Application of GIS and SPSS for prostate cancer and health disparity detection in Texas

Jose Huerta

College of Business,
Management Program,
The University of Texas,
Permian Basin, USA
Email: huerta_j@utpb.edu

Gayle Prybutok*

Department of Rehabilitation and Health Services,
College of Health and Public Service,
University of North Texas,
1155 Union Circle #311456,
Denton, TX, 76203, USA
Email: gayle.prybutok@unt.edu
*Corresponding author

Victor R. Prybutok

Department of Information Technology and Decision Sciences,
G. Brint Ryan College of Business and Toulouse Graduate School,
University of North Texas,
1800 Chestnut St., Suite 103, Denton, TX 76201, USA
Email: prybutok@unt.edu

Abstract: This study uses a geographic information system to create and analyse choropleth maps determining the distribution of prostate cancer in Texas and uses statistical package for the social sciences (SPSS) software to analyse social determinants of health that may explain prostate cancer mortality. The data, collected for period 1999–2009, was furnished by the Texas Health Rankings and VitalWeb. The dataset was for 1999–2004 and 2004–2009. It comprised age-adjusted data specific to the 2000 US Standard Population data, based on an age-distributed and -weighted methodology to create age adjustments for statistical purposes. The study found there was a statistically significant ($P < 0.05$) percentage of African Americans with age-adjusted prostate cancer mortality, but no statistically significant correlations were found in other races. The study indicates a number of ways medical communities and public health agencies can employ geographical information system (GIS) and SPSS to screen for and treat prostate cancer more effectively.

Keywords: GIS; geographical information system; prostate cancer; social determinants of health; spatial patterns; SPSS; statistical package for the social sciences; choropleth maps; geographic information systems.

Reference to this paper should be made as follows: Huerta, J., Prybutok, G. and Prybutok, V.R. (2023) 'Application of GIS and SPSS for prostate cancer and health disparity detection in Texas', *Int. J. Healthcare Technology and Management*, Vol. 20, No. 4, pp.287–325.

Biographical notes: Jose Huerta is an experienced Healthcare Information Technology professional with over 20 years of experience in the healthcare information technology (HIT) space. He holds a Master of Science (MS) degree in Business Management from Troy University and a Bachelor of Science (BS) degree in Business Management from Park University. He holds two certifications in electronic health record technology from the American Health Information Management Association (AHIMA). He is a board member of the Texas Health Information Management Association (TXHIMA) and is currently a PhD candidate in the College of Information at the University of North Texas in Denton, Texas.

Gayle Prybutok earned her PhD from the University of North Texas in Information Science. She is an Assistant Professor, Health Services Administration in the Department of Rehabilitation and Health Services, College of Health and Public Service, at the University of North Texas. She is an Associate Professor, Health Services Administration, and was instrumental in creating that program and the PhD in Health Services Research. She holds a Bachelor's in Nursing from Thomas Jefferson University, an MBA from Texas Woman's University and a PhD in Information Science with a focus on Health Informatics from the University of North Texas. She formerly served as the Chief Nursing Officer of a local hospital, Director of home health and hospice agencies, and was the Executive Director of a national non-profit funded by NIH to procure human tissue for research.

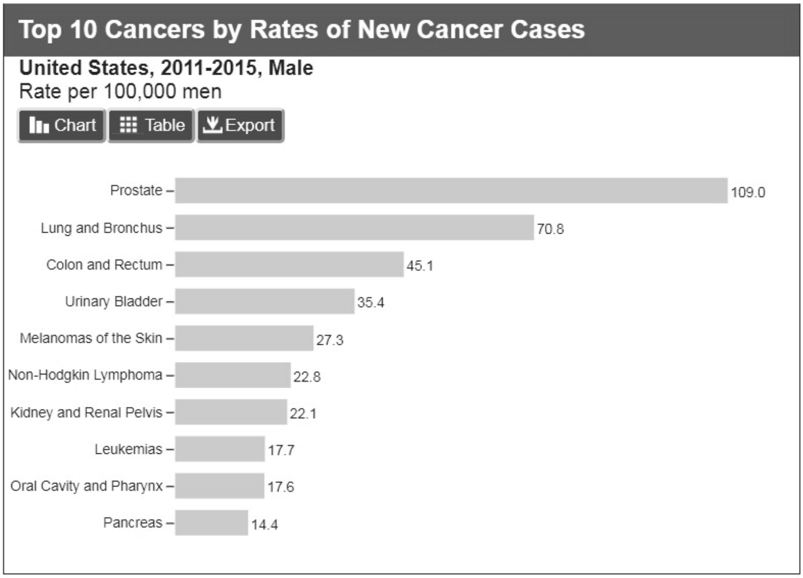
Victor R. Prybutok is a Regents Professor of Decision Sciences in the Information Technology and Decision Sciences Department, Vice Provost for Graduate Education, and Dean of the Toulouse Graduate School at the University of North Texas. He received from Drexel University, his BS with High Honours in 1974, an MS in Bio-Mathematics in 1976, an MS in Environmental Health in 1980, and a PhD in Environmental Analysis and Applied Statistics in 1984. He is an American Society for Quality certified quality engineer, certified quality auditor, certified manager of quality/organisational excellence, and an accredited professional statistician (PSTAT®) by the American Statistical Association. He has authored over 200 journal papers, more than 300 conference presentations/proceedings, and several book chapters. In 2017 he received the American Society for Quality Gryna Award for a co-authored manuscript and in 2018 was awarded the Decision Sciences Institute Lifetime Distinguished Educator Award.

1 Introduction

1.1 Prostate cancer in the USA

Prostate cancer has increased across the US. During 2011–2015, prostate cancer became the most prevalent cancer among males with lung and bronchus cancer following. US age-adjusted prostate cancer incidence was 109.0 per 100,000, which consisted of 953,204 cases (U.S. Department of Health and Human Services, 2018). See Figures A and B.

Figure A



Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
U.S. Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

Figure B

Top 10 Cancers by Rates of New Cancer Cases
United States, 2011-2015, Male
Rate per 100,000 men

Cancer Type	**Age-Adjusted Rate**	**Case Count**	**Population**
Prostate	109.0	953,204	778,060,201
Lung and Bronchus	70.8	572,602	778,060,201
Colon and Rectum	45.1	365,934	778,060,201
Urinary Bladder	35.4	275,807	778,060,201
Melanomas of the Skin	27.3	219,303	778,060,201
Non-Hodgkin Lymphoma	22.8	182,273	778,060,201
Kidney and Renal Pelvis	22.1	184,358	778,060,201
Leukemias	17.7	139,112	778,060,201
Oral Cavity and Pharynx	17.6	151,268	778,060,201
Pancreas	14.4	117,201	778,060,201

Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html U.S. Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

Prostate cancer is prevalent in the US, and the analysis of spatial patterns of prostate cancer distribution, along with an examination of their changes over time, promises significant insights into how the disease spreads geographically over time. Tools that analyse such patterns have helped determine the spatial distribution of disease and its geographic spread (Bui and Pham, 2016). This study examines the geography and spatiotemporal patterns of prostate cancer in Texas counties, drawing upon county-level, mortality-rate data during the decade 1999–2009. Through the utilisation of statistical package for the social sciences (SPSS) software and geographical information system (GIS) technologies, the study leverages analytical, mapping, and visualisation functionalities, providing new insights that can help explain health disparities in prostate cancer. Other GIS-like applications, such as web-based spatial processing tools, have been used successfully in other countries to measure spatial patterns in an effort to track disease incidences (Bui and Pham, 2016). In addition, social determinants of health include race, socioeconomic status (SES), and healthcare accessibility (Wilkinson and Marmot, 2003) are evaluated in an attempt to explain the existence and geographical distribution of the disease.

Other studies analysing the role of social determinants in healthcare studies (Shulan et al., 2015) have shown contradictory results regarding social determinants such as SES in explaining health disparities in prostate cancer, as well as its prevalence (Cheng et al., 2009). Nevertheless, these factors remain important to study further in cancer because in other geographic settings, the prevalence of malaria and other diseases have been found to be correlated with environmental and socioeconomic factors (Bui and Pham, 2016). In this study, the term *health disparities* refers to the differences found in incident cases, deaths, and healthcare access due to variables such as socioeconomic status, settlement or habitation, gender, or ethnic and racial makeup (LaVeist and Pierre, 2014).

To mitigate the problem of health disparities, the US Office of Disease Prevention and Health Promotion developed Healthy People 2020, which tracked rates for the following components of disease: illness, mortality, long-lasting conditions, and other factors in health outcomes that may correlate with factors including race and ethnicity, gender, geographic location, and the like ("Disparities"). This study advances that initiative, focusing particularly on delineating and understanding prostate cancer mortality geographically in the state of Texas in relation to social determinants including race, socioeconomic status, and healthcare access. The study specifically addresses how such factors influence disparities in the disease. Although gender is a social determinant, it was not evaluated because prostate cancer affects only males. The following research questions are addressed by the study:

1 What is the geographic distribution of prostate cancer deaths across Texas?

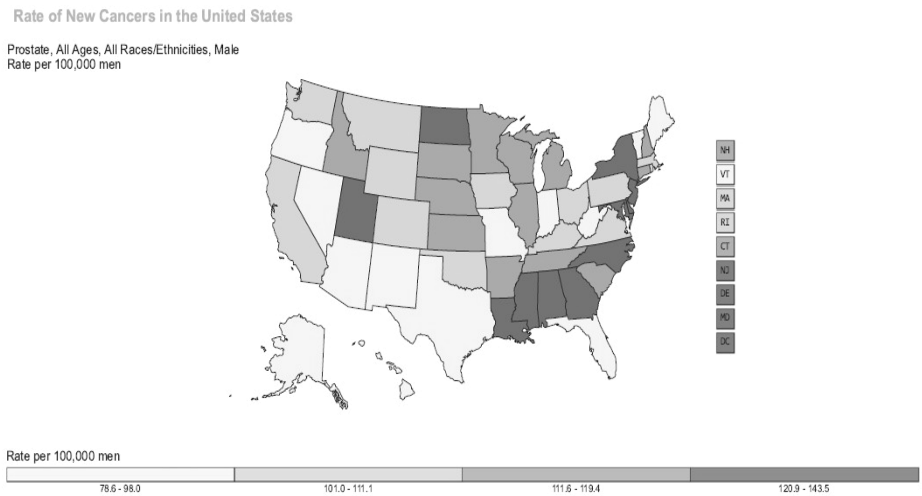
2 Why are prostate cancer deaths geographically distributed in that way?

3 How does the geographic, spatial-temporal pattern of the disease change over time?

2 Background

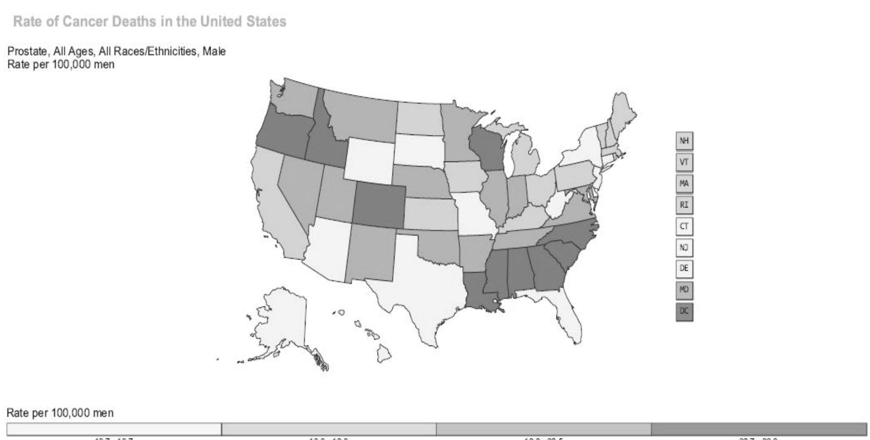
Prostate cancer affects the prostate gland cells, usually in the form of high cell-growth rate, and the prostate cancer risk increases with age (Klassen and Platz, 2006). In US, one in six males age 50 and above will be diagnosed with prostate cancer (Penson and Chan, 2007). It is number two in both incidence and mortality (Figure 2(A) and (B)). During 2011–2015, about 953,204 new cases of prostate cancer were reported, and 140,086 men died in US (Figure 1(A) and (B)).

Figure 1(A)



Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html Centers for Disease Control and Prevention
US Cancer Statistics: The Official Federal Cancer Statistics

Figure 1(B)



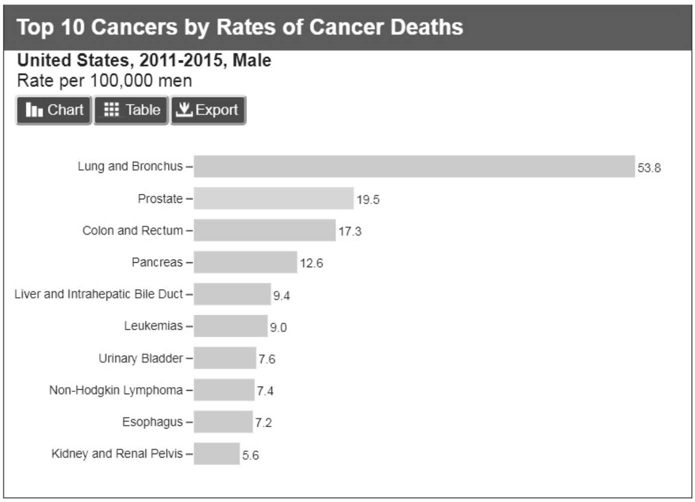
Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

2.1 Prostate cancer in Texas

The Texas Cancer Registry (TCR) provides data and cancer measures. TCR has technical and functional capacities leveraging geographical maps. This registry has assisted in trending prostate cancer morbidity and mortality data (Texas Department of State Health Services). According to the Texas Health and Human Services Cancer Registry dataset (Texas Department of State Health Services. "Cancer Incidence Leading Sites 2011–

2015"), during 2011–2015, there was an annual average of 11,572 new cases of prostate cancer in Texas. Non-Hispanic Whites comprised an annual average of 7367 incidents registered. Blacks comprised 1807. Asian/Pacific Islanders comprised 188. American Indian/Alaska Natives comprised 30. Hispanics comprised an annual average of 2059. Overall, prostate cancer was more conspicuous among those aged 50 and older (Texas Department of State Health Services. "Cancer Incidence Leading Sites 2011–2015").

Figure 2(A)



Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

Figure 2(B)

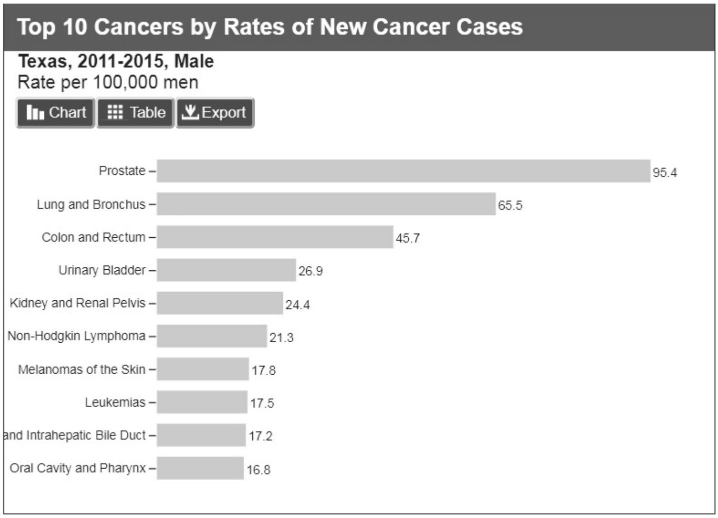
Top 10 Cancers by Rates of Cancer Deaths
United States, 2011-2015, Male
Rate per 100,000 men

Cancer Type	**Age-Adjusted Rate**	**Death Count**	**Population**
Lung and Bronchus	53.8	427,587	778,060,201
Prostate	19.5	140,086	778,060,201
Colon and Rectum	17.3	135,542	778,060,201
Pancreas	12.6	100,599	778,060,201
Liver and Intrahepatic Bile Duct	9.4	80,526	778,060,201
Leukemias	9.0	67,201	778,060,201
Urinary Bladder	7.6	55,652	778,060,201
Non-Hodgkin Lymphoma	7.4	56,402	778,060,201
Esophagus	7.2	59,082	778,060,201
Kidney and Renal Pelvis	5.6	45,076	778,060,201

Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

As in US as a whole, of new cancer cases, prostate cancer was the number one cancer in Texas, followed only by lung and bronchus cancer, but the age-adjusted rate in Texas was slightly lower than the national rate, at 95.4 per 100,000 population, which consisted of 57,860 cases (Figure 3(A) and (B)).

Figure 3(A)



Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

Figure 3(B)

Top 10 Cancers by Rates of New Cancer Cases
United States, 2011-2015, Male
Rate per 100,000 men

Cancer Type	**Age-Adjusted Rate**	**Case Count**	**Population**
Prostate	95.4	57,860	65,783,771
Lung and Bronchus	65.5	36,272	65,783,771
Colon and Rectum	45.7	26,655	65,783,771
Urinary Bladder	26.9	14,150	65,783,771
Kidney and Renal Pelvis	24.4	14,675	65,783,771
Non-Hodgkin Lymphoma	21.3	12,208	65,783,771
Melanomas of the Skin	17.8	10,173	65,783,771
Leukemias	17.5	9,986	65,783,771
Liver and Intrahepatic Bile Duct	17.2	11,048	65,783,771
Oral Cavity and Pharynx	16.8	10,481	65,783,771

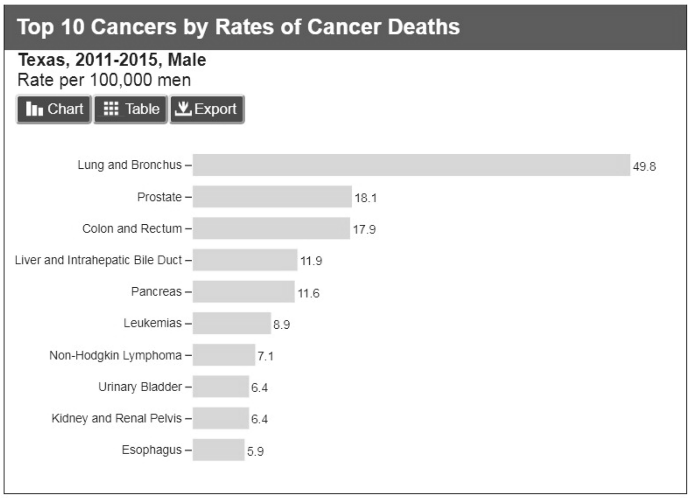
Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

The Texas Health and Human Services Cancer Registry dataset (Texas Department of State Health Services. "Cancer Mortality Leading Causes, 2011–2015") indicates that during 2011–2015, there was an annual average of 1,695 deaths. Prostate cancer mortality rates was higher among Blacks (35.7), followed by Whites (17.3), Hispanics

(15.2), Non-Hispanic Asian/Pacific Islanders (7.9), and American Indian/Alaska Natives (5.8). The higher rates of prostate cancer mortalities recorded were among age 50 and older. Among those aged 50–59, prostate cancer mortality rates were higher almost six times higher than those aged 40–49 (Texas Department of State Health Services. "Cancer Mortality Leading Causes, 2011–2015").

The age-adjusted deaths from prostate cancer deaths in Texas was 18.1, which consisted of 8,519 deaths (Figure 4(A) and (B)).

Figure 4(A)



Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

Figure 4(B)

Top 10 Cancers by Rates of Cancer Deaths
United States, 2011-2015, Male
Rate per 100,000 men

Cancer Type	**Age-Adjusted Rate**	**Death Count**	**Population**
Lung and Bronchus	49.8	26,917	65,783,771
Prostate	18.1	8,519	65,783,771
Colon and Rectum	17.9	9,897	65,783,771
Liver and Intrahepatic Bile Duct	11.9	7,267	65,783,771
Pancreas	11.6	6,432	65,783,771
Leukemias	8.9	4,579	65,783,771
Non-Hodgkin Lymphoma	7.1	3,663	65,783,771
Kidney and Renal Pelvis	6.4	3,558	65,783,771
Urinary Bladder	6.4	3,105	65,783,771
Esophagus	5.9	3,357	65,783,771

Source: https://gis.cdc.gov/Cancer/USCS/DataViz.html
US Cancer Statistics: The Official Federal Cancer Statistics
Centers for Disease Control and Prevention

2.2 Geographic impact and social determinants in health

Geography is essential to understanding disease and its spread. The geographical connection between people and their environments consists many components that affect the social, economic, and physical aspects of people's lives (Klassen and Platz, 2006). Social determinants such as race, environment, and socioeconomics have been determined to have correlations to health (LaVeist and Pierre, 2014). The present study articulates three main hypotheses to help validate the involvement of social determinants related to health disparities in prostate cancer.

Hypothesis 1: There is a relationship between minority race and the geography of prostate cancer mortality in Texas. A positive relationship is hypothesised between the percentage of African Americans and the age-adjusted death rate (AADR) of prostate cancer. The same is expected for the percentage of Hispanics and other races combined.

Hypothesis 2: A relationship exists between socioeconomic status (income level) and prostate cancer mortality. A negative relationship is hypothesised between income and age-adjusted death rate (AADR) prostate cancer mortality. A positive relationship is hypothesised between healthcare costs, unemployment, uninsured adults and age-adjusted death rate (AADR) prostate cancer mortality. That is, as each individual variable increases, AADR prostate cancer mortality also increases.

Hypotheses 3: There is a correlation between healthcare access and prostate cancer mortality. A negative relationship is hypothesised such that an increase in healthcare access, namely access to primary care physicians, results in a decrease in prostate cancer mortality.

3 Literature review

3.1 Prostate cancer

Three factors are considered risks for prostate cancer: age, race, and heredity (Attard et al., 2016). These factors are also important to assessing the distribution of the disease geographically. That is, the variables may be used to map out incidences and mortalities of prostate cancer using geographical components, such as demographics, to visualise where the disease occurs. For example, a person that has the disease can be categorised in a specific ethnic group or race, and race can be used to help determine the spread of the disease through the geographical space of interest. Race and ethnicity are based on genetic variations inherited from an individual's parents (Pearce et al., 2004). With these genetic variations, individuals can inherit other traits that increase the risk of diseases such as prostate cancer (Rebbeck, 2017). For example, prostate cancer is dominant in black men of African lineage. In fact, the highest mortalities occur among men with Afro-Caribbean and sub-Saharan African descent. In 2008–2011, the mortality rate for black men was 43 per 100,000, followed by whites (19.8), Hispanics (17.8), and Asians/Pacific Islanders (9.4) (Rebbeck, 2017). Therefore, because the disease has high hereditary risks, race/ethnicity is an important factor to study in prostate cancer.

3.2 Health disparities

Health disparity exists when health differences impact a group negatively. Measures of health disparity can encompass the portion of a population affected by disease, its severity, its symptoms, and its mortalities ("Disparities in Health and Health Care: Five Key Questions and Answers"). Differing healthcare access and capacity to receive disease screening are important factors ("Health Disparities: MedlinePlus"). Such disparities effect racial and ethnic minorities most severely, but socioeconomics, gender, age, geography and disability are also significant components (Braveman, 2014; Kumar et al., 2018). Healthy People 2020 highlighted the effects of these and other such forces, including mental health and religion ("Disparities").

3.2 Social determinants of health

Social determinants of health involve conditional settings and circumstances that affect populations. They may include factors such as birthplace, where people grow up, where they live, and where they work. Other factors include age, money, and political conditions. Social determinants can negatively affect health, producing health inequalities, which can be defined as unfairness in the health status of individuals or groups ("About Social Determinants of Health").

There are five main categories in social determinants of health. These are economic stability, education, health, community, and neighbourhood. Elements such as impoverishment and deprivation, stability in the housing market, security of food, and the stability and quality of employment make up economic stability. A degree from high school or higher education, childhood education access that is of high caliber, and the ability to read are components of education, and any of these can affect health results. Elements of health include health insurance, literacy in relation to health terms or health literature, and healthcare access. Discrimination, workplace conditions, civic participation, community belonging, and incarceration are elements of community. Quality of air and water, neighbourhood safety, access to healthy foods, transportation, and housing are neighbourhood elements. Anytime these are deficient, health is affected negatively ("Social Determinants of Health: What Medical Students Need to Know", 2019).

The social ladder can determine the life expectancy for groups. In this study, the term *social ladder* is a shorthand to describe a population's socioeconomic status, encompassing the typical social determinants that accompany a particular position. The lower on the social ladder, the greater risk for disease, and the lower expectancy for lifespan. Impoverishment in the social and economic dimensions of peoples' lives can affect quality of health. For this reason, it is imperative that health policies focus on changing negative outcomes surrounding determinants of health. Disadvantages can come in the form of poor education, employment insecurities, limited job mobility, poor housing conditions, family responsibilities in difficult social conditions, and insufficient retirement resources. The often unmanageable stress and anxiety that such disadvantages produce can lead to compromised health or early death. Diet and food supply are important factors shaped by social ladder position. Food shortage can lead to diseases of malnutrition, while overeating can lead to diseases such as cancer and diabetes. Security and satisfaction in employment can be conducive to good health, while high unemployment rates are associated with sickness and early death. In the form of poor education, financial difficulties stemming from employment insecurity can generate psychological stressors that affect health negatively (Wilkinson and Marmot, 2003).

3.3 Healthcare access

Healthcare access is an important factor in prostate cancer mortality. Access to healthcare facilities that have the clinical and technical capacities to conduct prostate-specific antigen (PSA) tests for early detection and treatment are imperative to fighting the disease (Major et al., 2012). The availability of primary care physicians or providers that specialise in prostate cancer is also important because a disproportionate availability of prostate cancer specialists throughout a geographic area is directly correlated to the mortality rate (Kim et al., 2017).

To help facilitate the discovery of healthcare access among researchers, geographic information systems can help to produce maps, not only of the disease but also of other important healthcare variables. This type of mapping can be a simple choropleth map, which is normally used to visualise the geography of variables such as healthcare accessibility (Sherman et al., 2014).

3.4 Mode of spread

The geographic distribution of prostate cancer has been analysed to determine the mode of spread across geographic regions. Typically, datasets from two or more time periods are used to analyse how the spatial pattern of the disease is changing over time. Other factors that are taken into consideration in the geographic distribution analysis are location, at-risk population, number of cases or incidences, and sometimes an age-adjusted rate (Gregorio et al., 2004). Tracking a spatiotemporal pattern can help researchers understand where the disease occurs and why. The findings of such studies can then help to pinpoint other factors that may be associated with increase or decrease of the disease. For example, prostate cancer mortalities may decrease as healthcare access and rate of PSA screening increase.

4 Methodology and data sources

The International Classification of Diseases, Tenth Revision (ICD-10-CM) code used for the study is C61, defined in this study as prostate cancer. The ICD-10-CM is based on a classification logical system that helps healthcare practitioners and healthcare workers identify disease through coding processes to determine disease diagnosis (https://www.cdc.gov/nchs/icd/icd10cm.htm). GIS-produced choropleth maps were used to visualise the geography of the variables. Two main datasets were obtained for the study. Both datasets were specific to Texas. The first dataset was obtained through VitalWeb, an online user-intuitive website that houses large and complex datasets and leverages the health data analysis software Vitalnet to conduct analysis (https://www.ehdp.com/vitalnet/overview.htm). The dataset obtained was comprised of age-adjusted prostate cancer mortality data per 100,000 from 1999 to 2009, and the variables used for the study were name of county (Name) and age-adjusted death rate (AADR). The calculated state average was 23.0. The age-adjustment was set to the 2000 US Population standard.

The second dataset contained the explanatory variables, or social determinants, and was obtained from the 2012 Texas Health Rankings published dataset. The published data was provided and compiled by the County Health Rankings & Roadmaps program, which was created as a central hub for reliable community health data (http://www.countyhealthrankings.org/about-us).

The explanatory variables used to address Hypothesis 1 were percent of African American, percent of Hispanic, and percent of Other Combined Races. The variable Other Combined Races is defined as percent of American Indian and Alaskan Native, percent of Asian, and percent of Native Hawaiian/Other Pacific Islander. The source of data was US Census Bureau records for 2009. The compiled dataset was most applicable to explain prostate cancer mortality for 1999–2009, the time period used in the study. Additional races could not be assessed and are addressed in the limitations section of this study. A Pearson bivariate correlation analysis was conducted and analysed against the AADR prostate cancer variable obtained from the VitalWeb dataset.

Variables analysed for Hypothesis 2 were median household income, percent of uninsured adults, healthcare costs, and percent of unemployed. The variables with their respective data sources and data year were: median household income (Small Area Income and Poverty Estimates [SAIPE], 2010); percent of uninsured adults (Small Area Health Insurance Estimate [SAHIE], 2009); healthcare costs (Health Resources and Services Administration [HRSA], 2007); and percent of unemployed (Local Area Unemployment Statistics, Bureau of Labor Statistics, 2010). Pearson's statistical bivariate correlation was applied to analyse the relationship between the variables and AADR prostate cancer.

To address hypothesis 3, the variable of primary care physicians was used. The source of the data was Heath Resources and Services Administration, Area Resource File (ARF) for 2009. This variable was used to determine healthcare access to counties. As in the statistical analysis of hypothesis 1 and 2, a Pearson bivariate correlation analysis was conducted between PCP rate and AADR prostate cancer to measure the relationship between both variables.

5 Results

5.1 Brief notable findings

Counties with the highest mortality rates and above the Texas state mean (23.0) were Menard (35.0), Crockett (34.9), Dimmit (43.2), Refugio (38.3), Harrison (26.5), Delta (35.7), Jack (38.5), Haskell (37.0), King (39.7), Garza (38.7), Floyd (44.2), Bailey (43.2), Cochran (43.2), and Mitchell (34.3). Similarly, very high concentrations showing high mortality rates were found in smaller, dispersed clusters, mainly in northwestern Texas, exemplified by Bailey (43.2), Cochran (43.2), Floyd (44.2), King (39.7), Haskell (37), Garza (38.7), and Mitchell (34.3) counties. In contrast, counties with the lowest rates of prostate cancer mortality and below the state mean were Loving (0), Glasscock (0), Oldham (0), Roberts (0), Briscoe (0), Kenedy (0), Zapata (6.8), La Salle (4.1), and Edwards (5.3), located in the panhandle and southern Texas.

The eastern region showed Harrison, Rusk, Cherokee, Anderson, Houston, Shelby, Nacogdoches, Sabine, Jasper, Tyler, Hardin, Jefferson, Liberty, Chambers, Galveston, Brazoria, Matagorda, Wharton, Colorado, Washington, Waller, Robertson, and Falls counties with high concentrations, and Panola county with a very high concentration. Likewise, AADR ranged above the state mean and was between 25.6–33.5 in western counties such as El Paso, Reeves, Ward, and Pecos. Brewster showed high concentrations, and Crockett showed a very high concentration. Additionally, high concentrations were also found in Dallam, Hansford, Wheeler, Donley, Hall, Swisher, Lamb, Crosby, Lynn, Dawson, Gaines, and Stonewall counties. The central region had high concentrations among Wise, Palo Pinto, Parker, Eastland, Erath, Comanche, Brown, Mills, Coryell, Runnels, Concho, Kimble, Sutton, and Gillespie, with very high concentrations showing for Jack and Menard counties. Eastern Texas had more counties with a greater concentration of prostate cancer in comparison to the western and central regions.

5.2 Change map analysis

The main time period of 1999–2009 was broken into two portions for further analysis in this section. Time Period 1 was 1999–2004 (see Figure A11), and Time Period 2 was 2005–2009 (see Figure A12). Prostate cancer mortality rates during Time Period 1 were mostly moderate for the eastern region. However, Time Period 2 showed a heavier presence for the eastern region.

In an effort to review the changes in the disease mortality rate through time, a new map was created, namely a choropleth map charting the difference between the two time periods. The new change map showed the areas where mortality rates tended to improve and areas where the disease tended to worsen (see Figure A13).

The greatest improvements between Time Period 1 and Time Period 2 were around the Panhandle region in counties such as Dallam, Cochran, and Borden. In the Permian Basin, counties such as Crane and Reagan also showed great improvements. Additionally, there were improvements in the southeastern region such as in Goliad County.

Areas that tended to worsen over both time periods were in the western region, in counties such as Hudspeth and Brewster. The Permian Basin region also showed Ward County worsening. Counties surrounding the Panhandle region that worsened were Sherman, Hartley, and Hansford. The north-central region showed Stephens, Shackelford, and Throckmorton counties worsening over the two time periods.

Overall, the majority of the regions showed slight improvements. Regional improvements appeared in areas of western Texas such as El Paso County, areas of the Permian Basin such as Andrews and Ector counties, areas of the Panhandle region such as Ector and Moore counties, areas of the southern region such as Starr and Cameron counties, areas of the Eastern region such as Harris and Fort Bend counties, areas of the North region such as Cook and Grayson counties, areas of the North East region such as Bowie and Cass counties, and areas of the central region such as San Saba, Mason, and Llano counties.

5.3 Race/ethnicity and prostate cancer mortality

The overall age-adjusted death rate (per 100,000) of prostate cancer from 1999 to 2009 for all races combined, including Whites, was 9.5, a total of 18,315 deaths across the state (Table A). Based on this data, there were 9.5 deaths per every 100,000 population. Blacks had the highest rate among the group at 18.8, followed by Whites (9.0), Hispanics (7.9), and Others (3.7).

Table A

Prostate Cancer Deaths by Race in Texas for 1999-2009					
Race	*White*	*Black*	*Hispanic*	*Other*	*Total*
Rate*	9	18.8	7.9	3.7	9.5
Deaths	12,456	3059	2646	154	8315

*Rate is Age-Adjusted Death Rate (per 100,000)

DataSource: https://www.ehdp.com/vn/rw/txu1/eg2/8a17xgqn-tbl.htm
Texas VitalWeb ICD-10 Underlying Cause of Death

The explanatory variables from the Health Rankings dataset were used to analyse several minority group populations to help explain the high concentrations of prostate cancer mortality and its geography. These groups were African Americans, Hispanics, and Other Races Combined, which were American Indian, Alaskan Native, Asian, Native Hawaiian, and Other Pacific Islanders.

5.4 *Analysis of African Americans*

The first minority group analysed in relation to geographic concentrations of prostate cancer mortality were African Americans. There was a high to very high population rate and concentration of African Americans in the eastern part of Texas (Figure A3). A comparison of the geography of the percentage of African Americans (Figure A3) and the geography of prostate cancer (Figure A1) showed a visual similarity between the two. That is, for eastern Texas, there was a high concentration of Blacks, which appeared geographically similar to the high concentrations of prostate cancer. For example, counties with percentage of Blacks to AADR (per 100,000) were as follows (see Table B):

Table B % of Blacks to AADR (per 100,000)

Counties	*% of Blacks*	*AADR*
Jefferson	34.7	31.3
Waller	26.0	28.3
Houston	25.9	32.0
Anderson	22.5	27.9
Harrison	22.4	26.5
Robertson	22.2	29.4
Rusk	18.1	31.3
Shelby	18.0	29.3
Washington	17.7	27.4
Jasper	17.2	30.9
Nacogdoches	16.6	28.4

Table B % of Blacks to AADR (per 100,000) (continued)

Counties	*% of Blacks*	*AADR*
Wharton	14.7	32.8
Cherokee	14.6	30.5
Mitchell	14.4	34.3
Galveston	14.3	27.9
Colorado	14.2	27.7
Liberty	12.1	29.0
Matagorda	11.8	26.2
Tyler	11.8	26.8
Brazoria	11.4	26.9
Chambers	10.8	29.4
Sabine	9.3	27.4
Cochran	7.2	43.2
Hardin	7.2	29.4
Garza	6.2	38.7
Floyd	4.6	44.2
Haskell	4.3	37.0
King	3.8	39.7
Bailey	2.0	43.2

In contrast, other pockets showing very high concentrations of prostate cancer did not show high concentrations of Blacks for the region. Examples were Bailey (2.0% of Blacks to 43.2 AADR), Cochran (7.2% of Blacks to 43.2 AADR), Floyd (4.6% of Blacks to 44.2 AADR), King (3.8% of Blacks to 39.7 AADR), Haskell (4.3% of Blacks to 37 AADR), Garza (6.2% of Blacks to 38.7 AADR), and Hardin (7.2% of Blacks to 29.4 AADR) county.

5.5 Analysis of hispanics

The second minority group analysed in relation to the geography of prostate cancer mortality was Hispanics. There was a high to very high concentration of Hispanics in the western, northwestern, and southern regions of Texas, especially along the Mexico-Texas border-state line (Figure A4). A comparison of the southern geography of the percent of Hispanics (Figure A4) and the geography of prostate cancer (Figure A1) did not show a visual similarity between the two. In fact, prostate cancer mortality in south Texas showed the lowest AADR (Figure A1), while Hispanics were a dominant group in the region (Figure A4). For example, counties with percentage of Hispanics to AADR (per 100,000) were as follows (see Table C):

Table C % of Hispanics to AADR (per 100,000)

Counties	*% of Hispanics*	*AADR*
Starr	97.2	11.8
Brooks	90.3	13.9
Hidalgo	89.8	14.5
Zapata	88.9	6.8
Duval	87.4	9.8
Willacy	86.8	14.5
Cameron	86.6	17.3
La Salle	77.1	4.1
Kenedy	68.8	0.0
Bailey	56.8	43.2
Floyd	49.6	44.2
Cochran	48.7	43.2
Garza	41.2	38.7
Mitchell	33.3	34.3
McMullen	33.1	13.4

In contrast, there were pockets in the western and northwestern regions with high concentrations of prostate cancer and moderate concentrations of Hispanics. Examples were Bailey (56.8% of Hispanics to 43.2 AADR), Cochran (48.7% of Hispanics to 43.2 AADR), Floyd (49.6% of Hispanics to 44.2 AADR), Garza (41.2% of Hispanics to 38.7 AADR), and Mitchell (33.3% of Hispanics to 34.3 AADR) counties.

Since mortality is reported by place and time of death, it is imperative to take into consideration the migration of these groups to locations where better healthcare access exists. For example, Hispanics in southern Texas may have migrated for more specialised cancer care to eastern Texas, where renowned cancer care facilities exist. Likewise, Hispanics in the western region may have migrated for more specialised cancer care to the far western region, namely El Paso County.

5.6 Analysis of other races

The combination of remaining nonwhite races was categorised as Other Combined or Other Races Combined in the analysis. Whites could not be analysed due to the limitations of the study (see Research Limitations). This combination was comprised of Alaskan Native, American Indian, Native Hawaiian, Asian, and Other Pacific Islanders. There was a high to very high concentration of small clusters for Other Races Combined in the northern, central, and eastern parts of Texas (Figure A5). A comparison of the northern geography of the percentage of other races combined (Figure A5) and the geography of prostate cancer (Figure A1) showed a visual similarity between the two. That is, northern Texas, specifically Denton, Collin, Tarrant, and Dallas counties, showed a high concentration of other races combined, which appeared geographically similar to the high concentrations of prostate cancer for their respective counties. For example, counties with percentage of other races combined to AADR (per 100,000) were as follows (see Table D):

Table D % of other races (excluding Whites) to AADR (per 100,000)

Counties	*% of Other Races*	*AADR*
Fort Bend	15.8	22.1
Collin	11.0	22.1
Harris	6.7	24.8
Denton	6.6	25.4
Travis	6.4	22.0
Dallas	5.7	24.3
Brazoria	5.5	26.9
Tarrant	5.4	25.4
Williamson	5.1	17.6
Calhoun	4.9	23.5
Brazos	4.8	22.0
Bell	4.5	24.9

In contrast, the central region showed high concentrations of Other Races Combined, while prostate cancer concentrations were shown to be light to very light. Examples were Bell (4.5% of Other Races Combined to 24.9 AADR), Williamson (5.1% of Other Races Combined to 17.6 AADR), and Travis (6.4% of Other Races Combined to 22.0 AADR) counties. There was also a moderate concentration of Other Races Combined in eastern counties such as Brazos (4.8% of Other Races Combined to 22.0 AADR) and Harris (6.7% of Other Races Combined to 24.8 AADR). The eastern region was not shown to have a strong similarity in the comparison of the geography of prostate cancer and percent of Other Races Combined.

5.7 Hypothesis 1 findings

Hypothesis 1 of this study posited a positive relationship between race and the geography of prostate cancer mortality, specifically among the percentage of African Americans, Hispanics, and Other Races Combined. A Pearson bivariate correlation matrix was conducted to analyse the relationship between three main minority groups and the age-adjusted death rate (AADR) in prostate cancer (Table E) specific to the state of Texas. A close examination of Table 1 indicates that the percentage of African Americans in Texas shows a weak but positive correlation with the percentage of African Americans and AADR. This positive correlation was statistically significant at the 0.001 level, demonstrating a correlation between higher death rates among the African American Population in Texas. Percentage of Hispanics shows a slight decrease in AADR, while the percentage of Other Combined shows a slight increase in AADR.

Table E

		Correlations			
		AADR	*African American*	*Hispanic*	*Other Combined*
AADR	Pearson Correlation	1	0.245**	–0.097	0.043
	2-tailed Sig.		0.000	0.125	0.491
	N	254	254	254	254
African American	Pearson Correlation	0.245**	1	–0.393**	0.194**
	2-tailed Sig.	0.000		0.000	0.002
	N	254	254	254	254
Hispanic	Pearson Correlation	–0.097	–0.393**	1	0.023
	2-tailed Sig.	0.125	0.000		0.717
	N	254	254	254	254
Other combined	Pearson Correlation	0.043	0.194**	0.023	1
	2-tailed Sig.	0.491	0.002	0.717	
	N	254	254	254	254

**Correlation is significant at the 0.01 level (2-tailed).

(Table shows percentage for African American, Hispanic, and Other Combined. AADR represents age-adjusted death rate).

5.7.1 Socioeconomics and prostate cancer

Socioeconomic factors were analysed to help explain the prostate cancer mortality rate in geographic setting and status. These factors were household median income, healthcare costs, unemployment, and uninsured adults.

5.7.2 Household median income

The 2012 Texas Health Rankings used data derived from the 2010 Small Area Income and Poverty Estimates (SAIPE) to determine the overall median household income of $48,622 for Texas (Table F. Texas Median Household Income). The study analysed the geography of household median income.

Table F Texas median household income

Texas median household income	
Summary information	
Data Year Used:	2010
Range in Texas (Min-Max):	$22,948–$81,113
Overall in Texas:	$48,622

Source: http://www.countyhealthrankings.org/app/texas/2012/measure/factors/63/data. County Health Rankings & Roadmaps program: 2012 Texas Health Rankings Compilation

Household median income was below the Texas median household income across the southern region, especially regions adjacent to the Mexico-Texas border (Figure A6), an area with a high population of Hispanics (Figure A4), but mostly low to moderate concentrations of prostate cancer (Figure A1). High levels of household median income were shown in the eastern parts of Texas, with other high levels clustering in central and northern Texas, as well as the western regions known as the Permian Basin. With the exception of the eastern region, regions showing high concentrations of household median income, such as the northern region, showed only low to moderate concentrations of prostate cancer. There were small pockets of clusters in the western (Permian Basin), northwestern (Panhandle), central, northern, and eastern parts of Texas that showed household median income above the state overall household median income.

5.7.3 Healthcare costs

The 2012 Texas Health Rankings used data from the Dartmouth Atlas of Healthcare 2007 data to determine the overall healthcare costs of $10,889 for Texas (Table G. Texas Overall Health Care Costs). The table shows the price-adjusted Medicare reimbursements per enrollee.

Table G Texas overall healthcare costs

Texas overall healthcare costs	
Summary information	
Data Year Used:	2007
Range in Texas (Min-Max):	$5,999–$15,429
Overall in Texas:	$10,889

Source: http://www.countyhealthrankings.org/app/texas/2012/measure/factors/86/data?sort = desc-0 County Health Rankings & Roadmaps program: 2012 Texas Health Rankings Compilation

Healthcare costs were at the highest levels in the southern region of Texas (Figure A7). This region was a heavily Hispanic-populated area with low household median income levels (Figure A6). However, prostate cancer mortality rates were low in this region (Figure A1), and healthcare costs may not have been a contributing factor of mortality rates recorded from this region. However, this may have been due to migration to more specialised cancer care facilities in the eastern region.

5.7.4 Unemployment

The 2012 Texas Health Rankings used the Bureau of Labour Statistics 2010 data to determine the overall 8.2% of unemployment for the state (Table H. Texas Overall Unemployment). The table shows the percent of the population 16 and older that are unemployed and seeking work.

Table H Texas overall unemployment

Texas overall unemployment	
Ranking methodology	
Data year used:	*2010*
Summary Measure:	Health Factors – Social & Economic Factors (Employment)
Weight in Health Factors:	10%
Summary information	
Top US Performers:	5.4% (10th Percentile)
Range in Texas (Min-Max):	4.1–17.9%
Overall in Texas:	8.2%

Source: http://www.countyhealthrankings.org/app/texas/2012/measure/factors/23/map?sort = desc-0 County Health Rankings & Roadmaps program: 2012 Texas Health Rankings Compilation

Areas with high rates of prostate cancer mortality had low rates of unemployment (Figure A4). Unemployment percentages were lowest in the northwestern parts of Texas, specifically within the Panhandle region (Figure A8). Moderate to high levels of unemployment were in the southern to eastern part of Texas, along the Texas state line, and in the Gulf of Mexico region. Prostate cancer mortality rates recorded for southern Texas region may not have been correlated to unemployment, assuming migration has been excluded, namely due to Hispanics showing a low rate for prostate cancer mortality in that region (Figure A4). However, African Americans showed a high concentration in eastern Texas (Figure A3) and may have been correlated to high concentration of prostate cancer mortality rates (Figure A1), namely because of the genetic factor in prostate cancer.

5.7.5 Uninsured adults

The 2012 Texas Health Rankings used data from 2009, provided by the Small Area Health Insurance Estimate (SAHIE), to determine the overall 31% of uninsured adults for Texas (Table I. Texas Overall Uninsured Adults). The table shows the percent of adults less than 65 who do not have health insurance.

Table I Texas overall uninsured adults

Texas overall uninsured adults	
Summary information	
Data year used:	*2009*
Range in Texas (Min-Max):	19–51%
Overall in Texas:	31%

Source: http://www.countyhealthrankings.org/app/texas/2012/measure/factors/3/data?sort = desc-0 County Health Rankings & Roadmaps program: 2012 Texas Health Rankings Compilation

Uninsured adults were highest among southern Texas counties along the southern state border (Figure A9), where there was a high concentration of Hispanics (Figure A4). Additionally, there was a high concentration of uninsured adults in northwestern Texas, specifically the Panhandle, where there was a low percentage of unemployment. It is important to note that this area is agricultural, and employers may not have provided insurance for employees in this region, as employment may have been seasonal. This assumption helps to explain why a low percentage of unemployment region had a high concentration of uninsured adults.

5.8 Hypothesis 2 findings

Hypothesis two of this study posited that other factors potentially contribute to prostate cancer deaths, specifically household median income, healthcare costs, unemployment, and adults who are uninsured, all of which are assumed to show a negative relationship between these variables and AADR, with the exception of household median income, which is assumed to show a positive correlation. Therefore, a Pearson bivariate correlation matrix was conducted to analyse the relationship between these socioeconomic factors and the age-adjusted death rate (AADR) in prostate cancer (Table J).

Table J

Correlations

		AADR	*Household income*	*Healthcare costs*	*Percent of unemployed*	*Percent of uninsured adults*
AADR	Pearson Correlation	1	–0.077	0.058	0.075	–0.002
	2-tailed Sig.		.219	0.360	0.235	0.980
	N	254	254	251	254	254
Household income	Pearson Correlation	–0.077	1	–0.008	–0.277**	–0.705**
	2-tailed Sig.	0.219		0.898	0.000	0.000
	N	254	254	251	254	254
Healthcare costs	Pearson Correlation	0.058	–0.008	1	0.208**	–0.020
	2-tailed Sig.	0.360	0.898		0.001	0.757
	N	251	251	251	251	251
Percent of unemployed	Pearson Correlation	0.075	–0.277**	0.208**	1	0.159*
	2-tailed Sig.	0.235	0.000	0.001		0.011
	N	254	254	251	254	254
Percent of uninsured adults	Pearson Correlation	–0.002	–0.705**	–0.020	0.159*	1
	2-tailed Sig.	0.980	0.000	0.757	0.011	
	N	254	254	251	254	254

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

Table J shows that no statistically significant correlations exist between AADR and the listed variables. However, inferences can be made from the table. An increase in household income showed a very slight decrease in AADR prostate cancer mortality, indicating a weak negative relationship between both variables. Healthcare costs and AADR showed a very weak positive relationship, showing that as healthcare costs increased, there was a very slight increase in AADR. Likewise, as the percentage of unemployment increased, AADR increased slightly but still showed a weak positive relationship. Additionally, as the percentage of uninsured adults increased, AADR decreased slightly, indicating a very weak negative relationship.

5.8.1 Healthcare access and prostate cancer

5.8.1.1 Primary care physicians

Primary care physicians were analysed to help determine the accessibility of healthcare in counties. High concentrations of primary care physicians were in the eastern, central, and northern region of the state. There was especially a very high concentration in the eastern region. In like manner, the high concentration of healthcare access was comparable to the findings of prostate cancer regional concentrations. That is to say, a high concentration of prostate cancer and primary care physician access were regionally comparable.

5.9 Hypothesis 3 findings

Hypothesis 3 of this study posited that there is a negative correlation between healthcare access and prostate cancer. That is to say, greater access to healthcare would indicate a reduction in prostate cancer deaths. This is measured in the form of available primary care physicians in the area, who have legal authority to prescribe PSA testing for patients.

Table K shows a very weak positive correlation between the rate of primary care physicians and AADR. This was a surprising finding since the assumption was that the more available access to healthcare, the greater impact on the mortality rate. However, this correlation does not appear to be statistically significant and would need further review and analysis.

Table K

Correlations			
		AADR	*PCP Rate*
AADR	Pearson Correlation	1	0.108
	2-tailed Sig.		0.105
	N	254	226
PCP Rate	Pearson Correlation	0.108	1
	2-tailed Sig.	0.105	
	N	226	226

6 Discussion

Communities that stand to benefit from the insights acquired in this study are several. First, public health officials can find them useful because they are able to determine the geographic locations of mortalities occurring from prostate cancer in the state of Texas. Second, the study can help the medical community, including hospitals, clinics, and providers, to create interventions that can help further define and determine necessary interventions and treatments to save lives. The medical community can find opportunities to work with public officials at the state, county, or city level, specifically to help increase PSA tests to find prostate cancer in its early stage. Furthermore, it can help the medical communities initiate further investigations into the high mortality rates among the African American population and to further determine if the high mortality rate among Blacks is specific to late PSA testing or other external factors that were not covered in this study. Third, the public in general benefits from this study, especially those among minority populations, because it helps to increase awareness and can drive a positive reaction among the high-risk population to get screened more often. Other groups that may benefit from the study are academic researchers who would like to know the benefits of utilising data science software to explore prostate cancer disease. Additionally, insurance groups may also be interested in the results of the study because it can help them mitigate costly treatment in the future by intervening early.

7 Conclusion

This study documents the value data science programs such as geographical information systems and statistical software have for communities in the public and medical sector. This study assessed concentrations of prostate cancer mortality against social determinants in health to help explain the health disparities of prostate cancer. The study encompassed included race/ethnicity, socioeconomic status, and healthcare access factors, in an attempt to explain the geography of the disease and the reason for its distribution throughout the state of Texas. Two main conclusions are developed from this study. The first is that high concentrations of prostate cancer mortality were found mainly in the eastern and central areas of Texas. In like manner, smaller clusters of high concentrations existed in the West Texas Permian Basin and the Panhandle. This is an unusual paradox because there were large numbers of Hispanics living in these regions, but the study did not make a statistically significant finding of correlation between prostate cancer mortality and Hispanic race/ethnicity. This may be because Hispanics tended to migrate, seeking better healthcare treatment. The Hispanic paradox is important to note and requires further research. A deeper study and comparison of the prostate cancer death rate among Blacks and Whites in these regions may help solve the paradox. The second conclusion is that there was a high mortality rate among Blacks in the eastern part of Texas, despite renowned healthcare facilities such as MD Anderson Cancer Center in the region. Because the percentage of Blacks in this study was found to be statistically significant when analysing prostate cancer mortality rate for Texas, more research is suggested because African Americans had poorer access to prostate cancer treatment and not necessarily a more aggressive form of the disease ("Black Race Not Associated with Worse Prostate Ca Outcomes", 2018). In actuality, race discrimination may have been an important contributor. Benjamins and Whitman (2013), on relationships between healthcare discrimination and outcomes, found that African Americans commonly suffered discrimination in healthcare. One study found that there were notable racial disparities in Texas and that both black and Hispanic males had greater likelihood of dying from prostate cancer than white males (White et al., 2010). Minorities, especially blacks, may not receive adequate treatments for disease. Other opportunities for future research include genetics and prostate cancer, specialised medicine, specific migration assessments, environmental considerations, and political factors, whether at the local, regional, or state level. The social determinants in health explanatory variables that were assessed were socioeconomic and healthcare access domains. They did not find a statistically significant relationship with prostate cancer. New variables, and combinations thereof, from these domains are in order for future research. For example, in lieu of assessing primary care physicians against prostate cancer mortality in general, a deeper study examining insured and non-insured African Americans and prostate cancer mortality may produce a fuller assessment of healthcare access. Additionally, other socioeconomic variables such as occupation, education, wealth, and place of residence can be helpful.

8 Research limitations

The dataset acquired from the Texas Health Rankings, specifically the measure for primary care physicians, had missing values. This resulted in an incomplete choropleth map, showing holes in the map for several counties. Additionally, for demographics measure, the percent of whites was missing from the dataset. Although this was a limitation, and perhaps worthy of mention in the study at a more detailed level, this research study focused on three minority groups, which excluded whites. This exclusion was based on the fact that whites already have lower rates than African Americans and were not a variable of interest in the study. One other limitation was that the unemployment data comprised teenagers and younger adults. Employment tends to be tied to other factors such as the ability to afford healthcare and employer-provided insurance. Therefore, unemployment data may misrepresent the relationship between employment and prostate cancer mortality. Finally, AADR was a limitation for counties with small populations, especially in rural areas. For example, two deaths from Hispanics that occur from prostate cancer could show a 50% mortality rate if there were only four Hispanics in the area.

References

Attard, G., Parker, C., Eeles, R.A., Schröder, F., Tomlins, S.A., Tannock, I. and Bono, J.S. (2016) 'Prostate cancer', *The Lancet*, Vol. 387, No. 10013, pp.70–82, doi: 10.1016/s0140-6736(14)61947-4.

Benjamins, M.R. and Whitman, S. (2013) 'Relationships between discrimination in health care and health care outcomes among four race/ethnic groups', *Journal of Behavioral Medicine*, Vol. 37, No. 3, pp.402–413, doi: 10.1007/s10865-013-9496-7.

Braveman, P. (2014) 'What are health disparities and health equity? we need to be clear', *Public Health Reports*, Vol. 129, No. 1, Suppl. 2, pp.5–8, doi: 10.1177/00333549141291s203.

Bui, T.Q. and Pham, H.M. (2016) 'Web-based GIS for spatial pattern detection: application to malaria incidence in Vietnam', *SpringerPlus*, Vol. 5, No. 1, doi: 10.1186/s40064-016-2518-5.

Cheng, I., Witte, J.S., McClure, L.A., Shema, S.J., Cockburn, M.G., John, E.M. and Clarke, C.A. (2009) 'Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California', *Cancer Causes and Control*, Vol. 20, No. 8, pp.1431–1440, doi: 10.1007/s10552-009-9369-0.

Gregorio, D.I., Kulldorff, M., Sheehan, T. and Samociuk, H. (2004) 'Geographic distribution of prostate cancer incidence in the era of PSA testing, Connecticut, 1984 to 1998', *Urology*, Vol. *63, No.* 1, pp.78–82, doi: 10.1016/j. urology.2003.08.008.

Kim, J.H., Sun, H.Y., Kim, H.J., Ko, Y.M., Chun, D. and Park, J.Y. (2017) 'Does uneven geographic distribution of urologists effect bladder and prostate cancers mortality? national health insurance data in Korea from 2007-2011', *Oncotarget*, Vol. 8, No. 39, doi: 10.18632/oncotarget.18036.

Klassen, A.C. and Platz, E.A. (2006) 'What can geography tell us about prostate cancer?', *American Journal of Preventive Medicine*, Vol. 30, No. 2, pp.S7–S15, doi: 10.1016/j. amepre.2005.09.004.

Kumar, S., Singh, R., Malik, S., Manne, U. and Mishra, M. (2018) 'Prostate cancer health disparities: an immuno-biological perspective', *Cancer Letters*, Vol. 414, pp.153–165, doi: 10.1016/j. canlet.2017.11.011.

LaVeist, T.A. and Pierre, G. (2014) 'Integrating the 3Ds–social determinants, health disparities, and health-care workforce diversity', *Public Health Reports*, Vol. 129, No. 1, Suppl. 12, pp.9–14, doi: 10.1177/00333549141291s204.

Major, J.M., Oliver, M.N., Doubeni, C.A., Hollenbeck, A.R., Graubard, B.I. and Sinha, R. (2012) 'Socioeconomic status, healthcare density, and risk of prostate cancer among African American and Caucasian men in a large prospective study', *Cancer Causes and Control*, Vol. 23, No. 7, pp.1185–1191, doi: 10.1007/s10552-012-9988-8.

Pearce, N., Foliaki, S., Sporle, A. and Cunningham, C. (2004) 'Genetics, race, ethnicity, and health', *BMJ*, Vol. 328, No. 7447, pp.1070–1072, doi: 10.1136/bmj.328.7447.1070.

Penson, D.F. and Chan, J.M. (2007) 'Prostate cancer', *The Journal of Urology*, Vol. 177, No. 6, pp.2020–2029, doi: 10.1016/j. juro.2007.01.121.

Rebbeck, T.R. (2017) 'Prostate cancer genetics: variation by race, ethnicity, and geography', *Seminars in Radiation Oncology*, Vol. 27, No. 1, pp.3–10, doi: 10.1016/j. semradonc. 2016.08.002.

Sherman, R.L., Henry, K.A., Tannenbaum, S.L., Feaster, D.J., Kobetz, E. and Lee, D.J. (2014) 'Applying spatial analysis tools in public health: an example using SaTScan to detect geographic targets for colorectal cancer screening interventions', *Preventing Chronic Disease*, Vol. 11, doi: 10.5888/pcd11.130264.

Shulan, M., Gao, K. and Moore, C.D. (2015) 'Thirty-day all-cause hospital readmissions – racial and income disparities and risk factors in a veterans integrated healthcare network', International Journal of Healthcare Technology and Management, Vol. 15, No. 2, p.112, doi: 10.1504/ijhtm.2015.074539.

White, A., Coker, A.L., Du, X.L., Eggleston, K.S. and Williams, M. (2010) 'Racial/ethnic disparities in survival among men diagnosed with prostate cancer in Texas', *Cancer*, Vol. 117, No. 5, pp.1080–1088, doi: 10.1002/cncr.25671.

Wilkinson, R.G. and Marmot, M.G. (2003) *Social Determinants of Health: The Solid Facts*, World Health Organization, Geneva, Switzerland.

Websites

About Social Determinants of Health, [online] https://www.who.int/social_determinants/sdh_definition/en/ (Accessed 9 September, 2019).

Black Race Not Associated with Worse Prostate Ca Outcomes. [online] https://www.medpagetoday.com/meetingcoverage/astro/75896 (Accessed 24 October, 2018).

Centers for Disease Control and Prevention and National Cancer Institute, U.S. Cancer Statistics Working Group (2018) *U.S. Cancer Statistics Data Visualizations Tool*, https://gis.cdc.gov/Cancer/USCS/DataViz.html (Accessed 18 October, 2019).

County Health Rankings & Roadmaps. Health Care Costs. [online] http://www.countyhealthrankings.org/app/texas/2012/measure/ (Accessed 18 October, 2019).

Health Disparities [online] https://medlineplus.gov/healthdisparities.html (Accessed 18 October, 2019).

Disparities [online] https://www.healthypeople.gov/2020/about/foundation-health-measures/Disparities (Accessed 18 October, 2009).

Disparities in Health and Health Care: Five Key Questions and Answers. [online] https://www.kff.org/disparities-policy/issue-brief/disparities-in-health-and-health-care-five-key-questions-and-answers (Accessed 18 October, 2009).

Overview—Vitalnet Data Warehouse Software. [online] https://www.ehdp.com/vitalnet/overview.htm (Accessed 18 October, 2019).

Texas Department of State Health Services. "Cancer incidence leading sites 2011-2015". [online] dataset, *Age-Adjusted Cancer Incidence in Texas, 2011-2015 Leading Sites by Sex and Race/Ethnicity*. https://www.dshs.texas.gov/tcr/data/incidence-and-mortality.aspx (Accessed 15 October, 2018).

Texas Department of State Health Services. "Cancer mortality leading causes 2011-2015". [online] dataset, *Age-Adjusted Cancer Mortality in Texas, 2012-2-16 by Sex and Race/Ethnicity*. https://www.dshs.texas.gov/tcr/data/incidence-and-mortality.aspx (Accessed 15 October, 2018).

Texas Department of State Health Services. *Welcome to the Texas Cancer Registry*. [online] https://www.dshs.texas.gov/tcr/home.aspx (Accessed 15 October, 2018).

Social Determinants of Health: What Medical Students Need to Know. [online] https://www.ama-assn.org/delivering-care/patient-support-advocacy/social-determinants-health-what-medical-students-need-know?matchtype = e&network = g&device = c&adposition = 1t3&keyword = social%20determinants%20of%20health&utm_source = google&utm_medium = ppc&utm_campaign = pe-digital-ads-membership&utm_effort = GG0001&gclid = CjwKCAjwnMTqBRAzEiwAEF3ndo4yIcrpKp0qwljLmkfoTzcdUQEE0ZAY_1n5_IxFcNNCwf0yqvTfzhoCL9UQAvD_BwE (Accessed 18 October, 2019).

Appendix

Figure A1 Age-adjusted prostate cancer mortality in Texas counties, 1999 to 2009

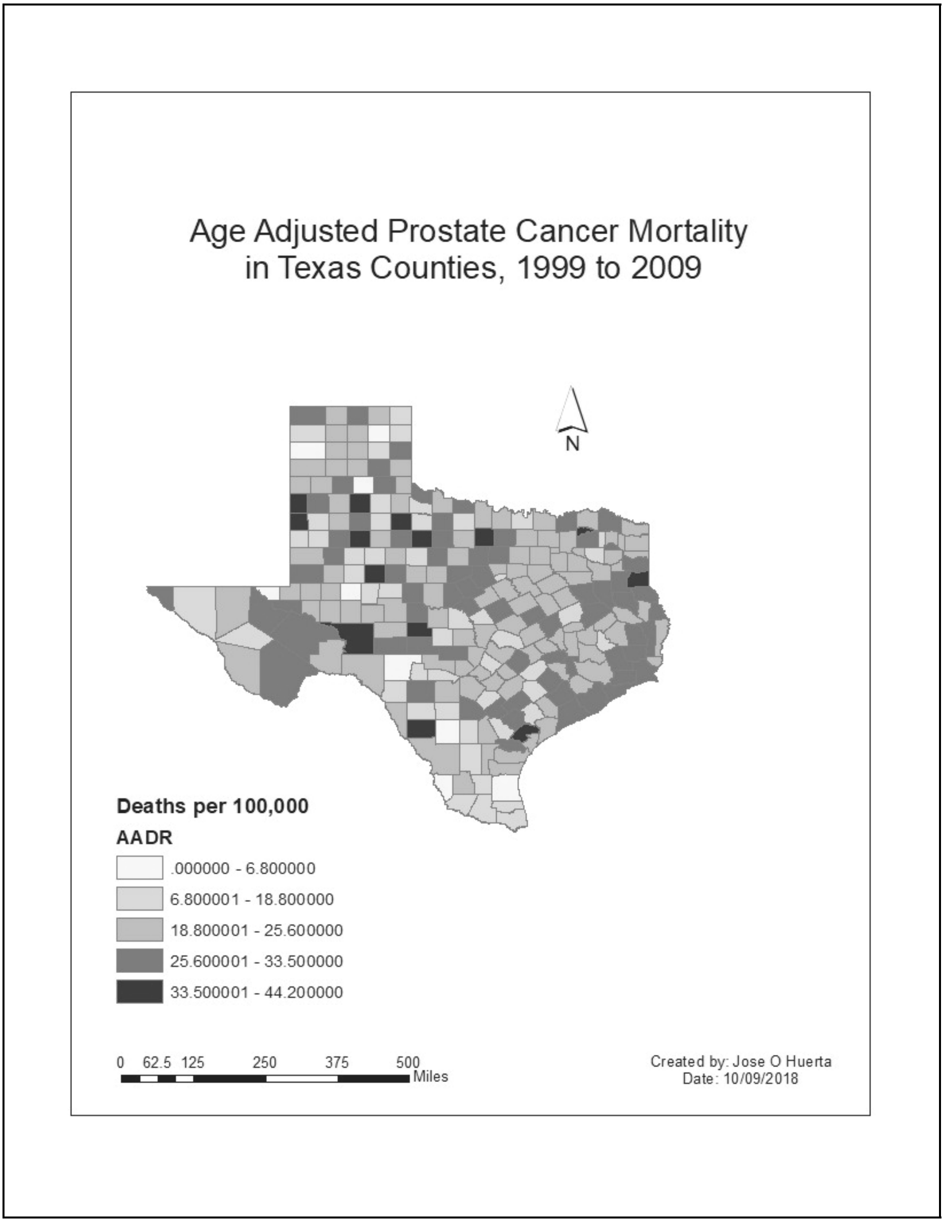


Figure A2 Geographic distribution of 2009 Texas U.S. Census (Population density)

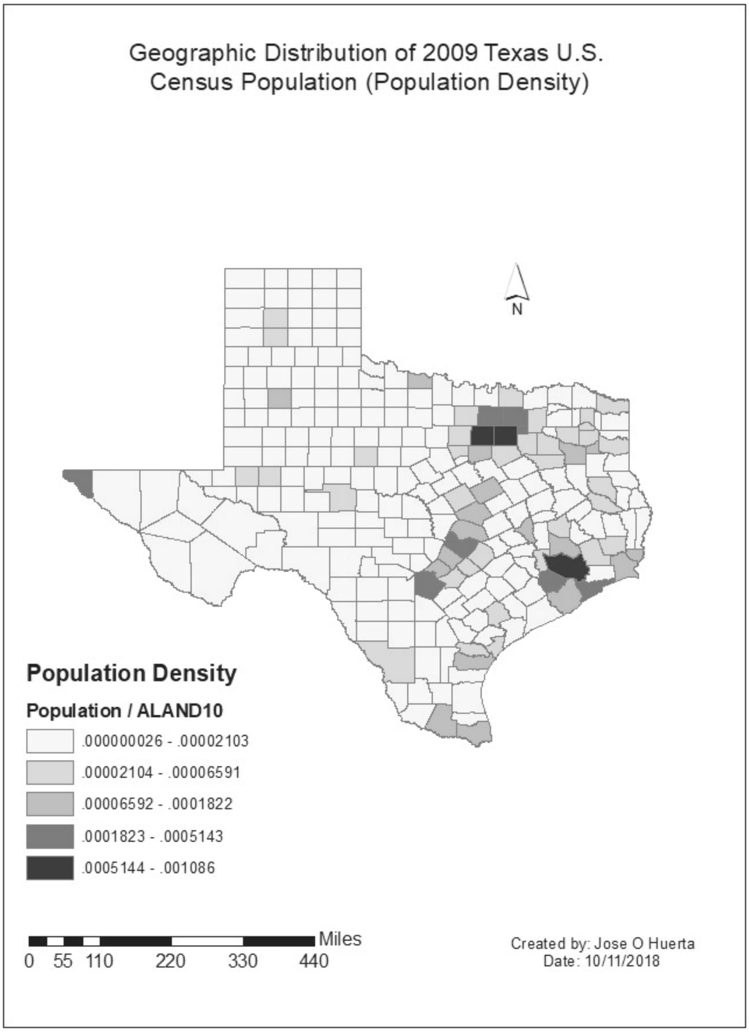


Figure A3 Geographic distribution of 2009 Texas US Census of African Americans

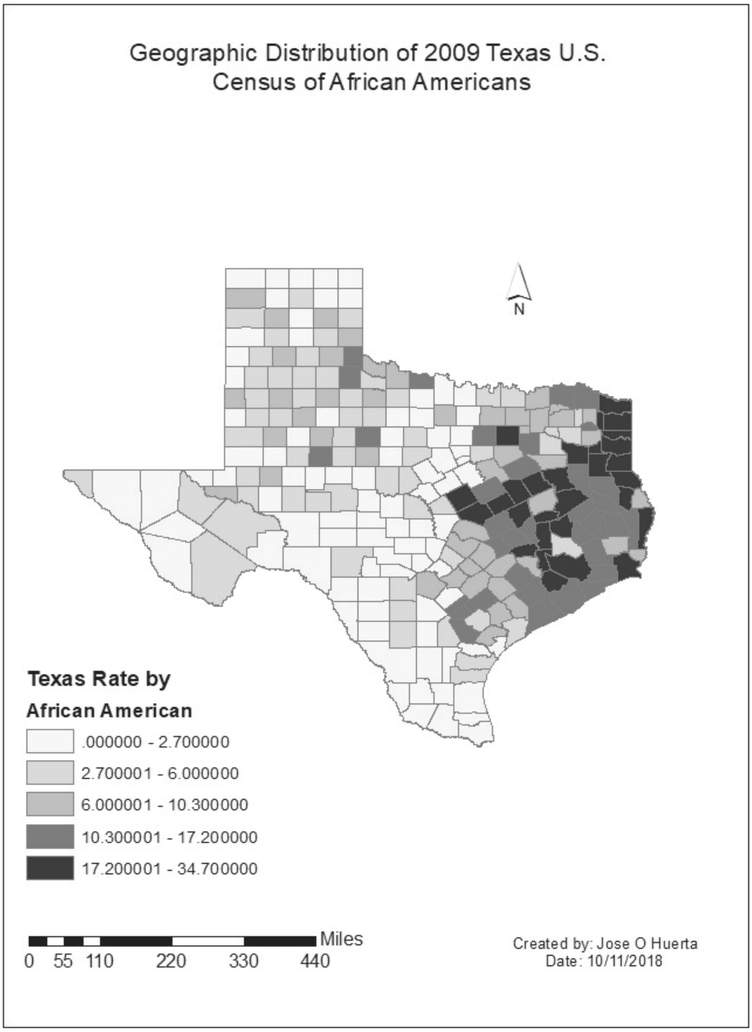


Figure A4 Geographic distribution of 2009 Texas U.S. Census of hispanics

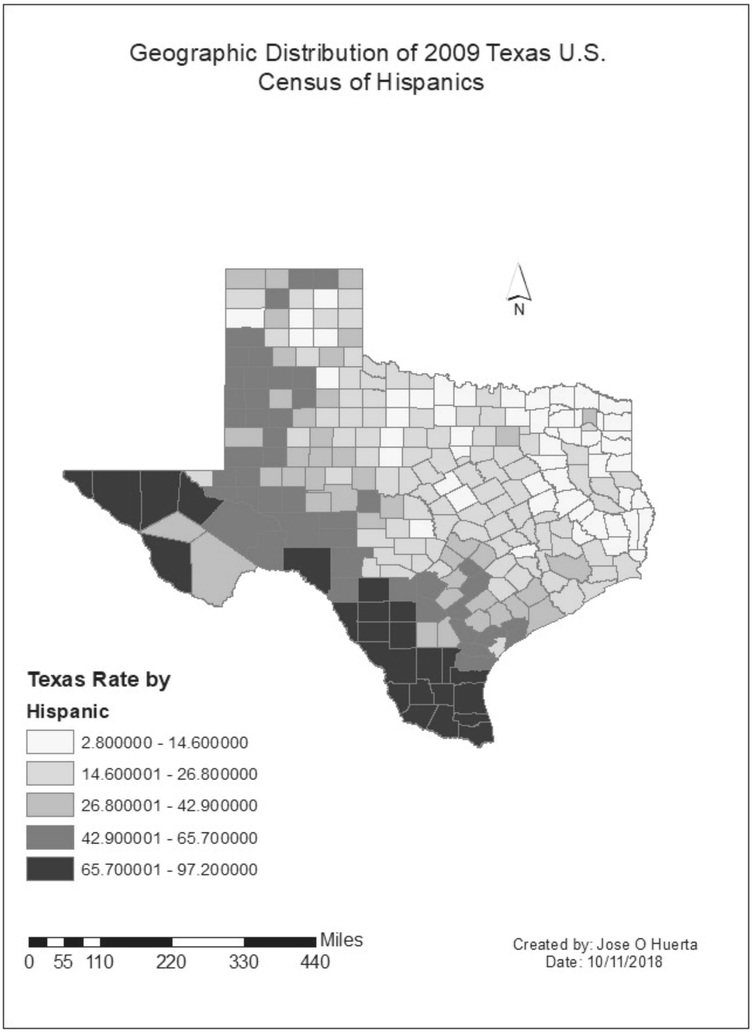


Figure A5 Geographic Distribution of 2009 Texas U.S. Census of Other Races Combined: American Indian, Alaskan Native, Asian, Native Hawaiian, Other Pacific islanders

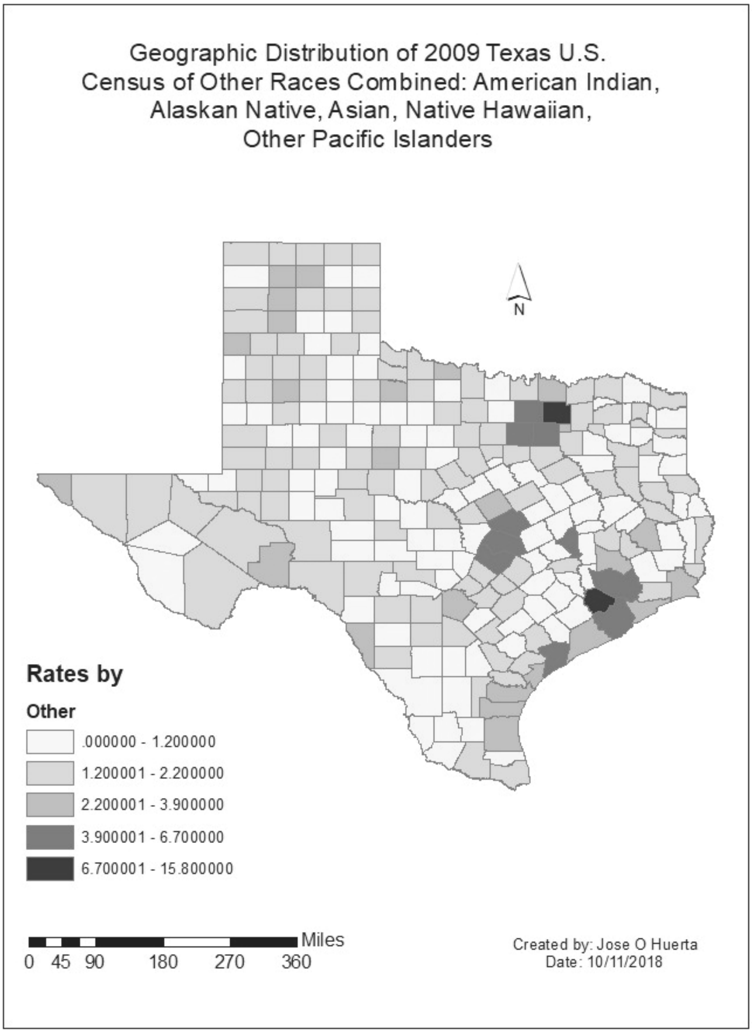


Figure A6 Geographic distribution of Texas household median income

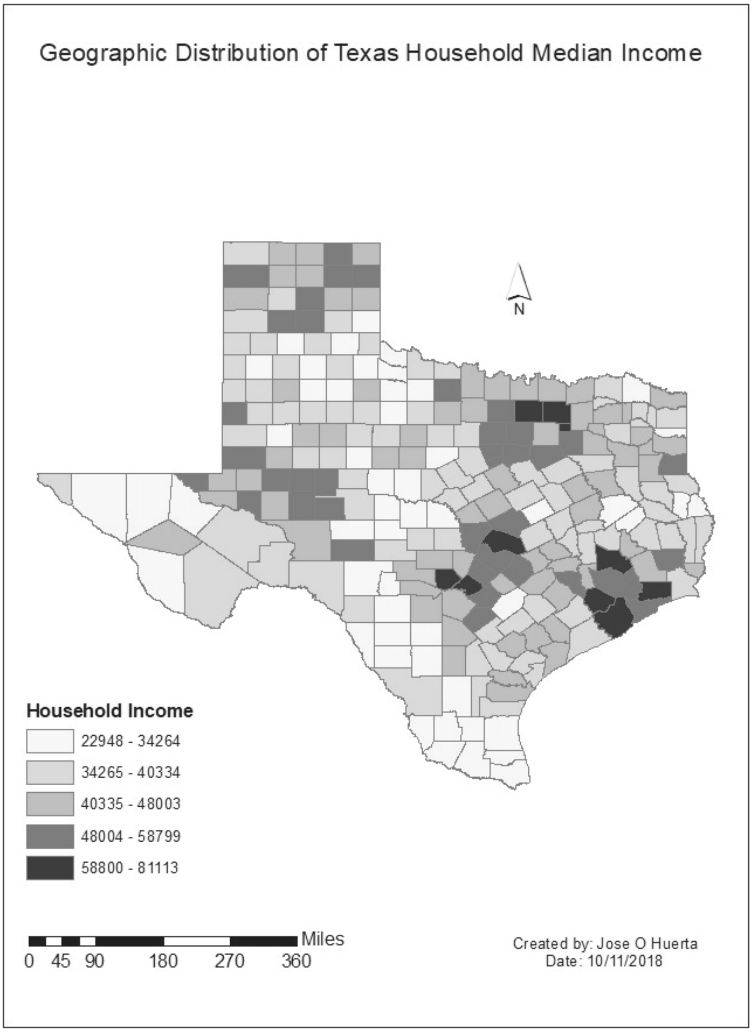


Figure A7 Geographic distribution of Texas healthcare costs

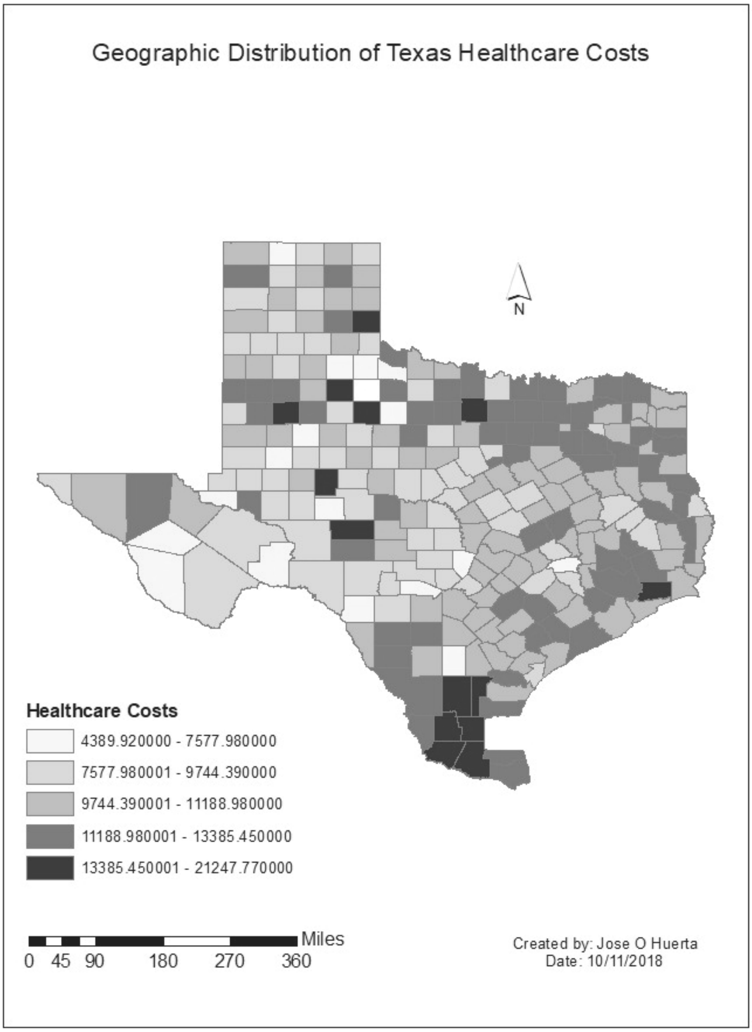


Figure A8 Geographic distribution of unemployment in Texas

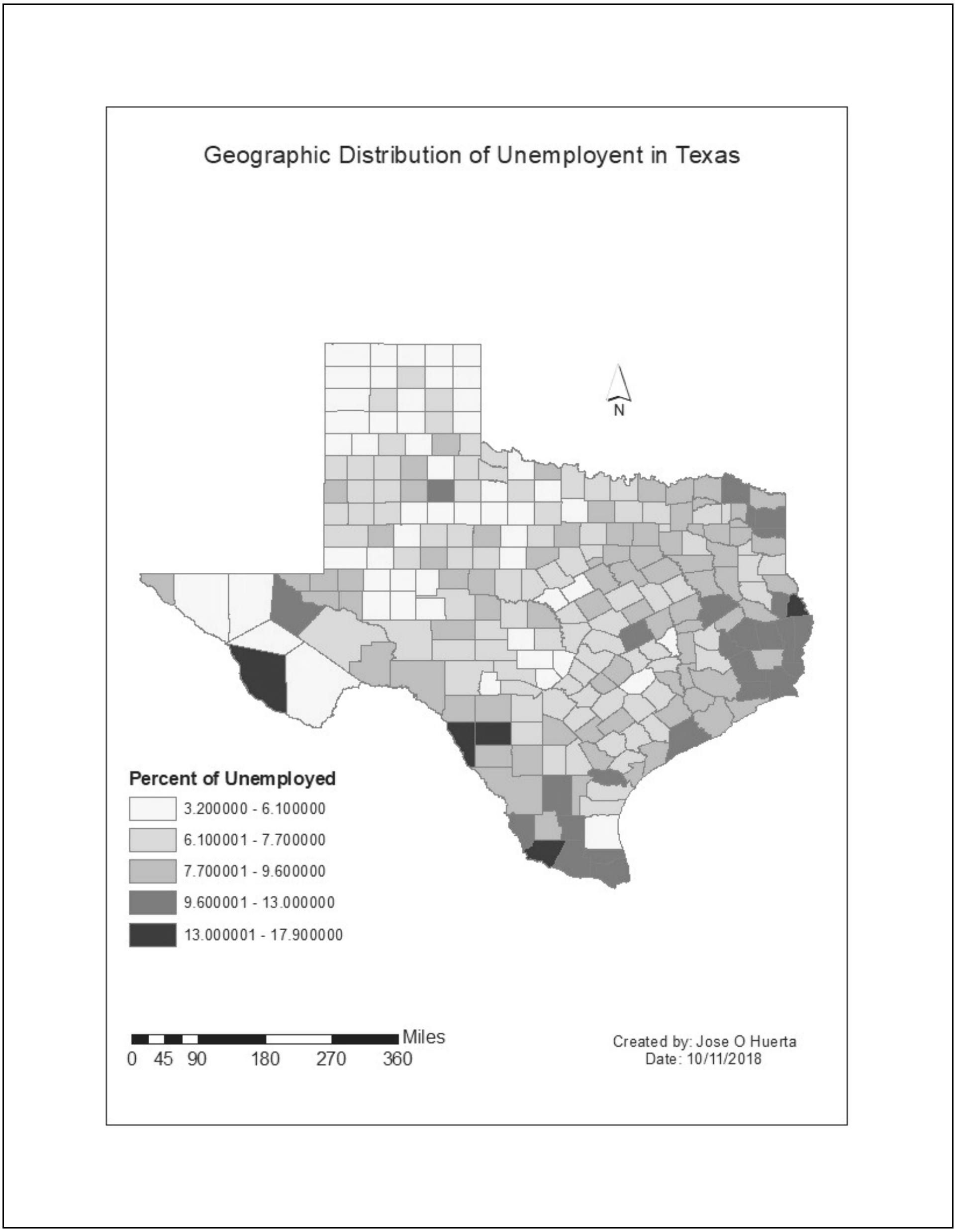


Figure A9 Geographic distribution of uninsured adults in Texas

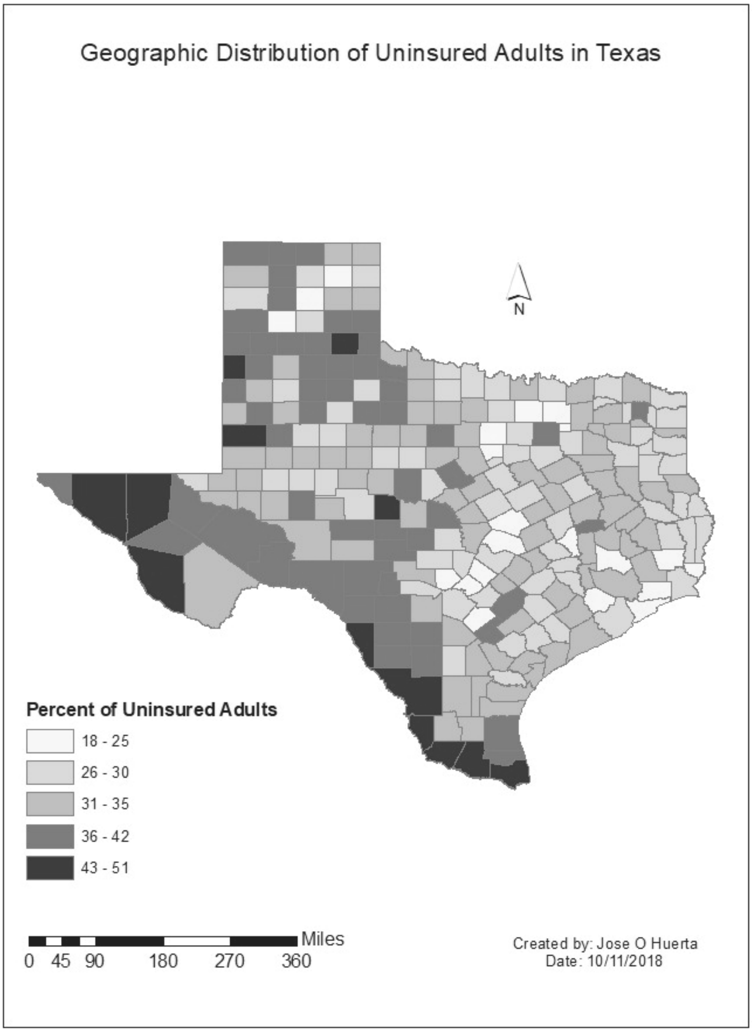


Figure A10 Geographic distribution of primary care physicians in Texas

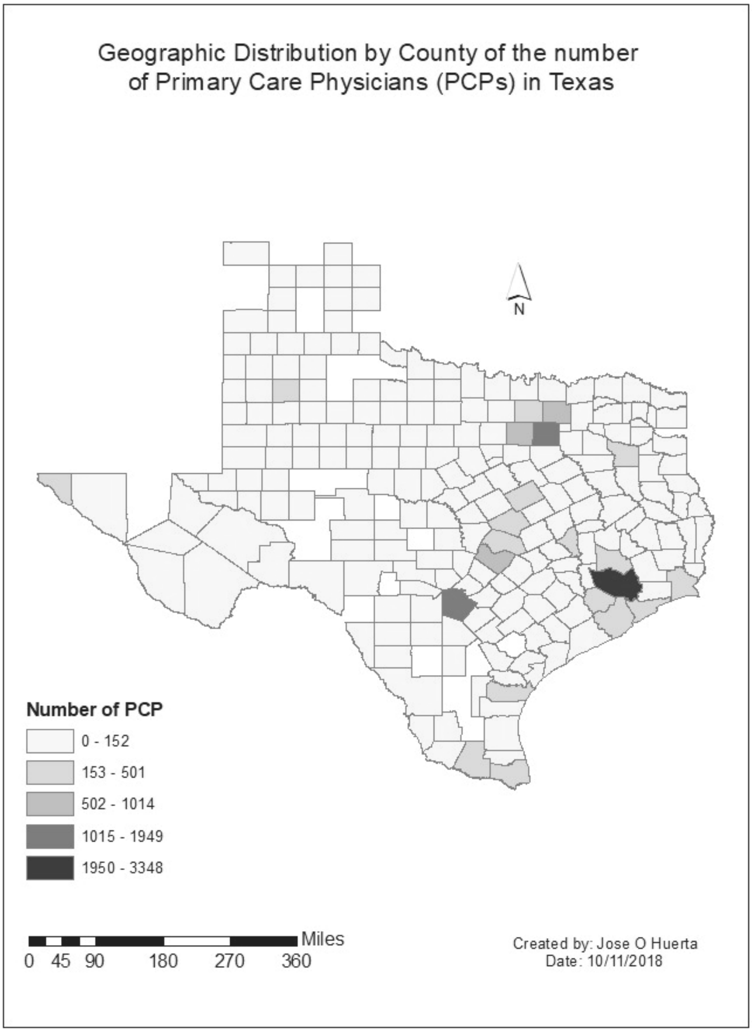


Figure A11 Age-adjusted prostate cancer mortality in Texas counties, 1999 to 2004

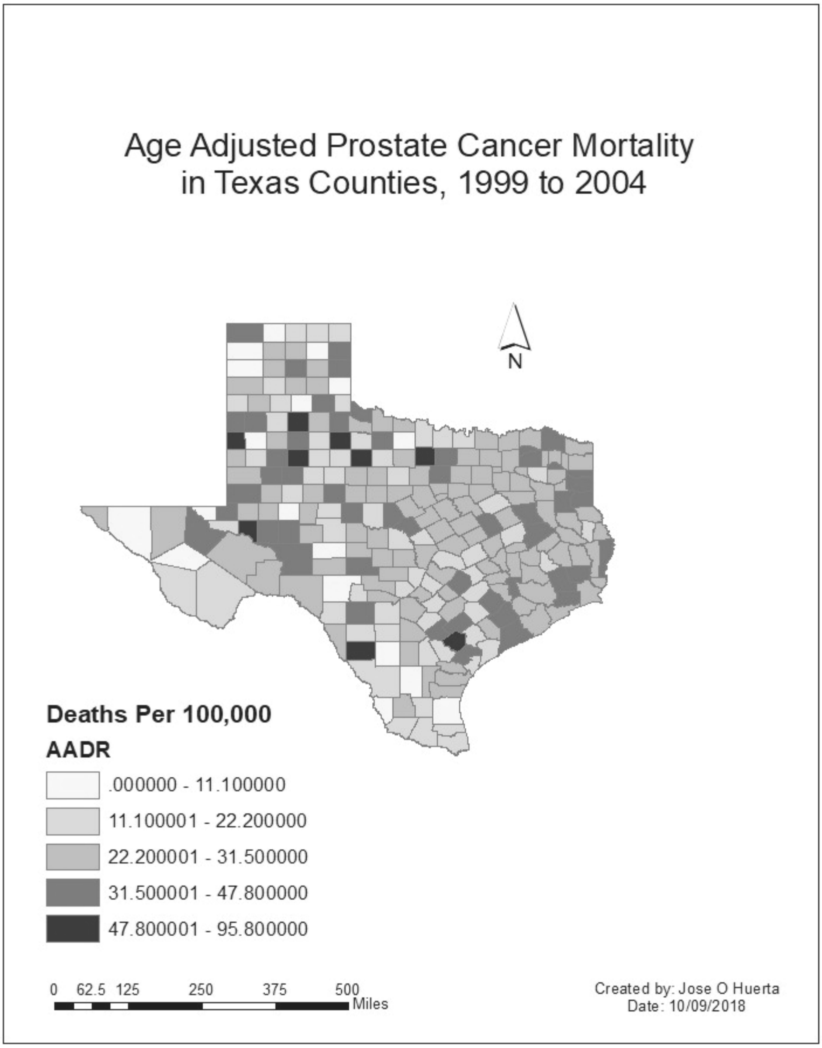


Figure A12 Age-adjusted prostate cancer mortality in Texas counties, 2005 to 2009

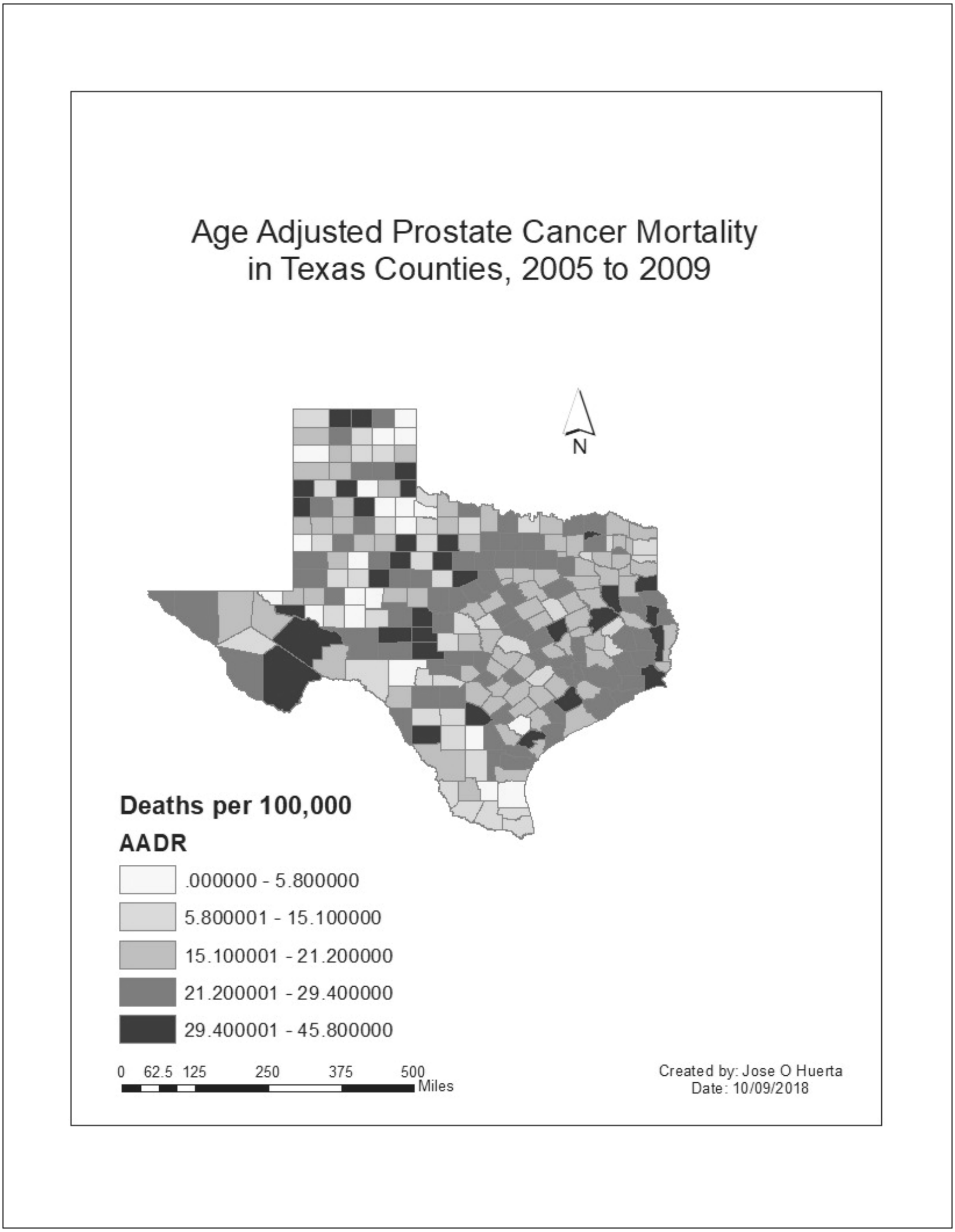


Figure A13 Age-adjusted prostate cancer mortality in Texas counties, change map