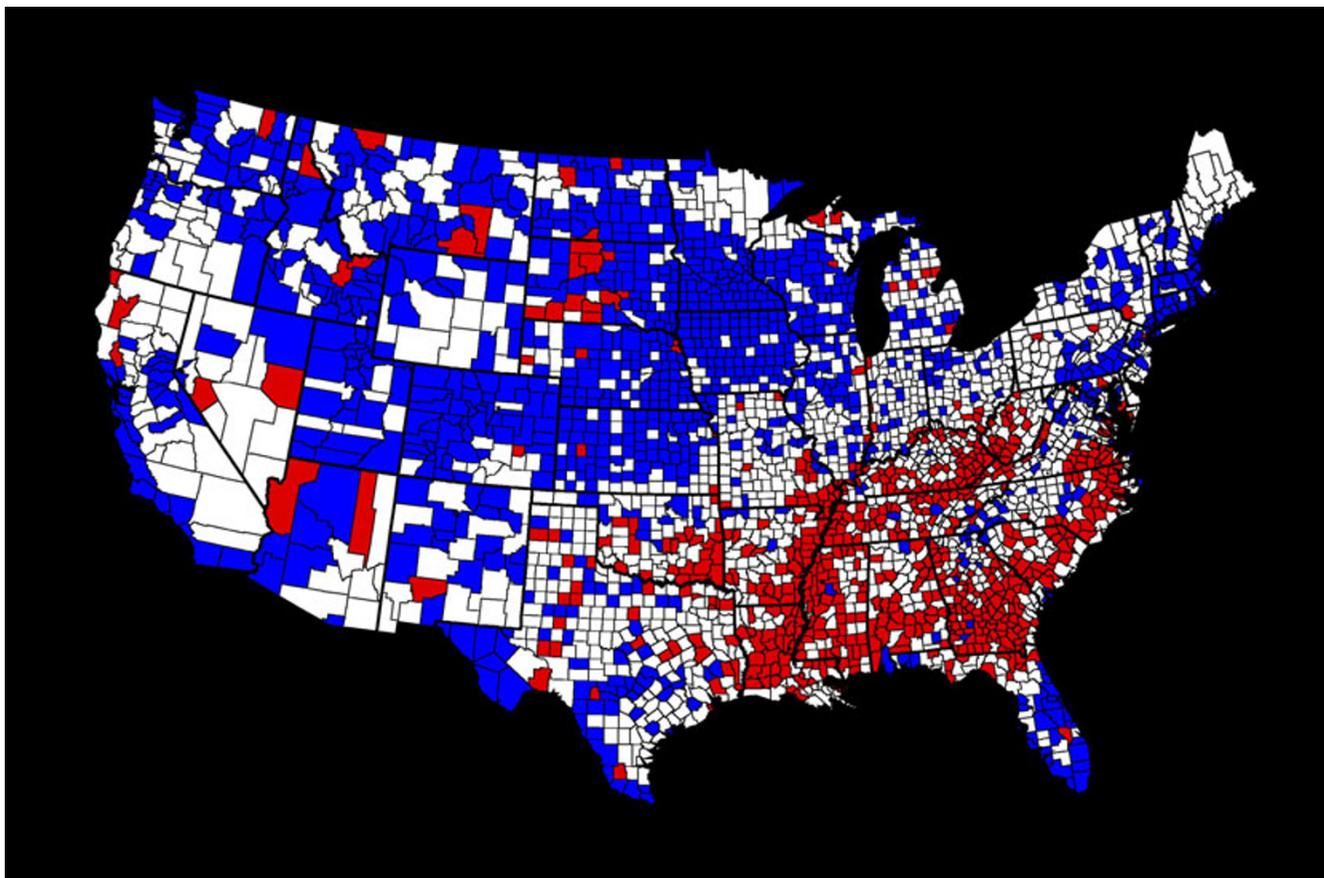
**Figure 8**  
United States Mortality: 3 Class Quantiles Red = High Mortality White = Normal Mortality Blue = Low Mortality

mortality counties, which are surrounded on each side by other relatively low mortality counties, classified as "low-low". These two classifications, red and blue counties, indicate spatial autocorrelation. Pink counties are high mortality counties adjacent to low mortality counties, and light blue counties are low mortality adjacent to high mortality counties. Counties whose mortality rate is statistically independent of its surrounding counties rates are colored white.

Analyzing this map (Figure 13), we see a statistically significant cluster of high mortality along the southern half of the East Coast. Other large belts of high mortality are the Mississippi Delta region and the region commonly known as Appalachia. Finally, a five-county cluster out West is the fourth significant high mortality cluster. Regarding significant healthy clusters, a broad area of the Midwest is the largest low mortality area. The low mortality counties bordering Mexico are believed to be an artifact of the dataset, where deaths may be underrepresented due to the exclusion in population and death data of non-res-

idents. Overall, these clusters are similar to many clusters demonstrated throughout this article, with slight differences throughout. One major difference of the Local Moran clusters is their statistical independence from other counties. Of the previous clusters derived from a variety of other techniques, the groupings were of a more general nature. Many of those clusters had included within them a mixture of county designations, for instance a few low mortality counties inside of a high mortality region. Figure 13 demonstrates a more pure representation and definition of "clusters".

In order to determine significant clusters across space, data methodology must be precise and logical in the preliminary stages. Figure 13 demonstrates that making the appropriate assumptions throughout the research process, as outlined in the previous sections of this article, lead to valid and reliable results. Based on the statistical significance of these clusters, the next step in our research was to test the temporal stability of mortality clustering.



**Figure 9**  
 United States Mortality: 3 Class Natural Breaks Red = High Mortality White = Normal Mortality Blue = Low Mortality

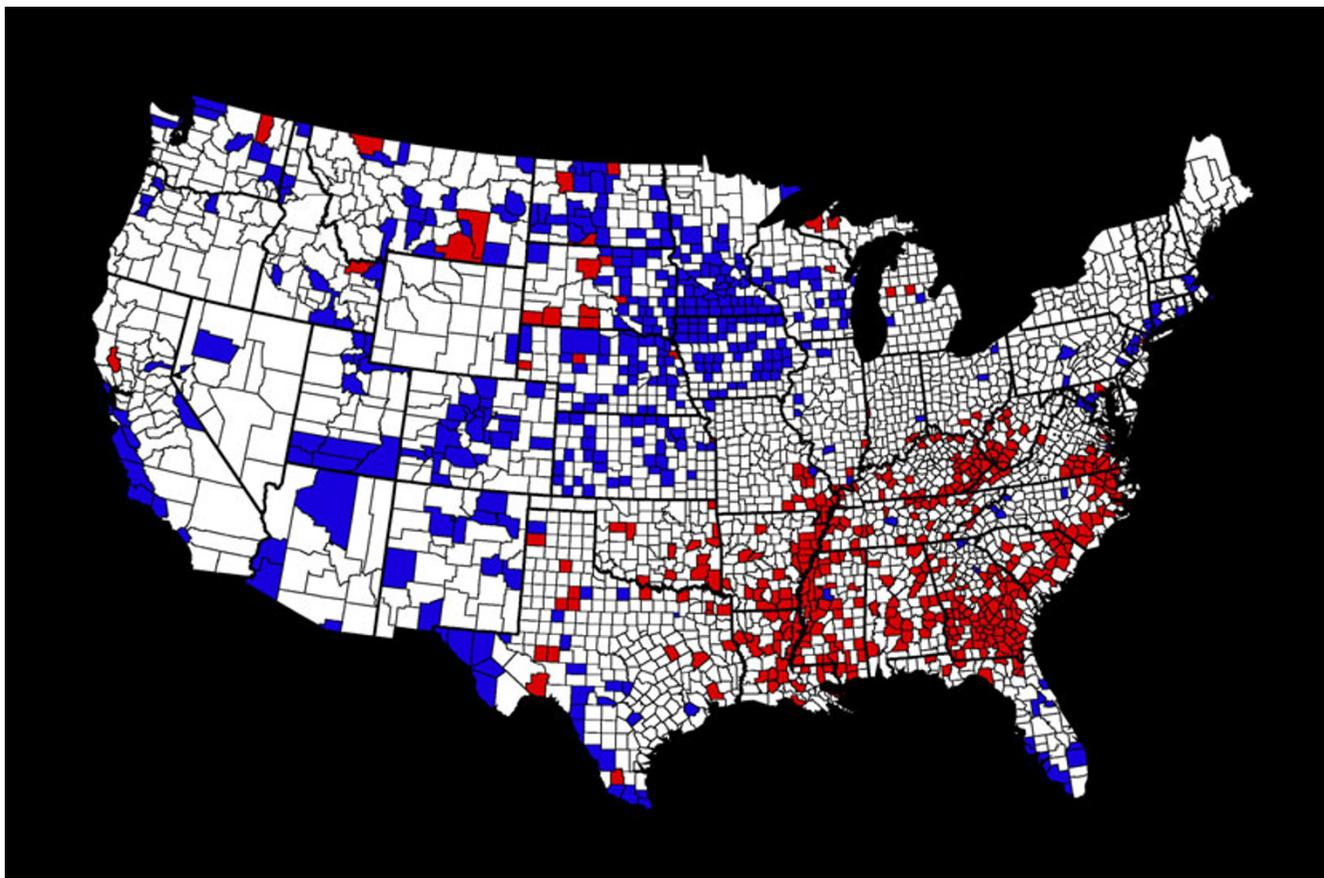
Measuring high and low mortality over time, 30-years to be precise, is valuable in finding trends and patterns that occur within the data.

**Persistence data**

After establishing our method of mortality rate standardization and operationalization, along with the presence of statistically significant healthy and unhealthy county clusters and how to portray them in a choropleth map, another investigation into the data was in order. Temporal stability, i.e. persistence of mortality over time was calculated [20]. For purposes of our research [20], we used age-adjusted death rates and standard deviation cut-points to measure high and low mortality and divided the thirty-year data into six 5-year time periods: 1968–72, 1973–77, 1978–82, 1983–87, 1988–92, and 1993–97. A significant advantage of using five-year time periods is to provide rate stability for small counties. A five-year average also eliminates potential outliers in deaths per year in a specific county. Any county one standard deviation above the

national mean in at least three of the six time periods is classified as persistently high mortality, whereas any county one standard deviation below the national mean in at least three of the six time periods is classified as persistently low mortality. Many counties change classification (high, average, low) over time; very few stay the same throughout. Because there exists a small number of counties designated as either high or low when the criteria is based on consistency throughout all six-time periods (6/6), a minimum of three time periods was assigned as the cutoff to measure persistence (3/6). Essentially, if a county is one standard deviation above or below the mean for at least 15 of the 30 possible years, it is classified as persistent.

Figure 14 demonstrates clusters of persistently high and low mortality in the contiguous United States. We have targeted high mortality clusters in similar regions to the previous maps. High mortality is concentrated along the East Coast, Appalachia, the Mississippi Delta, a handful of



**Figure 10**  
United States Mortality: 3 Class Standard Deviations Red = High Mortality White = Normal Mortality Blue = Low Mortality

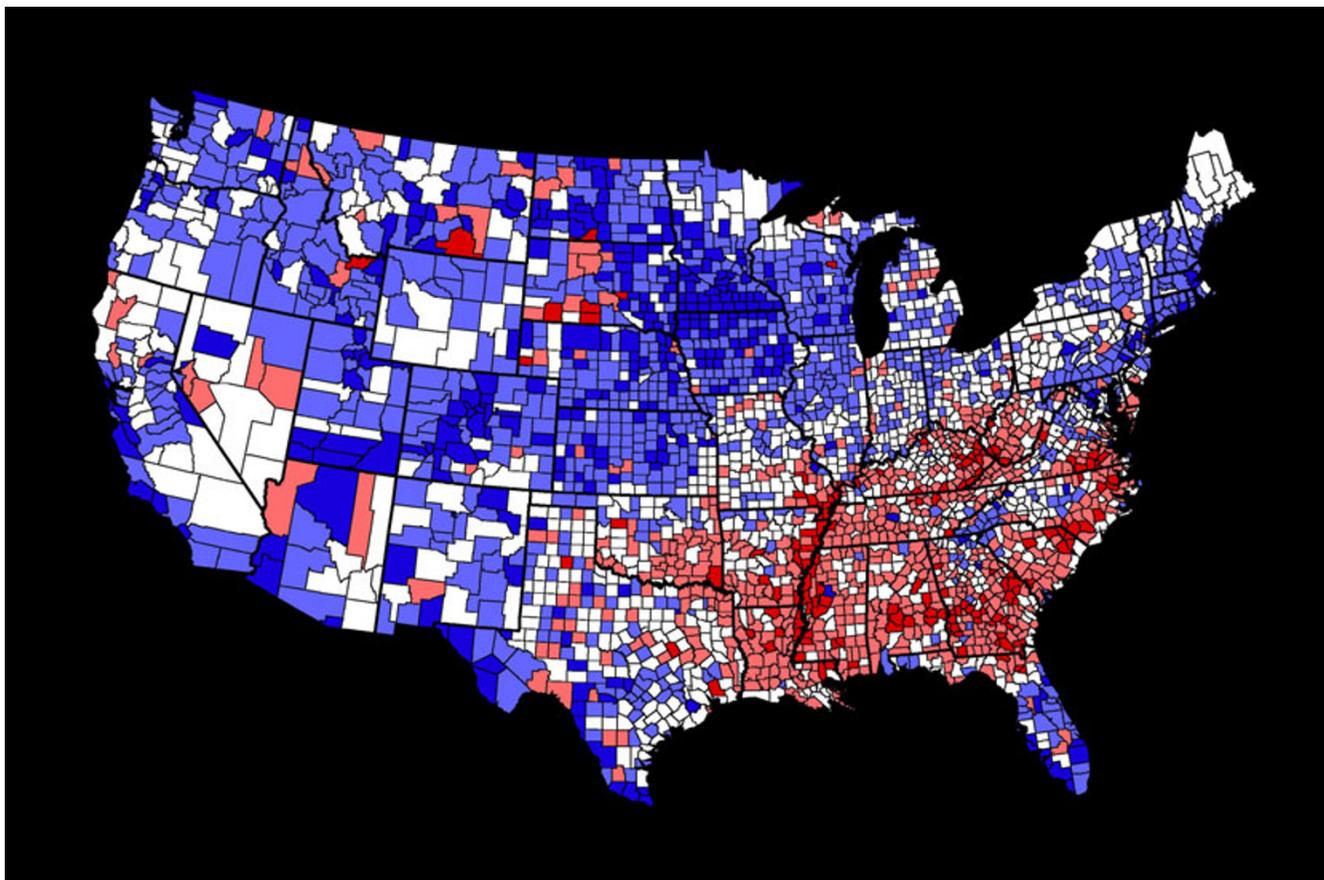
contiguous Southwestern counties, and a small cluster of counties in the North along the Canadian border. Once again, low mortality clusters include a large portion of the Midwest with another cluster encompassing a different area in the Southwest.

The purpose of this map was to identify persistent clusters of high and low mortality over time. The majority of these high and low mortality counties are clustered consistently with those of the Local Moran's I statistically significant clusters, leading to a reasonable conclusion that healthy and unhealthy places are deeply embedded in these particular health outcomes, thereby answering the third research question of our empirical research project. By implementing the mapping procedures discussed throughout this article, we targeted appropriate areas to which social science and demographic research can gain insights into similarities and differences within the social structure of these places, and what characteristics may be harmful or beneficial to the people who reside in these

areas. Cluster identification is relevant to this particular article because data construction methodology, standardization, and operationalization have a strong influence over which clusters arise; therefore they must be appropriately defined and targeted. Each methodology outlined in this article has led to the identification of significant and temporally embedded clusters of high and low mortality. Defining these clusters with a high level of confidence, validity, and reliability is a process that takes multiple steps, each of which must be approached with caution.

### Conclusions

This article has demonstrated the importance of data transformation and visual display to spatial mortality outcomes through our line of research in healthy and unhealthy places in America. Definition, operationalization, calculation and standardization of the variable being mapped (mortality) are crucial in providing valid and reliable spatial and statistical outcomes in research. Through a descriptive and visual analysis of changes across each of

**Figure 11**

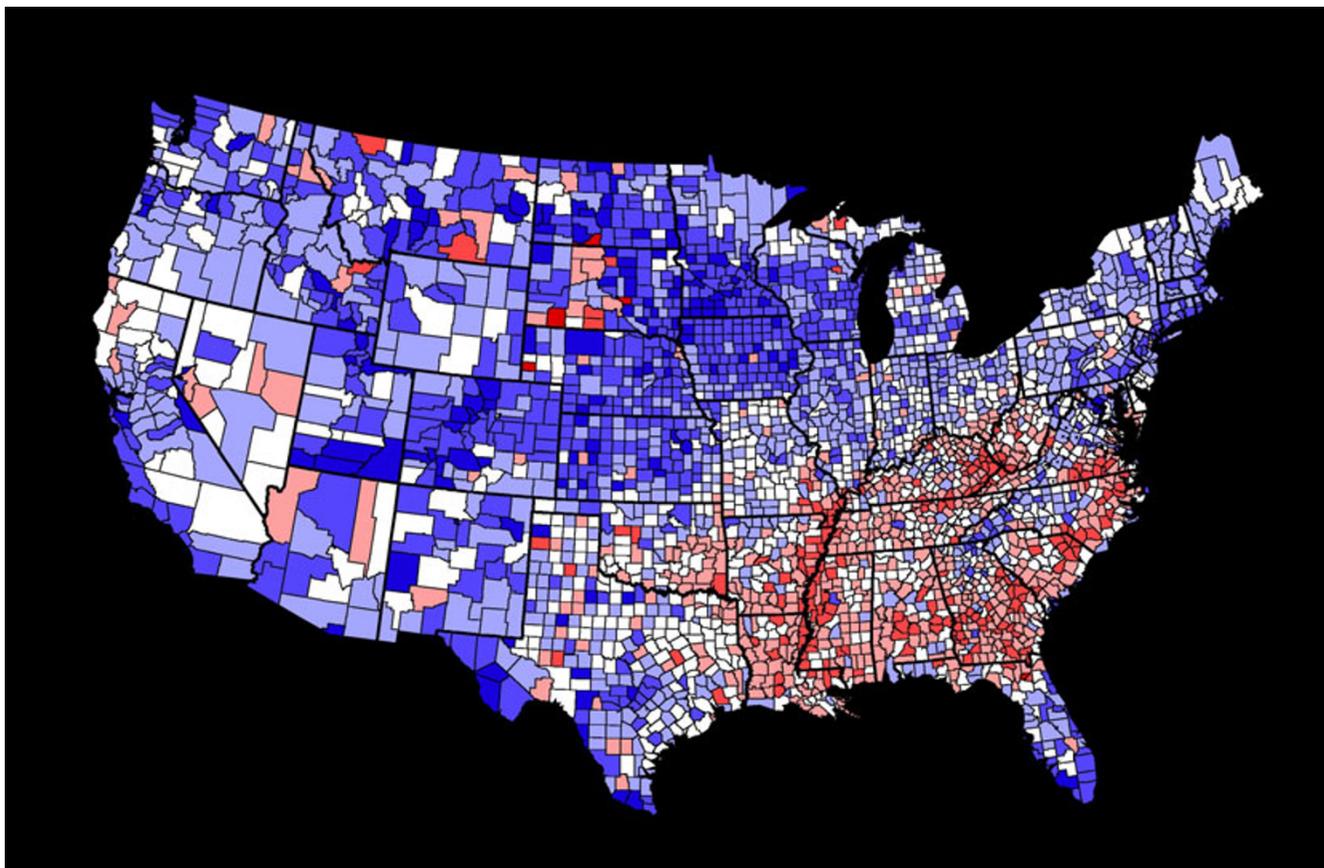
United States Mortality: 5 Class Natural Breaks Red = Highest Mortality Light Red = High Mortality White = Normal Mortality Light Blue = Low Mortality Blue = Lowest Mortality

these techniques, the differences that may occur in the spatial distribution of the data become apparent. Importance in standardization and calculation of the dependent variable is outlined, with emphasis on the appropriate methods of detecting trends and changes in mortality rates over time. Using a series of mortality maps demonstrates the stark spatial outcomes that exist between unadjusted and age-adjusted, age-sex adjusted, and age-sex-race adjusted mortality rates. Another necessary point of investigation is cut-points, or the manner in which the variable is operationalized. Quantiles, natural breaks, and standard deviations were summarized, along with the spatial implications provided by each technique and the differences that exist among these classifications. From here, significant mortality clustering was identified using the Local Moran's I, as well as mortality persistence and temporal trends in these clusters over a period of 30 years. Finally, summary maps were provided where necessary to highlight any dramatic changes in the spatial outcomes and patterns across varying methodological techniques.

The maps and graphics used to emphasize the descriptive and theoretical information presented in this article provided fundamental support for the impacts that these techniques have upon data analysis. Without the proper methodology, research results and conclusions may be critically flawed, resulting in an inappropriate investigation into potential policy-making and intervention to at-risk populations, communities, and counties. We hope these illustrations are useful for fellow investigators as they begin to fully employ data visualization and mapping techniques.

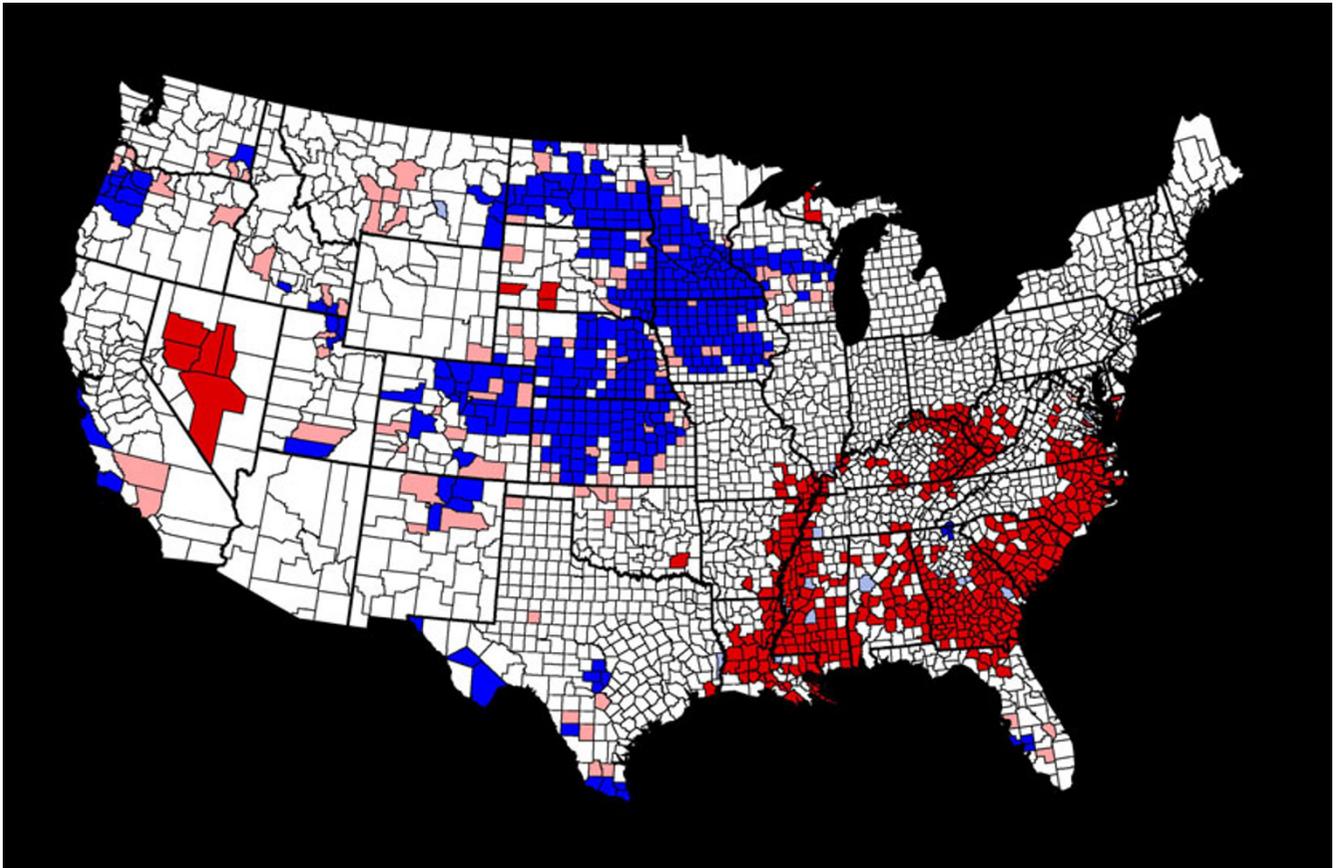
#### **Authors' contributions**

W.J. manipulated the mortality data, created the majority of the graphics, and drafted the manuscript. R.C. and J.C. provided valuable edits, organization, and structure to the manuscript. C.C. calculated the original mortality data, and T.B. constructed the Local Moran's I map of persistent mortality.

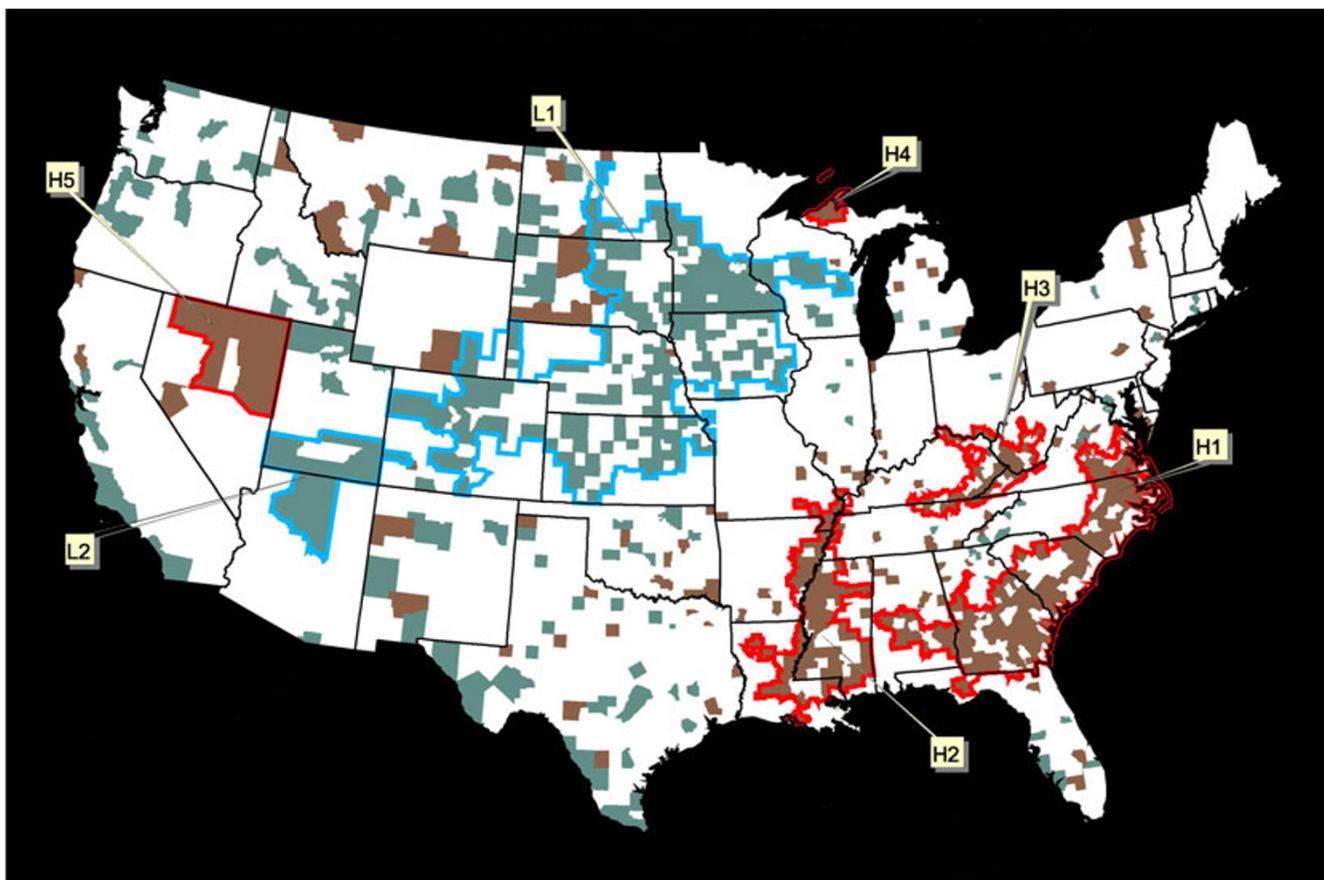


**Figure 12**

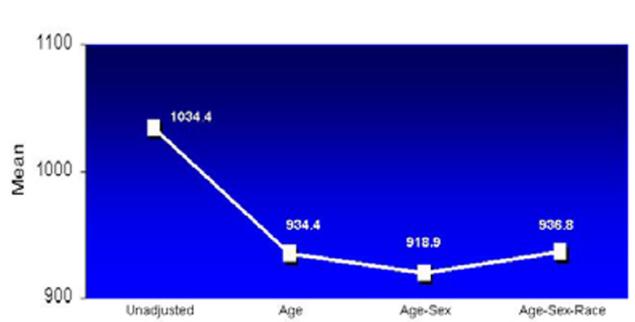
United States Mortality: 7 Class Natural Breaks Red = Highest Mortality Medium Red = Mid/High Mortality Lightest Red = High Mortality White = Normal Mortality Lightest Blue = Low Mortality Medium Blue = Mid/Low Mortality Blue = Lowest Mortality



**Figure 13**  
Local Moran's I Scatterplot of Persistent Mortality 1993–1997 Red = High Cnty-High Cnty Light Red = High Cnty-Low Cnty  
White = No Significance Light Blue = Low Cnty-High Cnty Blue = Low Cnty-Low Cnty



**Figure 14**  
 Clusters of Persistent Mortality 1968–1997 Dark Brown = High Mortality (5-6/6) Medium Brown = High Mortality (4/6) Light Brown = High Mortality (3/6) White = Normal Mortality Light Green = Low Mortality (3/6) Medium Green = Low Mortality (4/6) Dark Green = Low Mortality (5-6/6) Red Outline = Unhealthy Clusters Blue Outline = Healthy Clusters Yellow Box with H inside = High Mort Clusters Yellow Box with L inside = Low Mort Clusters



**Figure 15**  
 Change in Mean Across Standardization Model



**Figure 16**  
 Local Moran's I

## Acknowledgements

This research was made possible by grant number 4 DIA RH 00005-01-01 from the Office of Rural Health Policy of the U.S. Department of Health and Human Services through the Rural Health, Safety and Security Institute, Social Science Research Center, Mississippi State University. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Office of Rural Health Policy. Finally, we wish to thank the anonymous reviewers for their helpful comments.

## References

1. Harvey D: **Pattern, Process, and the Scale Problem in Geographical Research.** *Transactions of the Institute of British Geographers* 1968, **0(45):**71-78.
2. Gatrell A: **Structures of Geographical and Social Space and Their Consequences for Human Health.** *Geografiska Annaler* 1997, **79B(3):**141-154.
3. National Center for Health Statistics: **Compressed Mortality File: 1968-88.** Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, Hyattsville, Maryland 2000.
4. Cossman RE, Blanchard T, James W, Jackson-Belli R, Cosby A: **Healthy and Unhealthy Places in America: Are These Really Spatial Clusters?** *Proceedings of the 22nd Annual ESRI International User Conference 2002* [<http://gis.esri.com/library/userconf/proc02/abstracts/a1064.htm>].
5. Bond Huie SA: **The Concept of Neighborhood in Health and Mortality Research.** *Sociological Spectrum* 2001, **21:**341-358.
6. Ricketts TC, Pope DL: **Demography and health care in eastern north carolina.** *N C Med J* 2001, **62(Suppl):**S20-S25.
7. Wennberg JE: **Understanding Geographic Variations in Health Care Delivery.** *The New England Journal of Medicine* 1999, **340(1):**52-53.
8. Devesa SS, Grauman DJ, Blot WJ, Pennello GA, Hoover RN, Fraumeni JF: **Atlas of Cancer Mortality in the United States: 1950-94.** Hyattsville, MD: National Cancer Institute 1999.
9. Pickle LW, Mungiole M, Jones GK, White AA: **Atlas of United States Mortality.** Hyattsville, MD: National Center for Health Statistics 1996.
10. Pickle LW, Mason TJ, Howard N, Hoover R, Fraumeni JF: **Atlas of U.S. Cancer Mortality Among Whites: 1950-1980.** National Cancer Institute, Bethesda, MD 1987.
11. National Center for Chronic Disease Prevention and Health Promotion: **Men and Heart Disease. Women and Heart Disease.** 2001 [<http://www.cdc.gov/nccdphp/cvd/womensatlas/statemaps.htm>].
12. Dartmouth Medical School Center for the Evaluative Clinical Sciences: **The Dartmouth Atlas of Health Care.** 2000 [<http://www.dartmouthatlas.org/>].
13. National Cancer Institute: **Cancer Mortality Maps & Graphs.** 2000 [<http://www3.cancer.gov/atlasplus/>].
14. Murray CJL, Michaud CM, McKenna M, Marks J: **U.S. County Patterns of Mortality by Race: 1965-1994.** *Harvard Center for Population and Development Studies* 1998 [<http://www.hsph.harvard.edu/organizations/bdu/usbodi/>].
15. Snow J: **On the Mode of Communication of Cholera.** New York, NY: *The Commonwealth Fund* 21855.
16. Jenks GF, Caspall FC: **Error on Choroplethic Maps: Definition, Measurement, Reduction.** *Annals of the Association of American Geographers* 1971, **61(2):**217-244.
17. Monmonier M: **Mapping It Out: Expository Cartography for the Humanities and Social Sciences.** University of Chicago Press; 1993.
18. Brewer CA, Pickle L: **Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in a Series.** *Annals of the Association of American Geographers* 2002, **92(4):**662-681.
19. MacEachren AM: **The Role of Complexity and Symbolization Method in Thematic Map Effectiveness.** *Annals of the Association of American Geographers* 1982, **72(4):**495-513.
20. Cossman RE, Cossman JS, Reavis RM, Cosby A: **Forthcoming. Reconsidering the Rural-Urban Continuum in Rural Health Research.** *The Journal of Rural Health, Forthcoming* 2003.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

