Integrative Literature Review: A Review of Literature Related to Geographical Information Systems, Healthcare Access, and Health Outcomes

by Barbara Ann Graves, PhD, RN

Abstract

Differences in access to healthcare services and the resulting adverse health outcomes are major public health priorities. The Institute of Medicine and the Department of Health and Human Services have identified the need for strategies to improve access to healthcare services and to support the improvement of health outcomes. The literature documents health disparities associated with healthcare access and health outcomes from a geographic perspective. Place of residence, location of healthcare services, and geography in general are important factors in the analysis of health.

Geographical information systems (GISs) are an emerging technology in the analysis of health from a geographical or location context. As a type of information technology, GISs are potentially powerful assessment tools for the investigation of healthcare access, health outcomes, and the possible resulting health disparities. Their ability to integrate health data with mapping functions allows for visualization, exploration, and modeling of health patterns. Application of GIS technology using health data can help in describing and explaining disparities in healthcare access and health outcomes.

The studies reviewed demonstrated the use of GISs to investigate various aspects of healthcare access and health outcomes, including environmental variables of Lyme disease, sociodemographic variables and teen pregnancy, geographical disparities in breast cancer mortality by racial groups, PCP and AIDS prevalence, and factors of a leptospirosis disease outbreak. The literature reviewed shows effective integration and analysis of health data using GIS technology.

Key Words: Geographical information systems (GIS), healthcare access, health outcomes

Introduction

Differences in access to healthcare services and the resulting adverse health outcomes are major public health priorities. The Institute of Medicine and the Department of Health and Human Services have identified the need for strategies to improve access to healthcare services and to support the improvement of health outcomes.¹⁻³

The purpose of this article is to provide a review of the healthcare literature related to the use of geographical information systems (GISs) in the assessment of access to healthcare services and health outcomes. Because geography can be important to health, GIS technology can help in the evaluation of complex health relationships. The spatial analysis of health and disease and their relationship to the
geographic environment can provide valuable insight into patterns of health and health outcomes. Analysis of location of healthcare services and the relationship to health outcomes is important to the development of health-related policies and regulations for the prevention of health disparities.

The major topics for the review are GISs, healthcare access, and health outcomes. With this purpose in mind, the research questions that served as a basis of this literature review are as follows:

1. How have GISs been used in the analysis of healthcare access and health outcomes?
2. Are GISs effective in the assessment of healthcare access and health outcomes?
3. What are the implications of GIS analysis for healthcare professionals?

Methodology

The methodology for this study was a review of research literature in which GISs were used to evaluate healthcare access or health outcomes. A database search (CINAHL, Medline, EBSCOhost, and ProQuest) was conducted using the keywords “GIS AND healthcare access” and “GIS AND health outcomes” to search for research articles published between January 1995 and December 2004. Only articles with the paired keywords were extracted for full review. This strategy resulted in nine final articles for review and inclusion in this integrative literature review.

GISs and Health

A fundamental premise of health geography is that illness and health are unequally distributed across space and time. Spatial patterns of illness have been associated with many factors, including climate, microbes, exposures, culture, race/ethnicity, geography, and distribution of healthcare services. Understanding spatial relationships of health and illness is important.

A GIS is a “computer-based system for integrating and analyzing spatially referenced data.” GISs have the ability to capture, store, retrieve, analyze, and display spatial data. GISs can be used to display the spatial distribution of health data. Mapping of health data can be instrumental in visualizing patterns and generating questions that may have not otherwise occurred to researchers.

Historically, GISs have been used in the management of land and natural resources. More recently, GISs have emerged as a new technology in healthcare. They are now widely used in many areas of health and health research. In particular, they provide analytical tools for health geography and epidemiological research when geography is important. As an analytical tool, they serve to advance the knowledge base of health informatics. The benefits of GIS technology are becoming evident in the areas of public health and healthcare delivery. In the recent decade, GISs have been used to accurately and concisely provide computer-assisted cartography of disease patterns and outbreaks as well as racial, ethnic, and geographical disparities in health outcomes. Atlases produced by the United States federal health agencies concisely display maps of the spatial distributions of health outcomes. Within a framework of health information management, GISs are able to answer health questions regarding “where is what.”

GISs and Health Outcomes

GISs and Lyme Disease

Cromley et al. used an ecological study design to investigate the environmental variables that determine the distribution of Lyme disease in a 12-town region of south-central Connecticut from 1993 to 1995. The authors mapped the residential location of human cases of Lyme disease, a tick-borne disease, and evaluated the relative risk by type of residential setting. Data were obtained from the Connecticut Department of Public Health, and the cases of Lyme disease were decoded by residential street address using an address-matching procedure in the GIS software package ARC/INFO. Only 424 of 503 addresses were successfully geocoded due to incomplete addresses or the use of post office boxes. The missing data could introduce selection bias into the study.
This study successfully presented the relative risk (RR) or probability of Lyme disease occurring. A thematic map of Lyme disease case location was overlaid on a map of ecoregions (the South Central Hardwoods Region and the Eastern Coastal Region) of Connecticut to identify the number of cases per region. Population estimates of the ecoregions and villages were interpolated from 1990 Census block group data. This mapping showed that most of the cases of Lyme disease occurred in the Eastern Coastal Region (RR = 2.4; p < .01). Results also showed a significant RR associated with living outside a village (RR = 10.4; p < .01). GIS methodology was useful in the analysis of environmental conditions, the related incidence of disease, and the RR. This expression of the probability of Lyme disease occurring provided an advantage over previous tick counts from killed deer in identifying residential areas at high relative risk for Lyme disease and in the planning of new residential development.

**GISs and Unmarried Teen Births**

Blake and Bentov used a retrospective exploratory study design to analyze and map unmarried teen births and their associated sociodemographic variables. Data for Dallas County, Texas, were extracted from Texas Department of Health data and included date of infant birth, mother’s age at infant’s birth, marital status, zip code of residence, race, and ethnicity for all births to adolescents (<20 years of age). The sample size was 8,838 after the exclusion of those with unknown marital status or inaccurate zip code information. GIS software was used to map the spatial distribution of unmarried teen births and sociodemographic variables, which were extracted from the 1990 U.S. Census Bureau data for the 81 zip codes in Dallas County. Correlational analysis was used to determine the relationship between unmarried teen births and the sociodemographic variables. Maps provided visual representation of the spatial distribution of unmarried teen births as compared to median household income, percent of households receiving public assistance, percent and years of education, percent of households headed by females and by males (as proxy variables for family composition), percent black, percent white, percent Hispanic, and teen density.

The results of this geographical mapping showed strong visual relationships between unmarried teen births and ethnicity, socioeconomic level, and family composition. The statistical results of the quantitative analysis supported these findings, as shown by the correlation coefficients in Table 1.

The ability to generalize the results of this study is limited by the use of 1990 U.S. Census Bureau data and by changes that can occur due to mobility of the population as well as by the use of only one of 254 counties in the state of Texas. Still, the researchers concluded that while the study of unmarried teen births and its variables is complex, this sophisticated technology was able to integrate and analyze teen births and census data together to identify where problems exist geographically. GIS analyses provided valuable information to use in solving problems through policy deliberations and health resource allocations.

**GISs and Breast Cancer Mortality**

As observed in Hsu et al., GIS technology is a valuable method for the analysis of geographical disparities in mortality among racial groups. The authors examined geographical variations in breast cancer mortality among three racial groups (non-Hispanic white, black, and Hispanic females) in Texas using a GIS and county-level breast cancer mortality data from 1990 to 2001 (n = 28,813). Data were retrieved from the Texas Department of Health “Death Files by Race” and the U.S. Census Bureau 2000 Census data and analyzed by the SatScan program, version 4.0. County centroids were used as a proxy for county locality. The Spatial Scan Statistic was used to identify localized hot spots of excess events. It has the capability to search for clusters of cases and test for statistical significance across uneven geographical population densities and conditions.

Four regions of excess breast cancer mortality were identified. The most probable regions of excess mortality were West Texas for Hispanic women (RR = 1.18; p = .001) and 42 counties along the Gulf Coast and in central Texas for non-Hispanic white women (RR = 1.12; p = .001). Excess mortality was identified involving multiple racial groups in three counties in southwest Texas and nine counties in central Texas. While geographical variations in breast cancer mortality were evident among racial groups, this study failed to demonstrate hot spot clusters or persistent spatiotemporal trends of excess mortality due to breast cancer.
GISs and Pneumocystis carinii Pneumonia

Arno et al. used GIS technology to identify geographical areas and subpopulations at increased risk for Pneumocystis carinii pneumonia (PCP), the most common and highly treatable opportunistic infection, at the time of diagnosis of acquired immunodeficiency syndrome (AIDS)\(^9\). For this study a population-based PCP index was defined as the number of PCP-related hospitalizations divided by the number of persons living with AIDS. These data were obtained from New York State Department of Health hospital discharge records and AIDS surveillance data. Using the International Classification of Diseases, Ninth Revision (ICD-9) codes, the sample consisted of 39,740 and 2262 persons hospitalized with AIDS (042.0–044.9) and PCP (136.3), respectively, in 1997. Human subject protection was achieved by removing all personal identifiers. PCP indexes were calculated, geocoded, and mapped by zip code using Maptitude GIS (Caliper Corp., Newton, Mass.).

The PCP index for New York City (citywide) was .05691 for 1997. There was a wide geographical variation of PCP index values across neighborhoods with high AIDS prevalence (West Village = .02532 vs. Central Harlem = .08696). Other areas with moderate AIDS prevalence had unusually high PCP index values (Staten Island = .14035; northern Manhattan = .08756). No statistical correlation between AIDS prevalence rates and PCP index values was noted (Spearman rank correlation coefficient, \(r = –0.118; P = .13\)). Several implications were noted from this study. First, the high PCP index values found in the areas of moderate AIDS prevalence may indicate pockets of undetected HIV infections. Secondly, high AIDS prevalence and high PCP index values may indicate poor access to primary care services. And lastly, the mapping of the geographical distribution of PCP index values provides additional information beyond simple AIDS density that is valuable for targeting prevention and service delivery to improve access and reduce morbidity and mortality in persons with HIV. A weakness of any study of AIDS is an apparent reluctance to assign HIV/AIDS disease codes.

GISs and Leptospirosis Outbreak

Barcellos and Sabroza sought to describe a 1996 outbreak of leptospirosis (n = 87 cases) in the western region of Rio de Janeiro and to examine the relationship between socioeconomic factors and leptospirosis.\(^{10}\) They used GIS technology to integrate and analyze epidemiological and socioeconomic data and to geocode addresses of identified leptospirosis cases. GIS (ArcView 3.1) capabilities allowed flood risk areas to be digitalized onto 276 census tract polygons (1:50,000 scale) obtained from the Brazilian Census Bureau. Many socioeconomic variables were available in the census data that allowed the researchers to characterize the regional and environmental conditions associated with leptospirosis. This research used urbanization patterns, income, sanitation, and population distribution. Census tracts (unit of analysis) proximal to leptospirosis cases were identified using buffers around cases (i.e., 50 m, 100 m, 150 m, etc.). ANOVA was used to determine the influence of radius (50–1500 m) on the presence of leptospirosis cases (categorical variables). The sociospatial indicators associated with leptospirosis cases (by census tract) that were found to be significant at varying distances were flood risk area (50–1000 m), solid waste collection (150–600 m), population density (150–600 m), sewerage coverage (200–1000 m), water supply coverage (200–600 m), proportion of houses (400–600 m), and population per household (500–1500 m). This study demonstrated the effectiveness of using GIS technology to describe and analyze social and environmental vulnerabilities around a single “sentinel event.”\(^{11}\) The GIS was able to identify risk factors of disease and help health planners target resources to priority populations.

Healthcare Access, Health Outcomes, and GIS Technology

Disparities in health outcomes across populations and geographical regions may reflect differences in access to healthcare services. Many studies have shown the effect of access barriers on healthcare service access and the effect of low access on health outcomes. Studies that look at access and disparate health outcomes are important to understanding and changing how and where healthcare services are located.

GISs and Access to Primary Care

Luther et al. studied the impact of community-based primary care clinics on health outcome disparity (n = 84)\(^{12}\). In this study the authors sought to (1) identify geographical communities (zip codes) with high
and low access to primary care clinics that serve ethnic and racial minorities, (2) describe sociodemographic characteristics of high- and low-access communities, (3) compare the rates of selected health outcomes for blacks (the largest minority) between high- and low-access communities, and (4) develop a model to estimate number of lives saved by primary care clinics.

Proximity was used to measure high and low access, with zip codes containing or contiguous to a clinic classified as high access and all other zip codes as low access. ANOVA was used to compare mean values of demographic variables between high- and low-access locations for both maternal and child health and chronic disease mortality outcome measures. Linear modeling was used to compare rates between high- and low-access areas for the outcome measures. Results showed that blacks from high-access areas were 50 percent more likely to receive prenatal care (OR = 0.52) than blacks from low-access areas. The difference was highly significant at the p < .001 level. Of the five models used to model chronic disease mortality health outcomes, only one (diabetes) did not show a significant difference for predicted rates. Blacks from high-access areas had lower rates of all cancers, heart disease, pneumonia/influenza, and stroke than did blacks from low-access areas. In terms of number of lives saved, this study estimated that more than 130 deaths would occur among blacks each year if blacks in the area of study had only low access to primary care programs.

**GISs and Stroke Mortality**

The National Longitudinal Mortality Study (NLMS; n = 400,000) is a prospective cohort study conducted to evaluate excess stroke mortality and the association of socioeconomic status (SES) across the “stroke belt” and the United States in general. In this study, Howard et al. looked at stroke mortality, a devastating outcome from cardiovascular disease (CVD), using survey data collected by the U.S. Census Bureau during the period from 1979 to 1985. The stroke belt is an 8- to 10-state region (153 county regions) located in the southeastern United States. The existence of this region is not completely understood but may be due to increased prevalence of CVD risk factors, decreased access to healthcare services, or other geographical factors. As mentioned earlier, low SES has been associated with decreased access to care, and decreased access has been associated with increased mortality.

In this study the researchers investigated excess stroke mortality and its association with SES by adjusting for SES. They also estimated the relative risk of stroke in this 153-county region and sought to better define the “stroke belt” or the population at risk. Findings of the study showed that for ages 35 to 54 and 55 to 74 years, stroke mortality was 1.3 times greater in the “stroke belt” than in the rest of the nation, but less than 16 percent of the excess stroke mortality was attributable to SES.

Three major limitations of the research were noted. First is the lack of data in the data set to allow for closer study of the traditional risk factors and mortality. A second limitation is that the determination of death due to stroke was dependent on death certificate information. Finally, the use of family income and education as proxy variables for SES may not reflect the complexity of SES, and better measures could reflect a greater association between stroke mortality and SES. One observed strength of this study is the very large national sample size of the NLMS data set.

**GISs and Ischemic Stroke Treatment**

Scott et al. conducted a study to identify the proportion of the Canadian population with potential access to hospitals capable of offering intravenous thrombolysis to treat acute ischemic stroke (n = 203 hospitals). Guidelines for stroke treatment recommended a 180-minute treatment window from the onset of stroke symptoms until administration of tissue plasminogen activator. The guidelines allowed 60 minutes for stroke recognition and emergency room evaluation, which left 120 minutes for transport time. The researchers used ambulance databases to determine the correspondence between time and distance. Travel times of 60, 90, and 120 minutes were found to be equated to distances of 32, 64, and 150 kilometers respectively (or 20, 40, and 65 miles). Using GIS software and the Canadian census data (1991) and interim 1996 census counts, the researchers determined the proportions of the Canadian population living within 60, 90, and 120 minutes of hospitals capable of delivering intravenous thrombolysis for acute stroke. In the Canadian population, 67.3 percent, 78.2 percent, and 85.3 percent were within 32, 64, and 105 kilometers respectively of identified hospitals capable of treating acute ischemic stroke.
One flaw of this study was the use of crude empirical measures of accessibility. The study used buffer zones of Euclidean or straight-line distances that are not the best measure of travel time. Another limitation is the use of ambulance records to calculate the average relationship between distance and travel time. Even with the identified limitations, the GIS was able to integrate census data and provide measures of healthcare access. The availability of geospatial databases and recent improvements in microcomputer-based GIS technology has made the analysis of travel time more realistic and affordable. An even more sophisticated measure is the computation of the proportion of the population within a specified travel time of healthcare facilities as demonstrated in this study.\textsuperscript{16, 17}

\textit{GISs and Access to Emergency Departments}

In 1998, Parker and Campbell explored the potential use of GIS technology in defining and analyzing the use of healthcare services by researching patterns of utilization and their variance by both socio-organizational and geographical aspects of accessibility.\textsuperscript{18} Using a survey of patients in 18 general practices in Scotland (n = 8005), this study examined patients’ access to these services as well as to local Accident and Emergency Departments (AEDs). Information was obtained using a questionnaire about the patient’s home postcode, means of travel to the practice, and perception of accessibility of local AED services. Travel and perception were measured using a five-point scale of estimated travel time (<15, 15–30, 30–45, 45–60, and >60 minutes) and a four-point scale of estimated travel distance (<1, 1–2, 2–5, and >5 miles). They also obtained information from medical records of patients attending the local AED. Postcode data were geo-referenced using GIS software (ARC/INFO). This software was used to query and display (a) straight-line and network distances between patient homes and service centers, (b) the effect of distance on the utilization of services, (c) the overall accessibility of services, and (d) the patterns of patients’ utilization and perception of health service accessibility. The results showed a distance decay effect: the use of services decreased as the patients’ distance from the service increased. Furthermore, as much as a fourfold variation between the 20 postcodes surveyed was noted. The investigation of potential socioeconomic factors suggested that those with lower AED usage might be those with lower socioeconomic scores. Findings also reveal that patients appear to believe they live closer to services than they do. Several important limitations of this research were noted. The authors discuss statistical analysis and controlling for variables but do not give these statistical results. The limitations of data collection by questionnaire and the validity and reliability of GIS analysis were not discussed.

\textbf{Summary of Findings}

\textit{Research Question 1: GIS Use in Healthcare}

GISs have been used in a vast array of health applications. Health research using GIS technology has demonstrated environmental variables of Lyme disease, correlated sociodemographic variables and teen pregnancy, analyzed geographical disparities in breast cancer mortality by racial groups, demonstrated heart disease and stroke mortality, mapped PCP and AIDS prevalence, analyzed access to emergency departments, and identified factors of a leptospirosis disease outbreak.

\textit{Research Question 2: GIS Effectiveness}

Studies presented here have related health outcomes to the level of access to healthcare. GIS technology was useful in the evaluation of environmental conditions and Lyme disease. Studies have further related heart disease and stroke mortality to geographical location of healthcare services. Geographical mapping provided for visualization of relationships between teen births and other social factors, as well as showing geographical variations in breast cancer among racial groups. GIS technology was able to demonstrate both unequal access and unequal health outcome for many priority populations (blacks, women, and geographical region). GISs were effective analytical tools for the evaluation and study of healthcare access and health outcomes in the cited studies.
Research Question 3: Implications for Future Use of GIS Technology in Healthcare

Geography and health are unquestionably linked. Whether it is problems with environmental risk or exposure, poor health screening, or a lack of access to basic resources, geography can be vital in the assessment of health issues. GISs, as a technology and as a methodology, provide for effective geographical analysis of health problems. The implications for the use of GIS technology in healthcare are vast.

In the studies cited, disparities in access to healthcare services are associated with barriers to access such as age, sex, race, ethnicity, income, insurance status, and place of residence. Studies of medical geography and epidemiology well document significant variations in health over even small geographical areas. Geographical proximity to healthcare services was shown to be a strong predictor of disparities. But there remain many unexplained differences in measures of access. More studies are needed to evaluate the impact of geographical access and important outcome measures of health.

Implications for practice. The field of health geography could benefit from studies that provide greater understanding of patterns of geography, healthcare access, and health outcomes. GIS analyses can serve to guide policy deliberations and health resource allocations. They can be instrumental for targeting of interventions to improve geographical healthcare disparities.

Implications for research. The literature cited collectively indicates the interrelated aspects of geography, accessibility, and health. What is not yet understood is the specific relationship of specific populations in their unique geographical contexts. More research is needed to explore specific social and geographical variables of specific at-risk populations. Further research is also needed in the use of GIS technology to both visually identify and empirically measure spatial relationships of geographical, environmental, and social influences on disease and other health issues. GISs can provide the technology and methodology for the study of the web of causation of health disparities. Information regarding access to healthcare services for specific populations could better describe the healthcare needs of those at risk, such as rural, elderly, low-income, and black populations, as well as specific geographical areas. Research of this nature could serve to assist healthcare planners and those who make decisions about the location of healthcare services.

Implication for theory. The field of health geography is evolving through the use of evidence-based studies. Sufficient research is available to support the use of GIS as an effective technology for the study of healthcare access and health outcomes. Knowledge generated from empirical research can form a basis for the understanding of health access and health outcomes and for the development of intervention programs to resolve health disparities.

Conclusion

Health is an outcome of multiple determinants. Individual biology and behaviors, physical and social environments, policies and interventions, and access to quality healthcare are predisposing factors that can contribute to the health of people and communities. The predisposing factors of health status are often interdependent and interrelated, creating a complex web of causation. GISs are a technology and methodology that can answer questions about the complex web of causation of many health issues. They can be effective in the integration and analysis of physical, social, and cultural environments.

The purpose of this article was to determine the link between GIS use and the evaluation of healthcare access and health outcomes. The studies reviewed have demonstrated environmental variables of Lyme disease, correlated sociodemographic variables and teen pregnancy, analyzed geographical disparities in breast cancer mortality by racial groups, mapped PCP and AIDS prevalence, and identified factors of a leptospirosis disease outbreak.

GIS technology offers many advantages in data integration, interactive querying of databases and design, and presentation of findings in the form of maps. Both the visual impact and the data analysis provided by GISs are advantages that support their use. The ability to overlay data layers allows for interpretation beyond that seen with traditional research and statistical methods. The studies reviewed used GISs to both empirically measure and visually identify and explore spatial relationships of health and
health variables. GIS use was effective in the investigation of various aspects of healthcare access and health outcomes and therefore can be an asset in the understanding and resolution of health disparities.

Barbara Ann Graves, PhD, RN, is an assistant professor at the University of Alabama’s Capstone College of Nursing in Tuscaloosa, AL.
Notes

11. Ibid., p. 302.
13. Ibid.
22. Gatrell, A. *Geographies of Health.*
**Literature Cited**


Table 1

<table>
<thead>
<tr>
<th>Sociodemographic Variable</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median household income</td>
<td>–.663</td>
</tr>
<tr>
<td>Percent of households receiving public assistance</td>
<td>.705</td>
</tr>
<tr>
<td>Percent of population black</td>
<td>.527</td>
</tr>
<tr>
<td>Percent of population white</td>
<td>.677</td>
</tr>
<tr>
<td>Percent of population Hispanic</td>
<td>.679</td>
</tr>
<tr>
<td>Percent with &lt;9 years of education</td>
<td>.739</td>
</tr>
</tbody>
</table>