

RACIAL, SOCIOECONOMIC, AND GEOGRAPHIC DISPARITIES
OF FEMALE BREAST CANCER IN TEXAS

DISSERTATION

Presented to the Graduate Council
of Texas State University-San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Doctor of PHILOSOPHY

by

Nancy Tian, B.S., M.S.

San Marcos, Texas

August 2010

RACIAL, SOCIOECONOMIC, AND GEOGRAPHIC DISPARITIES
OF FEMALE BREAST CANCER IN TEXAS

Committee Members Approved:

F. Benjamin Zhan, Chair

Yongmei Lu

Jeff G. Wilson

Kenneth Chu

Approved:

J. Michael Willoughby

Dean of Graduate College

COPYRIGHT

by

Nancy Tian

2010

ACKNOWLEDGEMENTS

I am indebted to my outstanding dissertation committee which includes Drs, F. Benjamin Zhan (Chair), Yongmei Lu, Jeff G. Wilson, and Kenneth Chu for their insightful guidance, constant encouragement, and endless support. My advisor, Dr. Zhan, provided me tremendous support and valuable insights for my dissertation research. Dr. Lu was most helpful in providing me research materials that have inspired me with great research ideas. Dr. Wilson has been very generous to assist me with editing of my work which has resulted in a couple of publications. Dr. Chu was munificent with time and sharing of cutting-edge research on breast cancer.

In addition, I would like to extend my greatest gratitude to my wonderful mentor, Dr. Pierre Goovaerts (BioMedware), for his insightful help in the methodological development and interpretation of results. BioMedware gave me free access to the STIS software. Many thanks also go to Drs. Kyong Chee, Edwin Chow, Niem Huynh, Charles Yorke, and Ms. Jonna Beck for helping improve my writing skills. I am grateful to my mother and deceased father for raising me and providing me the opportunity to complete higher education. My husband, Les Stewart, is such a considerate person for taking care of housework and understanding my stressful work. The love and dedication of my in-laws have always given me support in my endeavors.

This manuscript was submitted on July 19, 2010.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
ABSTRACT	xi
CHAPTER	
I. INTRODUCTION	1
1.1 Background on Breast Cancer	1
1.2 Definitions of Disparity	3
1.3 Rationale of the Study.....	5
1.4 Research Questions and Hypothesis	8
1.5 Significance.....	9
1.6 Breast Cancer Disparities Research and the Disciplines of Geography	11
1.7 Outline of the Dissertation	13
II. LITERATURE REVIEW	14
2.1 Disparities in Health (not including cancer).....	15
2.1.1 Racial Disparities.....	15
2.1.2 Socioeconomic Disparities.....	17

2.1.3 Geographic Disparities.....	19
2.2 Disparities in Cancer (not including breast cancer).....	21
2.2.1 Racial Disparities.....	22
2.2.2 Socioeconomic Disparities.....	23
2.2.3 Geographic Disparities.....	25
2.3 Disparities in Breast Cancer.....	26
2.3.1 Racial Disparities.....	27
2.3.2 Socioeconomic Disparities.....	32
2.3.3 Geographic Disparities.....	36
2.4 Current Interventions.....	39
2.5 Current Limitations.....	40
2.5.1 Race Related Issues.....	40
2.5.2 Methodological Issues.....	41
2.5.3 Socioeconomic Status Measurement Issues.....	42
2.6 Conclusions.....	43
III. DATA AND METHODS.....	45
3.1 Data.....	45
3.1.1 Geography and Population.....	45
3.1.2 Data of Breast Cancer Incidence.....	47
3.1.3 Data of Breast Cancer Mortality.....	49
3.2 Methods.....	50
3.2.1 Disparity Measurement Approach.....	50
3.2.2 Regression Methods.....	54

3.2.3 Local Indicator of Spatial Autocorrelation	58
3.2.4 Factor Analysis	59
IV. SIGNIFICICANCE TESTS AND SPATIAL RELAIONSHIP OF RACIAL DISPARITIES IN LATE-STAGE DIAGNOSIS AND BREAST CANCER MORTALITY	61
4.1 Significance Tests of Racial Disparities in Late-stage Diagnosis of Breast Cancer	62
4.1.1 Late-stage Diagnosis Rates of Breast Cancer	62
4.1.2 Racial Disparities of Breast Cancer Late-stage Diagnosis among African-American Women.....	66
4.1.3 Racial Disparities of Breast Cancer Late-stage Diagnosis among Hispanic Women	70
4.1.4 Discussion on Racial Disparities of Breast Cancer Late-stage Diagnosis	72
4.2 Significance Tests of Racial Disparities in Breast Cancer Mortality	73
4.2.1 Breast Cancer Mortality Rates	73
4.2.2 Racial Disparities of Breast Cancer Mortality Rates among African-American Women.....	76
4.2.3 Racial Disparities of Breast Cancer Mortality Rates among Hispanic Women.....	79
4.2. 4 Discussion on Racial Disparities of Breast Cancer Mortality	81
4.3 Spatial Relationship of Racial Disparities in Late-stage Diagnosis and Breast Cancer Mortality	83
4.3.1 Visualization of Spatial Locations of Significant Racial Disparities	83
4.3.2 Descriptive Statistics of Racial Disparities in Late-stage Diagnosis and Breast Cancer Mortality	88
4.3.3 Linear Regressions, Spatial Autocorrelation Diagnosis, and Geographically Weighted Regression.....	93

4.3.4 Logistic Regression, Spatial Autocorrelation Diagnosis, and and Geographically Weighted Regressions	105
4.3.5 Discussion of Spatial Relationships of Racial Disparities of Late-stage Diagnosis and Breast Cancer Mortality	106
V. RACIAL DISPARITIES OF BREAST CANCER MORTALITY: RISK FACTORS	109
5.1 Descriptive Statistics of Demographic and Socioeconomic Factors	110
5.2 Spatial Accessibility of Mammography Facilities	116
5.3 Results of Factor Analysis	123
5.4 Results of Multivariate Logistic Regression Analysis among African- American Women	128
5.5 Results of Multivariate Logistic Regression Analysis among Hispanic Women	132
5.6 Conclusions	136
VI. CONCLUSIONS, INTERVENTIONS, AND FUTURE WORK	141
6.1 Conclusions	141
6.2 Intervention Programs	146
6.3 Contributions	148
6.4 Limitations and Future Work	149
APPENDIX A. SUMMARY OF SELECTED PUBLICATIONS	153
REFERENCES	161

LIST OF TABLES

Table	Page
Table 3.1 Number and proportion of breast cancer incidence cases by race and stage	49
Table 3.2 Geocoding results among race/ethnicity groups	50
Table 4.1 Number of census tracts with non-zero population and non-zero cases of breast cancer late-stage diagnosis by race	63
Table 4.2 Number of census tracts with significant racial disparities in breast cancer late-stage diagnosis for African-American women	67
Table 4.3 Number of census tracts with significant racial disparities in breast cancer late-stage diagnosis for Hispanic women	70
Table 4.4 Number of census tracts with non-zero population and non-zero mortality of breast cancer by race	74
Table 4.5 Number of census tracts with significant racial disparities in breast cancer mortality for African-American women using RD and RR measurements	77
Table 4.6 Number of census tracts with significant racial disparities in breast cancer mortality for Hispanic women using RD and RR measurements	79
Table 4.7 Number of census tracts with significant racial disparities in late-stage diagnosis and breast cancer mortality using RD and RR measurements	84
Table 4.8 Results of the logistic regression: odds ratios (ORs) and <i>p</i> values with the significance of racial disparities in late-stage diagnosis/ breast cancer mortality (dependent variable) and SES (independent variable) using RD measurement.....	85
Table 4.9 Descriptive statistics of RD and RR racial disparities in late-stage diagnosis and breast cancer mortality for both African-American and Hispanic women	89
Table 4.10 Linear regression results on normalized RD/RR statistics with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis for both African-American and Hispanic women	94

Table 4.11 Moran’s <i>I</i> results on residuals of the linear regression analysis with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis for both African-American and Hispanic women	95
Table 4.12 Results of logistic regression with racial disparities of breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable) using RD measurement for African-American and Hispanic women	106
Table 5.1 Descriptive statistics of individual demographic variables	111
Table 5.2 Descriptive statistics of individual socioeconomic variables	114
Table 5.3 Descriptive statistics of lack of spatial accessibility to mammography facilities	118
Table 5.4 Variables and communalities in the factor analysis.....	124
Table 5.5 The variance accounted for by the successive components in the factor analysis.....	125
Table 5.6 Rotated component matrix in the factor analysis.....	127
Table 5.7 Component score coefficient matrix in the factor analysis.....	128
Table 5.8 Logistic regression results for African-American women: racial disparities in breast cancer mortality as dependent variable (RD) and racial disparities in late-stage diagnosis (RD), demographic characteristics, socioeconomic status, and lack of spatial accessibility as independent variables	131
Table 5.9 Correlation coefficients of the logistic regression for African-American women	131
Table 5.10 Logistic regression results for Hispanic women: racial disparities in breast cancer mortality as dependent variable (RD) and racial disparities in late-stage diagnosis (RD), demographic characteristics, socioeconomic status, and lack of spatial accessibility as independent variables	135
Table 5.11 Correlation coefficients of logistic regression model for Hispanic women	135

LIST OF FIGURES

Figure	Page
Figure 3.1 Census tracts in Texas	47
Figure 4.1 Late-stage specific incidence rates per 100,000 (a) and proportion of late-stage cases (b) for breast cancer by race and age group	65
Figure 4.2 African-American breast cancer late-stage diagnosis: significant racial disparities according to the RD (a) and RR (b) statistics.....	69
Figure 4.3 Hispanic breast cancer late-stage diagnosis: significant racial disparities according to the RD (a) and RR (b) statistics.....	71
Figure 4.4 Age-specific mortality rates per 100,000 (a) and proportion of mortality cases (b) for breast cancer by race and age group	75
Figure 4.5 African-American breast cancer mortality: significant racial disparities according to the RD (a) and RR (b) statistics.....	78
Figure 4.6 Hispanic breast cancer mortality: significant racial disparities according to the RD (a) and RR (b) statistics.....	80
Figure 4.7 Geographic distributions of census tracts with significant racial disparities of late-stage diagnosis and breast cancer mortality (RD) for both African-American (a) and Hispanic women (b).....	87
Figure 4.8 African-American women: histograms and box plots of RD (a and b) and RR (c and d) in late-stage diagnosis and breast cancer mortality	91
Figure 4.9 Hispanic women: histograms and box plots of RD (a and b) and RR (c and d) in late-stage diagnosis and breast cancer mortality	92
Figure 4.10 LISA statistical results on the residuals of the linear regression analysis with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis: RD (a) and RR (b) for African-American women	97

Figure 4.11 LISA statistical results on the residuals of the linear regression analysis with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis: RD (a) and RR (b) for Hispanic women	98
Figure 4.12 Standardized correlate coefficients results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) for African-American women	101
Figure 4.13 Local R square results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) for African-American women.....	102
Figure 4.14 Standardized correlate coefficients results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) for Hispanic women.....	103
Figure 4.15 Local R square results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) for Hispanic women.....	104
Figure 5.1 Histograms and box plots of demographic variables: percentage of African-Americans (a), percentage of Hispanics (b), percentage of minority (c), and population density (d)	112
Figure 5.2 Histograms and box plots of socioeconomic variables: percentage of rural population (a), percentage of population with less than a college education (b), percentage of population under unemployment (c), percentage of population under the poverty line (d), median household income (e)	115
Figure 5.3 Spatial distribution, histogram and box plot of mammography facility density per 1,000 females	120
Figure 5.4 Spatial distributions, histograms and box plots: (a) travel distance to the closest facility, and (b) average travel distance to the five closest facilities	121
Figure 5.5 Spatial distributions, histograms and box plots: (a) travel time to the closest facility, and (b) average travel time to the five closest facilities	122
Figure 5.6 Scree plot of engenvalues by the component number in the factor analysis	125

ABSTRACT

**RACIAL, SOCIOECONOMIC, AND GEOGRAPHIC DISPARITIES
OF FEMALE BREAST CANCER IN TEXAS**

by

Nancy Tian, B.S., M.S.

Texas State University-San Marcos

August 2010

SUPERVISING PROFESSOR: F. BENJAMIN ZHAN

Breast cancer, as the most common cancer among women, has been reported to display remarkable health disparities in the continuum of late-stage diagnosis, utilization of mammography, treatment options, as well as survival and mortality. These inequalities are experienced by different subpopulations and related to a complex set of factors, such

as education, socioeconomic, race, geography, unequal access to health resources, and health-related policies. Existing knowledge about these factors and mechanisms causing these disparities in breast cancer outcomes is limited. Few studies have examined how racial disparities in breast cancer vary across geographic regions, which is key information for any attempt to allocate limited health resources more effectively and efficiently. To date, research on racial and socioeconomic disparities of breast cancer mainly proceeds by combining epidemiological data and investigating risk factors for the population under study. By doing so, one overlooks the geographical variations of breast cancer outcomes, a piece of information that is critical to target regions in intervention programs. Investigating racial disparities across regions can provide useful insights and reveal unknown risk factors for breast cancer, thereby helping health-policy makers to improve the overall health outcome of breast cancer among women.

Within the framework of Geographic Information Systems (GIS), this research conducted spatial and statistical analysis to identify the census tracts that displayed significant racial disparities in late-stage diagnosis and breast cancer mortality for both African-American and Hispanic women compared with their non-Hispanic white counterparts in Texas from 1995-2005. These disparities were measured in terms of rate difference (RD) and rate ratio (RR) and accounted for the population size of each census tract. The significance of these disparities was evaluated statistically and the results were corrected for multiple testing using the false discovery rate approach. For African-American women with the RD measurement, 278 and 188 census tracts displayed significant racial disparities for breast cancer mortality and late-stage diagnosis respectively in the 4,388 census tracts in Texas. These figures were larger for Hispanic

women, with 328 and 266 census tracts respectively. Fewer Census tracts tested significant for the RR measurement. Most of the census tracts with significant racial disparities were located in the metropolitan areas of Houston, Dallas, and Austin-San Antonio for African-American women. Hispanics were also found to have significant racial disparities in the Southwest border of Texas. Logistic regression between the significance of the RD statistic for the two types of health outcomes indicated that a census tract with significant racial disparities for late-stage diagnosis was 30 times more likely to test significant for racial disparities in breast cancer mortality.

Logistic regression was also utilized to investigate the spatial connection of significant racial disparities in late-stage diagnosis and breast cancer mortality. The socioeconomic status (SES) was categorized into groups of low, middle, and high based on the percentage of population living under the federal poverty line. For the two minority groups, low and middle SES census tracts were more likely to report significant racial disparities for both health outcomes. About 40% of the census tracts with significant racial disparities for breast cancer mortality also displayed significant disparities for late-stage diagnosis. Linear regression was then used to quantify the relationships between the magnitude of racial disparities in mortality and late-stage diagnosis for breast cancer. The correlation coefficient was 0.23 for the RD measurement and 0.45 for the RR measurement for both minority groups. Moran's *I*, however, indicated the presence of spatial autocorrelation in the regression residuals, which might reflect the non-stationarity of the regression coefficients and/or the existence of unknown spatial factors. Therefore, the regression models were weighted geographically to account for the spatial variations of observations.

Furthermore, potential risk factors such as demographic characteristics, SES, and spatial accessibility were added to racial disparities in late-stage diagnosis as covariates of a logistic regression to investigate their contributions to the significance of racial disparities in breast cancer mortality. Principal component analysis (PCA) was utilized to reduce the multicollinearity among covariates and summarized the correlation structure displayed by the fourteen variables that were used to measure socio-demographic conditions and spatial accessibility to mammography facilities. The logistic regression analysis revealed that a census tract with significant racial disparities in late-stage diagnosis was 4 times more likely to have significant racial disparities in breast cancer mortality. Lower SES played an important role in determining whether a census tract displayed significant racial disparities in breast cancer mortality. However, proximity to mammography facilities had no impacts on the presence of significant racial disparities in breast cancer mortality for Hispanics, while centroids of census tracts that were closer to mammography facilities were more likely to have significant racial disparities for African-American women. For these women, most census tracts with significant racial disparities were located within the metropolitan areas which had higher concentration of health care facilities. In addition to the metropolitan areas, significant racial disparities for Hispanics were also found along the Southwest border of Texas, which lacked health care and had longer driving distance and time to mammography facilities.

This research analyzed the spatial patterns of racial disparities in late-stage diagnosis and breast cancer mortality, which shed new insights on the location of problematic areas and could help prioritizing the areas for effective intervention programs by accounting for population distributions. The identified risk factors in racial

disparities could help develop community-based intervention models and lead to a more efficient allocation of limited health resources with the ultimate goal of saving women's life. Subsidized health insurances and free mammograms for disadvantaged African-Americans and Hispanics could be applied at local communities to reduce racial disparities in breast cancer. The long term goal to improve the African-American and Hispanic women's health is to boost their income and enhance their social status through educational attainment.

CHAPTER I

INTRODUCTION

Health inequality is well noticed within and between developing and developed countries due to the complicated factors such as unequal access to health resources and health-related policies. For example, the United States is confronting the serious challenges of breast cancer disparities as a result of multiple races and sophisticated socioeconomic stratification. One of the overarching goals of Healthy People 2010 is to reduce and ultimately eliminate disparities in health across diverse dimensions such as sex, race, education, and geography (Healthy People 2010 2000). Healthy People 2010, established on the previous initiatives since 1980s, is a comprehensive array of health objectives in the United States. Breast cancer is the most common carcinoma disease among women and it shows disparities in its distribution within the United States. By focusing on reducing disparities from geographic perspectives and explaining associated risk factors leading to the disparities in breast cancer, this study is designed to contribute to the objectives of the Healthy People 2010 framework.

1.1 Background on Breast Cancer

Breast cancer is a severe disease for women due to the uncontrolled growth of abnormal cells on the surrounding breast tissues. As the most common cancer in

women, breast cancer accounted for 31% of diagnosed cancer cases in 2001 (Greenlee *et al.* 2001); exceeded by lung cancer, breast cancer is ranked as the second cancer death among women (Jemal *et al.* 2007). In 2006, 40,970 deaths were expected out of 61,980 in situ cases and 212,920 new cases of invasive breast cancer among women in the United States (Smigal *et al.* 2006). In situ means that abnormal carcinoma cells are contained in breast tissue without spreading into other body parts, while invasive (infiltrating) breast cancer signals that carcinoma tumours may invade other parts of body by spreading from original breast tissue (Hunt 2007). The lifetime risk for women to develop breast cancer has tripled in the last 60 years from 1 in 22 to 1 in 8 (Feuer *et al.* 1993). The historical pattern of higher mortality in breast cancer in the Northeast United States has been attenuated due to the increase of breast cancer mortality in the South (Sturgeon 2004). Thirty percent of breast cancer can be explained by established risk factors including age, geographic variation, age at menarche and menopause, age at pregnancy, lifestyle and environmental exposure (McPherson *et al.* 2000). It is estimated that 6-7% of diagnosed breast cancer cases result from inherited family genetics when solely based upon the first-degree relatives including parents, offspring, and siblings (Claus *et al.* 1996). A first degree relative is defined as a family member sharing 50% of his or her genes.

Among all women in Texas, breast cancer is ranked as the most diagnosed cancer, followed by lung cancer. Among females, breast cancer accounted for 29.7% of cancer cases diagnosed between 2001 and 2005, followed by 12.8% lung and bronchus cancer cases (Risser *et al.* 2009). However, breast cancer responsible for 15.8% cancer mortality is ranked as the second cancer death after lung and bronchus cancer (24.8%). The goal of Healthy People 2010 is to reduce breast cancer mortality to 22.3 from 25.0 per 100,000 in

the United States. The average mortality rate is 23.8 per 100, 000 as well as breast cancer incidence and mortality rates in Texas are lower in contrast with the national level (25.0). In Texas from 1998 to 2002, breast cancer mortality among white women was 24.4 per 100,000, compared to 36.0 cases per 100, 000 among African-American women (Smigal *et al.* 2006). Out of 12 million in Texas, 15,132 women were diagnosed with breast cancer in 2008; and 18.4 % of them (about 2,780) is expected to die from the disease (American Cancer Society 2008b). In 2007, the total cost was US \$3.37 billion for breast cancer, out of which the direct medical cost including cancer care and cancer screening was US \$1.35 billion in Texas (Tan *et al.* 2009).

1.2 Definitions of Disparity

The terms —~~health~~ disparity”, —~~health~~ inequity”, and —~~health~~ inequality” have commonly appeared in the context of the public health literature over the last three decades (Adler and Rehkopf 2008). The term —~~health~~ disparity” is commonly used in the United States while the other two terms are often employed outside the United States. A health disparity signifies differences in health occurring within people’s life course. These differences include where they grow up and live, their historic and present health status, how to access health care, and their resultant utilization and quality of cancer treatment. Health determinants have been divided into biological differences, cultural and personal practices, and health services in institutions. In addition, environmental factors such as social and physical environments play an important role in shaping people’s health. Social factors include a number of factors such as income, education, employment, and working conditions that comprehensively influence people’s health and

health care services they receive (Federal, Provincial and Territorial Advisory Committee on Population Health 1999; Whitehead 1991).

In the context of public health, the term “disparity” implicitly places ethical judgement on health status (Cambridge Dictionaries Online 2001), which distinguishes it from “inequity” which generally means two quantities are unequal. The term “inequality” generally indicates differences in health. The phrase “health disparities” carries the assumption that there is no difference among different subpopulations such as racial and social groups. Health disparities emphasize that the difference in health is unnecessary and avoidable across different populations of interest. If the difference does exist, it could be modified and amenable to interventions. For instance, the inequalities of health among age groups are generally unavoidable. Deterioration in the living and working environment is considered as an avoidable underlying cause of health disparities (Brown 1995). Minority and the poor tend to live closer to the environmentally hazardous facilities and suffer the health burden from toxins exposure, which is commonly referred as environmental inequality (Pellow 2000; Szasz and Meuser 1997).

Black and his colleagues (1980) examined the first study of health inequality with respect to the health burden carried by minorities and underserved groups. Despite the decades of research, no clear definition of health disparities emerges in the body of research literature throughout the 1980s and 1990s. Carter-Pokras and Baquet (2002) listed eleven different definitions according to their resources. For instance, Healthy People 2010 addressed health disparities as “differences that occur by gender, race or ethnicity, education or income, disability, geographic location or sexual orientation” (US Dept of Health and Human Services 2000, 14). From a public health view, the Centers of

Disease Control and Prevention defines a disparity as “the quantity that separates a group from a reference point on a particular measure of health that is expressed in terms of a rate, proportion, mean, or some other quantitative measure” (Keppel *et al.* 2004, 7). Reference points could be the majority of whites, the national or state population average, the healthiest group, and the highest socioeconomic group, and this may result in different conclusions about the strength and direction of a disparity (Adler and Rehkopf 2007). Causes of disparities across a range of health indicators sometimes are not fully understood due to the interaction of social and demographic factors underlying biological vulnerability, individual life behavior, and institutional structural barriers. Moreover, differences in lifestyle and social and physical environment appear to suggest unidentified risk factors in breast cancer. For instance, Maskarinec and his colleagues (2004) found that the descendants of migrants from the countries with lower incidence would approach the rate in the host countries, indicating the greater importance of environmental factors on breast cancer incidence.

1.3 Rationale of the Study

Breast cancer is the most frequently diagnosed cancer in Texas, followed by lung cancer among women. Breast cancer accounted for 29.7% of cancer cases diagnosed and 15.8% of cancer mortality between 2001 and 2005 (Risser *et al.* 2009). However, not everybody is equal when faced with this disease: different racial and socio-demographic groups display striking disparities with respect to the incidence of late-stage diagnosis and mortality rates. From 1998 to 2002, breast cancer mortality among African-American women was 50% higher than for white women: 36.0 deaths versus 24.4 deaths per

100,000 women (Smigal *et al.* 2006). The massive burden of breast cancer caused by the disproportionate representation of African-American and Hispanic women in Texas, combined with the prevalence of lower socioeconomic status (SES) in these minorities, is posing a great challenge to health-policy makers. The risks are further enhanced by the rapid aging of Texan population. By 2040, the number of Texans older than 65 is expected to triple, which will greatly intensify the social and economical burden in the State. In 2007 the total cost of breast cancer in Texas, including cancer care and cancer screening, was US \$3.37 billion, an astronomical figure when compared to the total cancer cost of US \$21.9 billion (Tan *et al.* 2009).

Awareness about disparity issues in health care and cancer has been put into the public spotlight by an alarming increase in cancer mortality among African-Americans, first proposed by Henschke in 1973. Minority groups experienced a smaller decline of mortality burden in breast cancer relative to white females, although the decrease of breast cancer mortality was evidently observed over the 10-year period 1996-2005 (Ries *et al.* 2007). The Texas Cancer Registry reported that, in 1995-2005, non-Hispanic whites had the highest incidence rate of 125.4 per 100,000 and second highest mortality rate of 24.3 per 100,000 among all races; African-American women had the highest mortality rate of 35.6 although its incidence was ranked behind non-Hispanic whites (American Cancer Society 2008b).

Health disparities related to socio-demographic factors are more imperative in a large state like Texas. Disproportionately representative of African-Americans and Hispanics as well as the prevalence of lower socioeconomic status (SES) in these minorities challenge the state of Texas with massive burden of breast cancer. It is

estimated that Hispanics will overtake all other racial and ethnic groups and become the majority population in Texas by 2035. Hispanics account for 56% of 3.8 million Texans living in poverty in contrast to 17% African-Americans and 24% non-Hispanic whites (Texas Health and Human Services Commission 2007). Hispanics disproportionately represent people under the poverty line in Texas. A possible explanation for the elevated risk of disease for people with lower SES in part stems from the extensive exposure to stress and limited health resources for buffering its impact. Poverty is the most critical factor on receiving preventative care in cancer and maintaining health for people. Thus, prevalence of poverty as an indicator of options of cancer treatment and its quality negatively impact population's health in Texas.

Mammography preventive screening is a powerful determinant on early detection in breast cancer, which provides better prognosis survival rates, leading to significant decrease of breast cancer mortality in the 1990s (Lannin *et al.* 1998). Screening patterns across geographic regions reveal the substantial observed racial disparities in breast cancer. Although health promotion efforts to improve mammography utilization have greatly aided in the reduction of racial/ethnic disparities in the last twenty years (Centers for Disease Control and Prevention 2007), structural barriers including transportation difficulties, inadequate number of facilities and providers in some areas of the nation hamper fully-utilizing mammogram service (Breen *et al.* 2007; Chagpar and McMasters 2007), and this ultimately hinders reaching the overarching goal of Healthy People 2010 (Keppel 2007). Moreover, the number of mammography facilities in Texas has dropped from 575 in 1994 to 508 in 2009 since the inception of Mammography Quality Standards Act and Program, which maintains the database of mammography facilities in the Food

and Drug Administration (Eastern Research Group 2001). Overall, mammography utilization in Texas is 5% less than the national level among all races. Hispanic women in Texas have 5% lower mammogram use compared to all other races/ethnic groups and 10% less utilization of mammography compared to Hispanic women throughout the United States (American Cancer Society 2008b).

1.4 Research Questions and Hypothesis

This research contributes to current knowledge of the spatial patterns of breast cancer, in terms of late-stage diagnosis and mortality in Texas from 1995 to 2005. The research addresses the following questions.

- 1) Do geographic and socio-demographic disparities exist in the occurrence of late-stage breast cancer diagnosis at the census tract level across the state of Texas from 1995 to 2005?

Null Hypothesis: There is no statistically significant difference across geographic regions and socio-demographic groups in the occurrence of late-stage breast cancer diagnosis at the census tract level in Texas from 1995 to 2005.

- 2) Do geographic and socio-demographic disparities exist in breast cancer mortality at the census tract level across the state of Texas from 1995 to 2005?

Null Hypothesis: There is no statistically significant difference across geographic regions and socio-demographic groups in breast cancer mortality at the census tract level in Texas from 1995-2005.

- 3) If these disparities identified from research questions 1 and 2 exist, are there any spatial connections between the disparities shown in the answers to these

two questions? Could the locations of mammography facilities and other socio-demographic factors contribute to these disparities?

Null Hypothesis: There is no strong connection between the disparities in mortality and the occurrence of late-stage breast cancer across census tracts.

This research attempts to provide a better understanding of racial disparities across geographic regions and facilitate health-policy makers to allocate health education and health care resources more effectively with the ultimate goal of reducing and even eliminating racial disparities in breast cancer. Geographic information techniques are utilized to overcome the barriers of lack of individual socioeconomic information and examine racial disparities in breast cancer incidence and mortality at the census tract level in Texas, and this will fulfill the objective of Healthy People 2010: reducing and eliminating health disparities across enormous dimensions particularly in race, socioeconomic status, and geography. The proximities of mammography facilities and other socio-demographic factors are investigated to explain the underpinning racial differences across geographic regions, with a goal of shedding light on the intervention programs of improving health and saving thousands of women's lives.

1.5 Significance

Cancer disparity occurs when different types of population groups do not experience the same health status and health care due to socio-demographic factors such as age, race, and education level (Center to Reduce Cancer Health Disparities 2004). Despite a growing body of research evidence on cancer disparities (Redmond *et al.* 2005), the cause and mechanism of racial/ethnic disparities in breast cancer outcome

remain partially understood due to the limited epidemiologic research on non-white populations. This study contributes to disparities research in breast cancer by comparing African-American and Hispanic women with their white counterparts, using a large population-based incidence and mortality dataset from Texas Department of State Health Services (TDSHS).

Unfortunately, the etiological mechanism for 70% breast cancer remains unclear. Disparities research in breast cancer has the potential to save thousands of women's lives by identifying underlying risk factors and developing strategic intervention programs to reduce and ultimately elucidate the disparities. This study utilizes numerous risk factors including race, age, socioeconomic, rural/urban residence, and location of mammography facilities, in order to explain detected racial disparities in late-stage diagnosis and mortality of breast cancer across regions.

Current knowledge about the causes and mechanisms of racial disparities in breast cancer outcomes is limited. Few studies have examined how racial disparities in breast cancer outcomes vary across geographic regions, and identifying geographic variance should play an important role in attempting to allocate health education and resources more effectively and efficiently. Investigating racial disparities across regions may provide useful insights to reveal unknown risk factors for breast cancer. Due to rapid growth of elderly and large proportion of population under poverty line, age and socioeconomic status have been crucial factors to consider in Texas in order to reduce breast cancer burden. Little research has been done to examine how locations of mammography facilities affect racial/ethnic disparities in breast cancer mortality. Further research needs to be performed regarding access to and utilization of mammography

facilities from a variety of population groups in order to identify underserved areas in view of the fact that the use of clinical facilities often interacts with their neighborhood environment. Finally, this study offers a valuable guideline targeting tailored preventive screening to underserved racial groups and regions.

1.6 Breast Cancer Disparities Research and the Disciplines of Geography

According to Pattison (1964) and Robinson (1976), there are four traditions in Geography: spatial, area studies, man-land, and earth science. The main subjects of study are people, location, and environment in Geography. Geographers examine questions of patterns and process (Chapman 1979). Although the four traditions are distinct logically, they interact with each other. The discipline of geography is broadly split into human and physical geography. In the Association of American Geographers, human geography organizes itself in the specialty groups including Health and Medical Geography, Economic Geography, and Human Dimensions of Global Change. The discussion of physical geography is beyond the scope of this study. Spatial study concentrates on spatial distribution of phenomenon including natural and man-made features on the surface of the earth. The spatial tradition based on the root of spatial analysis has an arsenal of quantitatively analytic techniques, which are closely related to Geographic Information Science (GIS) in geography.

GIS has transformed spatial tradition by providing the capability of complex analysis and helps researchers to identify geographic patterns through techniques of overlay, buffer, and visualization. The spatial tradition benefits from the use of contemporary techniques and has been revolutionized in terms of data availability and

analytic power (Longley and Batty 1996). GIS can identify areas that are in shortage of health professionals across geographic regions, and help build a strategic platform, ensuring successful intervention programs from health policy (Juarez *et al.* 2003). Availability of current numerous spatial datasets from a variety of health related agencies enable health practitioners and research scientists to analyze spatial and temporal patterns of a range of diseases. Geographic information technologies can be integrated into public health research and help monitor the effectiveness of government policies on reducing health inequalities (Higgs 2004). In summary, GIS development allows researchers and scientists to access more geographic health data, therefore by improving the study of geographic disease patterns in the ultimate purpose of understanding biologic pathways of underlying factors.

The research is situated within spatial epidemiology, a sub-topic in medical geography. Medical geography explores health-related topics by applying the concepts and methods of geography from an ecologic perspective (Meade and Earickson 2005), and spatial epidemiology study serves as the foundation and logic of interventions in public health through investigating disease spatial variation in socio-demographic risk factors for health in populations (Elliott and Wartenberg 2004). The existing surveillance systems including medical records and death certificates fail to collect socioeconomic information necessary to systematically analyze racial disparities at an individual level (Krieger *et al.* 1997b). Area-based socioeconomic approach facilitated through geocoding techniques has assisted in overcoming the barrier of socioeconomic data paucity in the process of quantifying socioeconomic disparities in health. Thus, social inequality studies have been revived by heightening interest in health disparities due to emerging innovative

GIS techniques. The general research agenda of health inequality is to explore how social factors work together and separately shape an individual's health through the environment where he or she lives.

1.7 Outline of the Dissertation

To tackle the research questions proposed in Section 1.4, the rest of the study is organized as follows. Chapter Two offers a comprehensive overview of the literature regarding health disparity, particularly in breast cancer. This chapter further discusses the current interventions and limitations pertaining to the health disparity research. The research gap on breast cancer disparities is pointed out through the comprehensive examination of the literature.

Chapter Three presents the detailed description of breast cancer incidence and mortality datasets used in the research. The analytical methods used in this study are provided. Chapter Four describes the results of the significance tests of racial disparities in both late-stage diagnosis and breast cancer mortality for both African-American and Hispanic women at the census tract level in Texas. The spatial connections of the two types of racial disparities are examined. Chapter Five assesses racial disparities in breast cancer mortality using the explanatory risk factors including the significance of racial disparities of late-stage diagnosis, demographic characteristics, socioeconomic status, and spatial accessibility. Chapter Six concludes the research with potential intervention programs and highlights future research work with respect to health disparities in breast cancer.

CHAPTER II

LITERATURE REVIEW

This chapter reviews a range of literature related to health disparities, particularly in breast cancer which provides intellectual foundation for this research. Health disparities research has been revived because one of the overarching goals from Healthy People 2010 is to reduce and ultimately eliminate health disparities among diverse population groups. Consequently, research concerning disparity, inequality, and inequity in health has multiplied dramatically since the beginning of twenty-first century. Studies with “health disparities” as key words have grown from a few in the 1990s to hundreds in the 2000s. The first section of the chapter provides a general overview of health disparities not including cancer due to race, socioeconomic, and geographic determinants. It is followed by a second section which offers an extended review of previous studies in cancer with exclusion of breast cancer. The third section gives a detailed overview particularly in breast cancer. The chapter is concluded with discussion on the intervention strategies (section 4) and current research (section 5) limitations in health disparities.

2.1 Disparities in Health (not including cancer)

Health disparity research has significantly expanded in the past two decades since the Department of Health and Human Services launched “Healthy People 2010” to eliminate health disparities (Keppel *et al.* 2004). By 2015, the American Cancer Society (ACS) is striving for “the elimination of disparities in the burden of cancer” (Byers *et al.* 1999). Geographic disparity research has been fueled by the emergence of geographic information technologies (Clarke *et al.* 1996). In addition, the pervasive effect of race and socioeconomic issues on health in the United States provides a strong impetus to document and monitor a variety of disparities across a wide range of health indicators in order to develop effective intervention programs allocating health education and resources more fairly.

2.1.1 Racial Disparities

Race is defined by specific hereditary and biological descent; whereas ethnicity places a focus on cultural tradition. In a comprehensive review of racial/ethnic disparity in health care, the Institute of Medicine (IOM) described a model stating that health care disparities are consequences of the interaction among economic, social, and cultural factors (Freeman 2003). In an IOM landmark study, Smedley and colleagues (2002) documented the disparities in a range of health care services, including cardiovascular care and mental services. Clinician and patient characteristics in addition to the health care system have been proposed to explain the differences in the utilization and quality of health care (Klonoff 2009). According to Klonoff (2009), provider-patient communication, patient’s decision-making, and familiarity with health care system

deserve further investigation on the impact of health disparities. Not only health disparities exist among adults, but it was reported that African-American infants were 2.3 to 3.2 times more likely to have low-birth weight than their white counterparts due to maternal conditions (Kempe and Wise 1993).

Racial disparities should be diminished or even eliminated thanks to the absence of socioeconomic barriers if equal health care is provided to the public. Ross and Mirowsky (2000) argue that public health programs including Medicare and Medicaid lead to worse health among non-white minorities. For example, racial/ethnic disparities in hip and knee replacement surgery were remarkably noticed between African-Americans and whites in the Veterans Affairs health care system (Ibrahim 2007). Financial barriers were diminished within the health care system. In the equal-access of Veterans Affairs health system, disparities were most pronounced for medical adherence and surgery procedures involved with communication (Saha *et al.* 2008). Furthermore, a study of the national Medicare system measured 21 indicators in health care quality and found that Hispanics and African-American still needed health care improvement (Hebb *et al.* 2003).

Two major reports from IOM have spurred awareness of health disparities and stimulated prevention programs to improve the minority's health at national and state levels (Haynes and Smedley 1999; Marks 2000). Understanding health-seeking behaviors among minorities is also critical to approach the goal of eradicating health disparities. For instance, Latinos tend to be influenced by a combination of remedies, advice from family members, and strategic differences in health-seeking preferences (MacNaughto 2008). Established interventions to health racial disparities involve telephone outreach,

promotion of medical access, and medical adherence (Chin *et al.* 2007). However, few studies tested the effectiveness of interventions in reducing racial disparities in the vascular disease setting (Davis *et al.* 2007).

2.1.2 Socioeconomic Disparities

There are two explanations for the interaction between Socioeconomic Status (SES) and health. The first explanation is that SES influences health status (social causation); and the second reason is that health status ascribes to socioeconomic status (social selection). Existing health databases include SES as a crucial factor in shaping individual health status (Fox *et al.* 1985; Haan *et al.* 1989). Human behaviors such as cigarette smoking, physical inactivity, poor diet, and substance abuse are highly related to SES and play determining roles in health outcomes (Marmot *et al.* 1984).

Social inequalities in health reflect the unequal distribution of wealth, deprivation and privilege in population (Lloyd 1978; Veith 2002). In general, people live longer and have better health care in more affluent countries than in poor countries. Relative deprivation is associated with poor health and health is a non-linear negative function of absolute income. Ill health and income may have an inverse relationship over life course of people. The poverty hypothesis implies that absolute income is critical for people's health in poor countries and relative income (comparing with groups) is more important in rich countries (Ravallion 1997).

The influence of social inequality on health is evident through the main pathway of psychosocial factors including stress and discrimination (Brondolo *et al.* 2009; William and Mohammed 2009). John Lynch's study (2000) illustrated that a death rate

caused by health problems from income inequality is as high as the combined loss of life from lung cancer, diabetes, and motor vehicle crashes in the United States in 1990.

Cassel (1976) argued that the improvement of social support could ultimately reduce disease by changing human resistance to environmental agents. To answer the question “who and what drives current and changing patterns of social inequalities in health” (Krieger 2001, 672), Krieger called upon the importance of establishing theoretical frameworks to identify the causes and barriers of reducing social inequalities in health.

Social inequality studies may indicate etiologic mechanisms of neighborhood effects on health (Diez-Roux 1998; Duncan and Brooks-Gunn 1997). The effects of social inequalities on health have been investigated, and health-related policies have been established in an attempt to reduce these inequalities (Leon and Walt 2001; Smith 2003). Public health professionals and researchers are investigating how social and racial inequalities impair health among the poor due to adverse living and working conditions along with inadequate health care accessibility throughout life. Social interventions such as better health insurance could reduce the inequality both in health and income (Deaton 2001).

Socioeconomic status and race are two independent but interrelated variables which influence health outcomes. Moreover, socioeconomic status alone plays an important role in shaping health outcomes. LaVeist (2005) demonstrated that within racial groups, there were negative relationship between income status and elevated blood level among men 18 and older. Across the same income group, African-American men had higher lead level compared to non-Hispanic white men. However, even in South Korea with racial homogeneity, death rates associated with accidents and diseases were

found to be explained by lower education attainment (Khang *et al.* 2004). The possible explanation for the elevated risk of disease for people with lower SES may stem from the extensive exposure to stress and limited health resources.

Economic impoverish restrains people among underserved and disadvantaged groups to access and obtain adequate health care. The example of Steve Jobs, the CEO and cofounder of Apple underwent liver transplant in 2009, which vividly emphasized the importance of health insurance barriers and financial support (Hainer 2009). Despite a nominally universal health system in New Zealand, Maori (indigenous population) and Pacific people have worse access to health care and health services than Europeans in New Zealand Pacific women. As a result, the indigenous people suffer from a mortality rate that is three times higher than non-Maori non-Pacific (European dominantly) (Marrone 2007; Sarfati *et al.* 2006).

Health inequality research aims to explore how social factors work together and separately in shaping an individual's health through the environment where he or she lives. SES has been treated as a control variable instead of an independent etiologic factor in most current literature (Elster *et al.* 2003). Health insurance coverage, predicted by race and SES, is strongly associated with access and utilization of health care.

2.1.3 Geographic Disparities

Racial disparities concerning health may be compounded by geographic variation of health treatment. Racial disparities persists in diverse health treatments after controlling for access to treatment and SES, and this may suggest some other unknown factors underlying geographic regions (Gross *et al.* 2008). Dartmouth Atlas of Health

Care documented a variation in the use of Medicare health system across 306 Hospital Referral Regions (Wennberg and Cooper 1994). Barnett (2001) and Casper (1999) examined race and gender-specific overall mortality, and disease-specific death rates at county and state levels, and the researchers found that both treatment patterns and health outcomes substantially deviated across geographic lines.

Reducing geographic disparities in the quality of care and treatment can in part eliminate racial disparities in health outcome and benefit all Americans. In contrast to non-Hispanic whites, minority groups such as African-Americans and Hispanics visit different hospitals to seek care within their own neighborhoods (Lillie-Blanton and LaVeist 1996). Aggregated data can identify the community characteristics of health care and treatment by accounting for region-specific effects (Chandra and Skinner 2003). Improved access to hospitals and quality of health care can help alleviate health disparities among different racial groups mainly resulting from regional difference in treatment and consequent health outcomes. However, the causes of racial disparities cannot be eradicated by delivering universal health insurance or equal access to health care at smaller regions because the lower quality of health care is often offered in areas where minority groups tend to reside.

The Dartmouth Atlas Project has produced a report that demonstrated regional and racial variations in the health care of leg amputation, the management of diabetes, and ambulatory care-sensitive hospitalization rates across Dartmouth Atlas hospital service areas (Fisher *et al.* 2008). Strong evidence of substantial variation in hospital use stresses the importance of exploring and addressing underlying causes of disparities within and across regions. Disentangling health disparities varying by race or region has

important implications on health policy-making in terms of allocating resources and targeting prevention programs. However, only a few studies have so far been conducted with respect to geographic disparities in health in that most quantitative studies compressed all cases together across geography because of issues involving confidentiality as well as difficulties in interpreting geographic differences.

2.2 Disparities in Cancer (not including breast cancer)

Cancer rates steadily increased until the early 1990s. The rates have, however, decreased in both men and women, especially in recent years. The sharp decline of cancer incidences is mainly attributed to a decrease in male cancer as a result of tobacco use intervention starting in the early 1990's, while incidences in female cancer were increasing until the early 1990s and then have leveled off. Cancer deaths accounted for 23.2% of overall fatal cases in 1998, preceded only by heart disease. Mortality of all cancers combined reached its peak in 1991 for both men and women and declined at 1.1% annually during the period from 1991 to 1997 thanks to better cancer treatment options (Greenlee *et al.* 2001).

Monitoring and documenting health disparities in cancer can help identify underlying mechanisms by exploring the causes of the disparities. Cancer variations by race, socioeconomic status and geographic region have drawn much research interest in detecting the sources of inequality and in implementing intervention program more efficiently. Cancer disparities could be reduced or even be erased through implementing potential interventions (reduction in tobacco use, physical inactivity, and obesity), early

detection (mammography, colorectal screening, and pap smear tests), as well as appropriate treatment and care (Ward *et al.* 2004).

2.2.1 Racial Disparities

Cancer incidence and mortality vary across racial/ethnic groups. For example, African-American women have the highest incidence and mortality rates in all cancers combined (Devesa *et al.* 1999); they are 33% more likely to die of cancer compared with that of white and twice as much to die of cancer as any other minorities (American Cancer Society 2004).

Racial disparities persist across diverse cancer diseases and cancer outcomes. Alexander and his colleagues (2007) found that African-Americans experienced higher mortality and lower survival rates in colorectal cancer compared with Caucasians. Nevertheless, the magnitude of racial disparities differs across study groups in that African-American women are 10% to 80% more likely to die of colorectal cancer than white women (Alexander *et al.* 2007). Racial disparities in cancer therapy have not decreased since 1990s. Gross and colleagues (2008) evaluated receipt of cancer treatment among African-American patients diagnosed with colorectal, breast, lung, and prostate cancer from 1992 to 2002 in the Surveillance, Epidemiology and End Results (SEER)-Medicare linked database. The authors found that the cancer therapy in Medicare beneficiaries did not improve the magnitude of racial disparities, which indicated failure of initiatives to alleviate cancer therapy inequalities.

Cancer disparities among races are associated with other risk factors. For instance, in a retrospective cohort study, African-American patients are 1.36 times more

likely to die from colon cancer than their white counterparts with equivalent life-threatening symptoms (Ahuja *et al.* 2006). The worse cancer outcomes in minority groups may be due to the treatment difference and consequent lower survival. Using SEER-Medicare database in the period of 1985 and 1999, African-American patients had 12.7% lower surgical rate for the early stage of lung cancer as compared with their white counterparts, leading to 7.7% lower in survival rate. However, there was no survival difference observed among groups undertaking surgeries (Bach *et al.* 1999).

2.2.2 Socioeconomic Disparities

Health effects of socioeconomic status (SES) are pervasive at every level regardless of what the indicators or cutoff points are. SES is often obtained from the Census at certain geographic level due to the absence of individual SES. Poverty status is the most common single measure of SES used in health research with the assumption that health will not be improved significantly above the poverty line. Instead, every level of SES categorized by different social measurement complicatedly shapes people's health (Adler and Ostrove 1999). SES could also be created based on summary ranking of composite variables including median income, poverty, education, and employment at a geographical scale such as zip code (Eggleston *et al.* 2006). Kriger and colleagues (2005) demonstrated that SES at the census tract level provides similar results as individual level data in addition to considering neighborhood effects. Wang and colleagues (2008) assigned patients diagnosed with Non-Hodgkin Lymphoma between 1992 and 1999 with an index of SES according to their residence at the census tract level, which was derived from the 1990 Census is composed of education, poverty level, and median income

categorized by quartiles with equal weights. This study found that African-Americans did not suffer mortality burden after controlling for treatment and socioeconomic status.

Using the same SES index, Du and colleagues (2007) concluded that a lower SES reduced survival rates among African-Americans who had marginally higher mortality rate relative to Caucasians.

The direction and strength of the association between SES and cancer depends on the type of cancer and whether one is interested in incidence, mortality or survival.

Cervical and lung cancer occur most commonly among people with lower economic status. Smoking prevalence strongly associated with SES explains 80% of lung cancer (Blot and Fraumeni 1996). On the other hand, breast cancer and melanoma are more likely to be diagnosed in affluent people (Shack *et al.* 2008). Overall, existing literature has an unbalanced focus on racial disparities with absence of research on social inequalities in the cancer continuum of prevention, etiology, and access to clinical trials (Palmer and Schneider 2005).

Race and socioeconomic status interact with each other to impact on cancer continuum from incidence, diagnosis, intervention, and mortality. Minority people tend to be affiliated with lower SES of income, education attainment, poverty, employment and occupation. Studies in the current body of literature provide more evidence that race and socioeconomic status are two independent indicators of cancer outcomes. In the breast, cervix and uterine corpus cancer data, the interaction of race and SES marginally impacts all three cancer survivals and SES projects only breast and uterine corpus cancer (Greenwald *et al.* 1996). Income and education, but not race, are significant predictors of survival from multiple cancer diseases (Cella *et al.* 1991). In a population-based cohort

study of colon cancer, no significant interaction was reported between race/ethnicity and socioeconomic status using the product term of the two variables in a Cox proportional hazards regression model (Du *et al.* 2007). From methodological perspectives, multivariate regression models are the most common used approach to deal with the confounding of race and socioeconomic status.

2.2.3 Geographic Disparities

The likelihood of developing and surviving cancer is associated with where people live and what races/ethnicities they are. The neighborhood they live in could be treated as crude surrogates for underlying unknown factors. Cancer incidence varies by socioeconomic status and geographic locations which reflect underlying effects of living environment, lifestyle factors, and access to as well as utilization of health care (Macleod *et al.* 2000, Neal and Allgar 2005). Mortality and survival variations likely reflect differences in access to health care and quality of cancer treatment across regions (Farrow *et al.* 1996; Morrison *et al.* 2000).

The Australian Institute of Health and Welfare in conjunction with the Australasian Association of Cancer Registries (2003) found 7% better survival in urban areas for all cancers than in remote centers and larger rural areas in Australia. The Queensland Cancer Registry reported 5% difference in cancer survival between inner regional and remote centers (Baade *et al.* 2005). Sabesan and Piliouras (2009) hypothesized that lack of health providers, delay in referral, transportation barriers, and support services are contributing factors to cancer difference between rural and urban

areas. Observed cancer survival and mortality are associated with rural and urban residence.

Existing literature suggests inconsistent relationship between rural-urban and cancer outcomes. Place of residence is highly associated with proximity to health care and convenience of sustainable cancer therapy treatment. Urban population has more likelihood to obtain better quality cancer treatment and access to more screening facilities. Reversely, Eggleston and colleagues (2006) found no association between rural residence and cervical cancer stage as well as survival in contrast to the findings in Australia that women living in rural areas had odds ratio of 19.4 to die from cervical cancer (O'Brien *et al.* 2000).

2.3 Disparities in Breast Cancer

As the most common cancer among U.S. women, breast cancer alone accounted for 31% of diagnosed cases in 2001; breast cancer was preceded only by lung cancer among women for cancer deaths (American Cancer Society 2004). The lifetime risk for women to develop breast cancer has tripled in the last sixty years from 1 in 22 to 1 in 8 (Feuer *et al.* 1993). The contributing risk factors include BRCA1 and BRCA2 mutations (5.0%) (American Cancer Society 2004), family history of breast cancer (9.1%), late age at first birth and nulliparity (Madigan *et al.* 1995), high alcohol consumption (10.7%), low beta-carotene intake (15.0%) , low vitamin E intake (8.6%), low levels of physical activity (11.6%) (Mezzetti *et al.* 1998), and smoking (2.5%) (Ishibe *et al.* 1998). However, unknown factors are responsible for the remaining 30% breast cancer cases.

This section attempts to provide an extensive review of disparities research on female breast cancer and its potential explanations. The existing research on breast cancer can be divided into three categories: racial, socioeconomic, and geographic disparities. This section also identifies the research gaps in the literature that are worthy of further investigation. Appendix lists the primary research regarding breast cancer disparities.

2.3.1 Racial Disparities

In general, white women have a higher incidence rate in breast cancer than African-American, while African-American women have a higher mortality rate than whites (Chu *et al.* 1996). Female breast cancer incidence increased rapidly for all racial groups during the period of 1980 to 1987, which could be explained by the increase of mammography use in the United States. However, incidence remained at a much slower increase from 1987 to 2002 (Smigal *et al.* 2006).

Even within the racial groups, incidence rates increase differently across various age groups. In the United States, approximately 23% of all breast cancer cases occur among women younger than age 50 which represents 73% of the total female population. Incidence rates increased dramatically by age and reached its peak at the age group 75-79 for both white and African-American women during the period of 1975 to 2002. It may reflect the efficacy of mammography use to detect the slow growing tumors that seem more prevalent in elderly (Smigal *et al.* 2006).

White women have a heightened incidence trend for ages 50 and older, but less applicable to African-American women who show a stable incidence trend in this age group. Hormone Replacement Therapy (HRT) is a risk factor for developing breast

cancer which may explain the increased trend for white women. In addition, HRT is associated with larger and more advanced carcinomas tumors (Chlebowski *et al.* 2003). For women under age 50, incidence rates have been stable for whites since 1986 and have decreased for African-Americans since 1991 (Smigal *et al.* 2006). Mammography screening variation among racial groups may account for some of the observed difference in breast cancer incidence (May *et al.* 2000).

Obesity remains an important variable to explain racial disparities in breast cancer. The descending trend for premenopausal African-Americans (Magnusson *et al.* 2005; Vainio and Bianchini 2002) may be associated with multiplying occurrences of obesity (Flegal 2005). The prevalence of weight gain may result in an elevated risk for postmenopausal women due to the extra estrogen exposure produced by fat tissues (Feigelson *et al.* 2004). However, Hall and his colleagues (2000) did not find any evidence of body mass index for increased risk among postmenopausal women of both African-Americans and whites. For example, a survival difference study between Japanese and Caucasian patients suggested a weak negative association between obesity and breast cancer survival (Marchand 1991). Low fat intake has been implemented in adjuvant therapy for postmenopausal women with early breast cancer in the United States. Contrarily, obesity impact breast cancer stage differently due to difficulty in detecting carcinoma tumors. A study of breast cancer in Connecticut reported that one third of observed racial differences in breast cancer stage could be explained by the severe overweight among African-Americans (Jones *et al.* 1997). Another replicated study in North Carolina also demonstrated that the odds ratio of the stage for African-

Americans was reduced by 27% in a multivariate logistic regression model by considering hip ratio and severe obesity (Moorman *et al.* 2001).

There has been a growing body of research on evaluating racial/ethnic disparities in receiving cancer treatment (Shavers and Brown 2002). Optimal cancer treatments may potentially be influenced by structural factors, such as insurance coverage, medical facilities, and staff to provide standard treatments (Freeman and Reuben 2000). Moreover, the clinical factors impacting breast cancer treatment were identified as menopausal status, stage, auxiliary lymph node status, histology, and nuclear grade of primary tumors as well as estrogen receptor and progesterone receptor status. Despite a seemingly minor racial/ethnic variation in the efficacy of breast cancer treatments, there is a declining trend with recommended adjuvant therapy among some rural African-American and white women (Tropman *et al.* 1999). Consistent evidence revealed the trend that African-American women had less radiation therapy after breast conserving surgery relative to whites (Ballard-Barbash *et al.* 1996; Riley *et al.* 1999; Steve *et al.* 2008). In addition, Asian and Hispanic women frequently received fewer breast conserving surgeries compared with white women (Morris 2000). Owusu and colleagues (2006) found that women older than 75 were less likely to receive guideline therapy which were highly associated with decrease in age-specific breast cancer survival. White women also experienced shorter intervals of diagnosis and treatment than other racial groups (Caplan *et al.* 2000). The good news is that government screening programs provide similar treatment for medically underserved women compared with all women diagnosed contemporarily in the same regions (Richardson *et al.* 2001).

Estrogen Receptor (ER) status has been established as a treatment predictor for breast cancer patients (Clarke *et al.* 1998), whereas Progesterone Receptor (PR) is a predictor for survival (Elledge and Fuqua 2000). ER+PR+ tumors are more likely to detect at stage I except in African-American and Hispanic groups. More stage I tumors detected among white females has been attributed to mammography use. Stage II tumors are more frequently detected in African-American and Hispanic groups in that these two groups have not undergone extensive mammography use. However, ER-PR- tumors have a higher likelihood in stage II for both whites and other ethnic groups. This finding reflects that mammography screening may not be able to discern ER- tumors as early as ER+ tumors (Chu *et al.* 2001).

According to estrogen and progesterone characteristics of carcinogen, Basal-like tumors (ER-, PR-, and HER2-) are more prevalent among African-American women, and they occur more often in premenopausal women than postmenopausal (24% versus 15%) (Perou *et al.* 2000). Among African-American women, basal-like breast cancer accounts for 40% of cancer cases for premenopausal women, but only 6% for postmenopausal women (Carey *et al.* 2006). Because basal-like tumors are an extremely aggressive subtype of breast cancer, their prevalence may contribute to the high mortality rates of younger African-American women. No survival difference observed among older African-American women suggests that the Medicare program helps alleviating racial disparities in breast cancer treatment (Chu *et al.* 2003).

In the mid-1980s and the early 1990s, a declining trend from regional stage to localized stage of breast cancer might reflect the effectiveness of mammography promotion efforts and earlier reporting of cancer symptoms. However, the aggregation of

regional stages since 1993 may imply more accurate classification due to improvement of sensitive technology to identify lymph node metastasis (Mullenix *et al.* 2005; Reintgen *et al.* 2000). Among younger white women, the increase at distant stages may reflect more aggressive tumors that are not easy to detect by screening prevention (Talley *et al.* 2002).

In general, cancer survival rate is primarily determined by cancer stage that varies significantly within the racial groups. African-Americans, Hispanics, and American-Indians had more advanced stages of breast cancer in comparison to whites, Asians, and Pacific Islanders. These minority groups are more likely to have larger and higher grades as well as estrogen receptor and progesterone receptor negative tumors (Li *et al.* 2003). Even within the same general racial group such as Asians and Pacific Islanders, Japanese and Chinese women have better survival rate, whereas Hawaiians have the worst survival relative to non-Hispanic whites (Meng 1997).

The survival rate differences between white and African-American women may indicate more late-stage tumors and inferior accessibility to beneficial treatments among African-American women (Chu *et al.* 2003). Survival variation between white and African-American women in breast cancer reflects more discrepancy from access to cancer treatment and quality of care than from biological differences of carcinogen tumor characteristics (Brawley and Freeman 1999). Utilization and quality of cancer care are potentially influenced by structural barriers: insurance coverage, treatment recommendations from physician, and patients' own decision-making (Shavers and Brown 2002). Racial disparities in breast cancer mortality have been widened since 1980s due to a smaller decline in mortality within the minority groups. Using Surveillance Epidemiology and End Results (SEER) data from the National Cancer

Institute between 1990 and 2003, Menashe and his colleagues (2009) found that increased racial disparities in breast cancer mortality were mainly driven by higher death rates among African-Americans.

2.3.2 Socioeconomic Disparities

Socioeconomic Status (SES) has a paradoxical influence on breast cancer incidence and mortality. Higher incidence rates are often found associated with higher SES (Clarke *et al.* 2002; Hsu *et al.* 1997), which has been attributed to risk factors such as late-age childbearing or lifestyle behavior, and promotion of mammography screening among higher-SES population (Althuis *et al.* 2005; Hermon and Beral 1996; Kelsey *et al.* 1993; Mcpherson *et al.* 2000). On the other hand, the higher mortality rates among African-American could be explained in part by the lower SES which determines population with access to health resources. The historical pattern of higher breast cancer mortality in the Northeast U.S. has been attenuated due to increase of the breast cancer mortality in the South region that generally has lower SES (Sturgeon *et al.* 2004). Residents living in poor counties had lower death rates of breast cancer than those from affluent counties in 1975 because of lower incidence rate (Singh *et al.* 2003). In 1999, the relationship was reversed that people living in affluent counties had lower mortality resulting from the favorable survival effect of higher SES.

Freeman and Chu (2005) proposed social determinants of health disparities model to guide researchers to identify the social, economic, and cultural barriers in order to develop effective strategies to reduce racial disparities among African-American women in breast cancer mortality. Gerend and Pai (2007) identified social barriers of breast

cancer mortality disparities between African-American and white women. Serious poverty and deprivation may restrain people from easy access to primary care physician, medical care, competing survival priorities, comorbidity, health insurance, information, and knowledge about health, risk-promoting lifestyles, and provider-system-level factors. Culture related barriers include spirituality, perceived susceptibility to breast cancer, cultural beliefs, and attitudes. Additionally, medical mistrust and barriers related to social injustice are induced by racial discrimination, prejudice, and physicians' perceptions of patients. Gerend and Pai point out that one important challenge lies in understanding the pathways of these factors to promote difference in breast cancer outcomes.

Chu and his colleagues (2007) explored cancer disparity measured by cancer rate ratio and ratio difference by race and socioeconomic status during 1990-2000. Poverty level at resident counties was used as a measurement of SES within each racial group. The positive ratio difference for African-American women indicated the disparity increased from 1990-1994 to 1995-2000. The increased disparity may be explained by the fact that mortality decline is larger for white women than African-American women (Chu *et al.* 2007). Vona-Davis and Rose (2009) examined existing literature and concluded that strong evidence exists regarding how low SES exerts aggressive biomarker on breast cancer tumors and worse prognosis. The highest socioeconomic group among African-American women benefitted most from mortality decline (known as positive-high socioeconomic gradient). Hispanic and Asian Pacific women also had positive-socioeconomic gradients for mortality rates during this period. These findings postulate that the highest socioeconomic group takes the most advantage of health resources and benefits most from favorable biomedical interventions (Freeman and Chu

2005). However, the normal pattern of socioeconomic inequality on mortality rates did not apply to minority groups such as American Indians or Alaska Natives. It is possible that additional socio-cultural factors such as language, environment, beliefs, and religion could play an important role in shaping SES patterns and mortality rates of other ethnic groups (Brach and Fraser 2000).

Lannin *et al.* (1998) interviewed 743 women diagnosed with breast cancer at the University Medical Center of Eastern Carolina and found that education and income per capita in a family in conjunction with cultural belief factors, largely justified the stage difference in African-Americans. Gwyn and colleagues (2004) also discovered that the odds ratio of treatment delay among African-Americans reduced substantially after adjusting poverty index and insurance status that used as a measurement of SES. However, even among the insured population without barriers to access medical resources, African-American women still have 7.8% less in five-year survival rate than whites (Field *et al.* 2005).

If individual socioeconomic information is not available, aggregate measurement based on patient's residence could be used as a crude surrogate for SES. For instance, in a study of cancer incidence rate by socioeconomic and region in England, income domain at zip code was assigned to each patient and ranked at quintile order which includes 20% of the total population in England (Shack *et al.* 2008). Moreover, SES is commonly derived from socioeconomic index of median household income, poverty rate, education attainment, household crowding, employment, and Townsend deprivation index at the census tract level (Krieger *et al.* 2006). Using the percentage of population living under

the federal poverty line within a census tract, Bradley and colleagues (2002) found that socioeconomic status is more important predictor for breast cancer outcomes than race.

Individual records may cause individualism bias, while aggregated data may be associated with ecological fallacy. The method to estimate SES based on the neighborhood people live is less precise than individual data (Krieger *et al.* 1997a). However, in a large prepaid health plan dataset, Krieger validated area-based measurement methods and observed the similar association in health outcomes between individual and census-tract economic measurements (Krieger 1992). Hence, the area-based measurement of SES is an appropriate approach to estimate individual SES given that the current surveillance system does not collect individual SES information such as income and education levels. SES derived from different geographic scales may lead to Modifiable Area Unit Problems (Krieger *et al.* 2002; Openshaw 1984). Depending on the scale and zoning of individual and aggregated data, the results (e.g. correlation between income inequality and health) may not be consistent (Rogot *et al.* 1992).

Social differences are intertwined with racial/ethnic disparities concerning health. Social inequality can be taken into consideration using Poisson regression model to adjust for cancer racial disparities (Krieger *et al.* 2005). Breast cancer is more prevalent among white women and higher SES groups. Breast cancer incidences have higher rates in affluent counties than in impoverished counties for all age groups among both white and African-American women. White women had higher incidence rates than African-American women in all the socioeconomic groups during the period of 1998-2002 (Singh *et al.* 2003). Breast cancer incidence has increased for all socioeconomic groups measured by poverty level during the period 1975-1999 (Singh *et al.* 2003). Vainshtein

(2008) reviewed five studies pertaining to breast cancer incidence stratified by race and socioeconomic and found that the magnitude of racial disparities decreased as SES increased.

There is substantial evidence supporting the argument that the variation in breast cancer diagnosis, treatment, and survival resulting from socioeconomic disparity is larger than racial/ethnic disparity (Ayanian *et al.* 1993; McGinnis *et al.* 2000; Roetzheim *et al.* 2000). Bradley *et al.* (2002) conducted a logistic regression analysis based on a database of Metropolitan Detroit SEER Registry in order to disentangle the effect of race and socioeconomic status on breast cancer. African-American women were found similar to whites at diagnosis, treatment, and survival of breast cancer when controlling other covariates such as age, insurance coverage, and socioeconomic status.

Higher SES groups are reported to have lower breast cancer mortality by benefitting the most from intervention programs. For example, Chu *et al.* (2007) found that racial disparities increased the most among minorities with the lowest SES group. In a study of New Zealand, breast cancer mortality disparities among Maori women was broadened relative to non-indigenous people, which was attributed by a significant decrease in breast cancer mortality among higher income and education groups (Sarfati *et al.* 2006). Latinas were reported to undergo less mammography screening due to lack of health insurance and low SES (Wells and Roetzhei 2007).

2.3.3 Geographic Disparities

Geographic disparities of cancer mortality and survival indicate a noticeable discrepancy in access to cancer facilities and quality of cancer treatments across the

United States. Based on the SEER dataset, Grann *et al.* (2006) assessed breast cancer mortality according to SES measured by education, income, and employment at the county level. Among all the SEER regions, Detroit in Michigan has the worst breast cancer survival rate while Hawaii has the best. Therefore, geographic variation of cancer due to SES and race/ethnicity can identify targeting subpopulations and facilitate an unbiased basis for health care policy-making and health resource allocation.

Disparate breast cancer outcomes differ across regions (Tian *et al.* 2010). Shack and his colleagues (2008) assigned an index of income domain to each cancer patient according to their postal code of residence at diagnosis in England by age and region. The authors found that the modest variation of socioeconomic-specific incidence in breast cancer existed between and within regions. A cluster detection study across Texas revealed that the elevated breast cancer mortality occurred along the Gulf coast and Central Texas for non-Hispanic whites. Moreover, Hispanics had a relative risk of 18% with excessive mortality burden in western Texas (Hsu *et al.* 2004).

Mammogram intake is highly associated with geographic regions because it is directly affected by the locations of mammography facilities. Utilization of mammography among urban and rural women demonstrates that rural residence and Hispanic ethnicity are risk factors for underuse of mammography screening (Coughlin *et al.* 2002). Geographic patterns of mammography use in Toronto displayed strong relationships with income and immigration status (Glazier *et al.* 2004). The neighborhoods where mammography facilities are located also lower the likelihood of late-stage breast cancer (Tarlov *et al.* 2009).

The disparity assessment of breast cancer outcomes could be measured at different geographic levels. The Dartmouth Atlas project utilized local hospital referral region to analyze racial disparities in health care (Baicker *et al.* 2004). The inconsistencies across geographic scales have been evaluated for breast cancer survival in Michigan using federal house legislative districts, state house legislative districts, and community-defined neighborhood (Meliker *et al.* 2009).

Considering cancer incidence, survival, and mortality rates, race and region are significantly interconnected with each other. Geographic variation of cancer outcomes is an integral part of racial disparities in that minority tends to live in poor neighborhood with less health resources and limited access to health care. Compared to white women, African-American and other disadvantaged minority groups tend to live in and seek care from different regions and physicians due to financial barriers and inferior accessibility to health resources (Lillie-Blanton *et al.* 2001). Therefore, equal access to health care at local or hospital levels may not elucidate racial disparities, which may explain why racial disparities still persist in areas with universal health care. Even after adjusting for cancer stage variation across the regions, persistent and substantial differences in treatment patterns and health outcome remain (Barnett *et al.* 2001).

Most studies obfuscated racial disparities and geographic disparities in health across the US because of ineffective sampling strategy of a national database. Racial disparities in health should be examined within and across regions to understand whether it is caused by less treatment for minorities in the same area or minority groups living in areas with less coverage of health care (Baicker *et al.* 2004). Racial disparities within and between regions will have implications in policy making to assure equal access to health

treatment across regions by reducing racial disparities and increasing quality of health care.

2.4 Current Interventions

Georgia *et al.* (2007) examined major health care databases and found that enhancing knowledge in treatment could help patients make proper treatment options in the health-care delivery system. A tracking and feedback registry system demonstrated a significant increase of oncology consultations by 14% and a decrease of 9% in underuse of adjuvant treatment (Bickell *et al.* 2008). This finding suggests increasing knowledge would help reducing racial disparities in breast cancer treatment. Regardless of race, poor people are more likely to have worse cancer outcomes. Public intervention programs may reduce breast cancer disparities in health for women by extending health care coverage to uninsured and poor individuals.

The increase of prevention research is positively associated with state variation of mammography use (Legler *et al.* 2002). Social support and reminders from physician were reported to improve the reschedule screening rate on time (Bobo *et al.* 2004). Advices from primary care physicians regarding the use of age-appropriate screening care may save female lives in a more cost effective way among Asian women. For example, some subgroup women including Chinese non-U.S. citizens or citizens without usual sources of health care are least likely to have mammography screening (Gomez *et al.* 2007). A retrospective cohort study of 265 patients, who underwent operative therapy at university hospital in Newark (Kim *et al.* 2008), suggests that African-American women had higher proportion of failure to complete adjuvant therapy, indicating follow-

up after treatment may improve the survival rate. Improved treatment among African-Americans is a more cost-effective approach to reduce racial disparities than increased screening (Mandelblatt *et al.* 2004).

Moreover, mammography screening comparison may provide valuable evidence as to whether practice pattern and subsequent diagnostic evaluation influence breast cancer mortality (Ballard-Barbash *et al.* 1996). Community-initiated and action-oriented research is more likely to lead to action of improving community health (Cook 2008). Research on cancer disparities can facilitate the health decision makers to design intervention programs and allocate health resources more efficiently and effectively.

2.5 Current Limitations

This section describes the current limitations on cancer disparities in the following topics: race related issues, methodological issues, and SES measurement issues. It identifies the research gaps in health disparities and points out the future research work for potential improvements.

2.5.1 Race Related Issues

The two approaches in determining race identity include self-identification and race assignment by the health administrators. However, there are some problems in the collection of race/ethnicity data and its consequent analysis: 1) growing diversity within the minority populations; 2) more non-English-speaking individuals; 3) the rising number of multi-racial individuals (Sequist and Schneider 2006). The quality of racial disparity data could affect the measurement errors such as observer bias, change in racial identity,

multiracial status, definition(s) of racial group, census undercount, and the availability of adequate data for minority groups (William 1996). The variability of race/ethnicity identification has affected the interpretation of research pertaining to racial disparities in breast cancer (Kaplan and Bennett 2003). Long *et al.* (2006) recognized the importance in handling datasets missing data in race/ethnicity and found that over 40% of 114 articles did not address issues of missing data in race/ethnicity by examining literature in the Veteran Health Administration System.

2.5.2 Methodological Issues

No consensus has been reached on the specific measurement of health disparities to be used. Careful choices have to be justified in using relative or absolute disparities, unequal weights among population groups, reference point, pair-wise or summary comparisons, favorable or adverse events, and inherent ordering of the SES groups (Keppel *et al.* 2005). Guidelines have been described clearly by the Centers of Disease Control (CDC) in order to make appropriate choices pertaining to the above six issues. In the current health inequality research, simple comparison of rates (Chu *et al.* 2007), relative risk (Hsu *et al.* 2004), and hazard ratio (Grann *et al.* 2006) are commonly used for pair-wise measures on health gradients. Percy and Keppel (2002) developed the index of disparity (ID) to summarize variation across groups such as race, education and income for a population. Their study implemented the ID measurement using several health indicators and found that cardiovascular disease mortality did not change across race/ethnicity but decreased across gender between 1989 and 1998.

Cancer disparities are recommended to be measured at both absolute and relative scales. Using lung cancer incidence, Harper *et al.* (2008) provided an overview of absolute and relative measurements by social groups and race/ethnicity groups. Percentage of population under the poverty threshold at the county level was used as a social index. Using age-adjusted prostate and lung cancer mortality in the counties of Southeastern US between 1970 and 1994, Goovaerts and his colleagues (2007) assessed the statistical performance of power and specificity of six different relative and absolute measurements across geographical regions through simulation approach. The authors concluded that two statistic methods of weighting population outperformed others. Moreover, false discovery rate approach is recommended to correct multiple testing issue involved across geographic regions. Meliker and Goovaerts (2009) applied the method into datasets of breast and prostate cancer in Michigan and evaluated the geographic scale effects on the magnitude and direction of absolute and relative measurements.

2.5.3 Socioeconomic Status Measurement Issues

Due to the unavailability of socioeconomic status (SES), few population-based studies quantitatively examined socioeconomic gradients in breast cancer incidence. Only 7% of cancer registries collect information about the education level of patients (Krieger 2002). Lack of individual socioeconomic data has been a barrier to quantify the socioeconomic effects on health. Areal-based socioeconomic measures have been a promising approach to resolve the issues of paucity of socioeconomic data in public cancer health research. Moreover, the conventional aggregated method has been used to demonstrate that social conditions had a decisive role on mortality patterns among

neighborhoods in France (Coleman 1982). In the absence of individual SES, ecologic measures have their inherited fault because SES derived from residence areas may not be representative to individual's current and historical SES. However, people living in the same area tend to have similar deprivation levels which are robust over time (Woods *et al.* 2005).

In summary, the current research about SES effects on cancer health have the following methodological issues: 1) the whole range of the SES hierarchy has not been examined extensively including the effect variation between different SES levels; 2) SES is commonly measured by single variables such as income and education; 3) SES indicators have been measured at only a specific scale, either at individual level or aggregate level; 4) the vast majority of studies have used simple correlation or regression analysis to examine primary effects of SES on health outcome. Therefore, there is a gap of knowledge regarding the appropriate measurement of SES and which index to be used to better understand how SES shapes individual and neighborhood health.

2.6 Conclusions

Monitoring and documenting disparities in cancer across race, SES, and geographic regions can provide quantitative evidence on the progress of "Healthy People 2010" in attaining its goals (Keppel *et al.* 2004). In the early twenty-first century, health disparities gained more research attention across diverse health dimensions (Redmond *et al.* 2005). Disparities research regarding breast cancer ranges from incidence, hormonal and lifestyle factors and genetic/biologic factors to treatment, survival and mortality (Brody *et al.* 2007). Most literature has been directed to health disparities of breast cancer

between white and African-American women, while the growing Hispanic population is calling for research efforts to identify and assess associated risk characteristics for its race. A variety of breast cancer outcomes may result from established and unknown risk factors among racial groups across regions. Compared to white, early child-bearing may be a protective factor for Hispanic population, while the religious and cultural beliefs and lower SES may be risk factors for the more advanced carcinoma stage. Institutional, cultural and linguistic barriers are important factors for Asian women who are reported not to follow mammography use recommendations in addition to their lower SES (Gomez *et al.* 2007).

In terms of racial disparity assessment, public health researchers confront challenges due to data availability, difficulty in race/ethnicity identification, sample size limitation, Modifiable Area Unit Problem (MAUP), availability of cultural and socioeconomic data at the individual level (Sequist and Schneider 2006). The maturity of geographic information techniques offers a new and different platform for research related to health disparities. Disentangling regional variation from racial disparities with regard to breast cancer can provide a guideline for well-organized prevention programs. Even though areal-based measurement approach overcomes the barrier of individual SES information shortage, it inevitably introduces MAUP, a common issue for geographic scale research.

CHAPTER III

DATA AND METHODS

This chapter presents the detailed description of the data and methodology in this study. Section 3.1 describes the data sources of the geography and population, breast cancer incidence, and breast cancer mortality. Section 3.2 describes the methodology of this research. The approach of racial disparity measurement was discussed to quantify the absolute and relative differences of cancer rates adjusted by the population size. The linear regression and logistic regression were further elaborated to explain the predictors which were adopted to model racial disparities. Local indicator of spatial autocorrelation was used to test spatial dependency within the spatial datasets and to provide the evidence of a needed geographically-weighted scheme in a regression analysis. Factor analysis was utilized to detect the underlying structure in order to minimize the multicollinearity essential for a regression analysis.

3.1 Data

3.1.1 Geography and Population

Texas is the largest state in the contiguous United States with an area of 268,820 square mile. The estimated population was 24,236,974 in 2008 with an annual increase

of 2.0% since the year 2000. The population density is 34.8 people/square kilometer in 2008. In 2006, the demographic distribution was estimated as follows: 70.6% white, 11.5% African American, 3.3% Asian American, 0.5% Native American, and 12.3% other racial group; by ethnicity the population was categorized as 35.5% Hispanic or Latino and 64.5% non-Hispanic in Texas.

The cartographic boundary shapefile was retrieved from the US Census 2000 (Figure 3.1). Census tracts are denser within metropolitan urban areas such as Houston, Dallas, and the Austin-San Antonio area. With an average of 4,000 people in each unit, Census tract is designed to contain homogenous populations in terms of socioeconomic composition within the neighborhood. Census tract is also a common administrative geographic unit used to determine eligibility and resource allocation for diverse programs. A total of 4,388 census tract were indentified in Texas for the year of 2000.

Female population by race and age was derived from the 2000 Census Summary File 1 (SF 1) 100-Percent Data. Race is defined as specific physical, hereditary and cultural traditions within social-political construct, while ethnicity is categorized by one of two ethnicities: Hispanic or Latino, and not Hispanic or Latino. Three races/ethnicities groups of non-Hispanic white, African-American, and Hispanic were considered in the study. In 2000, the female population for non-Hispanic white was 5,555,694; that of African-American was 1,244,302; and the female population of Hispanic or Latino was 3,273,458. The population was grouped as five categories of age 1-14, 15-24, 25-44, 45-64, and 65 above to calculate age-adjusted cancer rates.

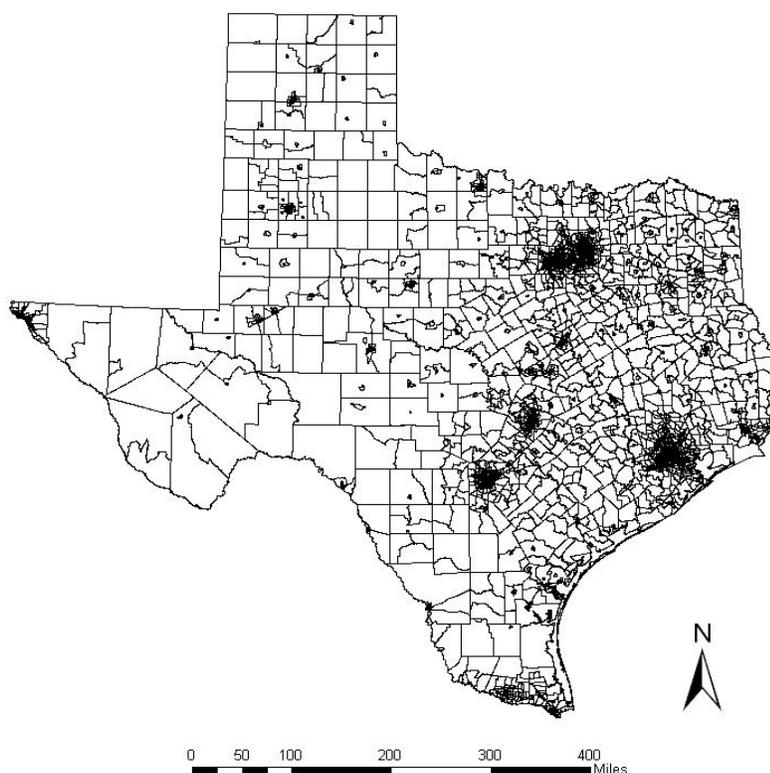


Figure 3.1 Census tracts in Texas.

3.1.2 Data of Breast Cancer Incidence

The study aims to examine racial disparities of breast cancer from 1995-2005 in Texas. Breast cancer incidence data were provided by the Texas Cancer Registry, Texas Department of State Health Services. This dataset includes all diagnosed breast cancer cases with various characteristics of stage, resident address, diagnosed date, race, and Hispanic origin. The Surveillance, Epidemiology, and End Results (SEER) Summary Staging, as one of the most fundamental staging systems, categorizes cancer cases based on how far it spreads from its original point (Young 2001). According to SEER Summary Stage Coding Scheme (National Cancer Institute 2001), breast cancer stages are coded as follows: 0-in situ; 1-localized; 2-regional by direct extensions; 3-regional to lymph

nodes; 4-regional (direct extension and lymph nodes); 5- regional, not otherwise specified; 7- distant metastasis; 9-unstaged, unknown, unspecified. In situ refers to that the carcinoma cells are still retained within the cell groups. A localized cancer is constrained to the organ of breast and has no spread further than the breast tissue. Regional cancer extends beyond the limits of the breast organ. Distant stage is tumor cells metastasizing to other body parts (Shambaugh and Platz 1977). Early and late stages in breast cancer are general terms defined using many clinical and pathological staging systems. In this study, codes “0” and “1” indicate early stage. It is difficult to distinguish regional and distant cancer. Regional and distant stages with codes “2-3” are defined as late stage (DeChello and Sheehan 2007). Breast cancer is reported at unknown-stage, which is excluded in this study.

About 150,087 total cases among all races were reported with the exclusion of 190 records from Veteran’s Administration facilities due to confidentiality issues. In addition, 437 cases with incomplete address information were excluded from this study as they were not successfully geocoded based on information provided from the original dataset obtained at the Texas Cancer Registry. Thus, 92,655 cases (61.91%) are identified as early-stage breast cancer, 44,515 (29.75%) as late-stage, and 12,480 (8.34%) as stage-unknown/unspecified. Breast cancer varies by race and stage. About 109,453 non-Hispanic white women developed breast cancer; 15,147 African-Americans and 21,211 Hispanics were diagnosed. Moreover, 2,896 cases (1.94%) and 943 cases (0.63%) were identified with breast cancer for other and unknown racial groups respectively, which was not considered in this study. Table 3.1 presents the number and proportion of breast cancer cases by race and stage. Non-Hispanic white women were 11.72% and 9.98%

more likely to be detected as the early-stage than African-American and Hispanic females. On the other hand, late-stage breast cancer was 10% more prevalent in these two minority groups relative to their white counterpart.

Table 3.1 Number and proportion of breast cancer incidence cases by race and stage

Race	Early-Stage		Late-Stage		Unknown-Stage		Total
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
non-Hispanic white	70,627	64.53	30,210	27.6	8,616	7.87	109,453
African-American	7,999	52.81	5,605	37.00	1,543	10.19	15,147
Hispanic	11,570	54.55	7,755	36.56	1,886	8.89	21,211
Total	90,196	61.86	43,570	29.88	12,045	8.26	145,811

3.1.3 Data of Breast Cancer Mortality

Breast cancer persists as the leading cause of cancer death among women in Texas. The mortality data of female breast cancer were provided by the Center of Health Statistics, Texas Department of State Health Services (TDSHS). Demographic information and causes of death were collected in death records by the Vital Statistics Unit from TDSHS (Risser *et al.* 2009). A total of 27,162 death cases were reported in Texas through the year of 1995-2005. All cases had information on street address, age group, year of death, race, Hispanic origin, and geographic location. In this study, 3,228 cases (about 11.88%) were disqualified as a result of geocoding process failures owing to lack of complete address information. Table 3.2 displays the geocoding results across races/ethnicities. The matching rate was slightly different among racial groups. African-American and Hispanic groups had as much as 3-4% more cases to be successfully

geocoded in comparison to non-Hispanic white females. To examine the age-adjusted mortality rates by races, the US 2000 Standard Million population was employed to take into account the heterogeneity of age distribution.

Table 3.2 Geocoding results among race/ethnicity groups

	Non-Hispanic white	African-American	Hispanic	Total
Matched	16,419	3,616	3,655	23,690
Unmatched	2,458	350	412	3,220
Matching rate	86.98%	91.17%	89.87%	88.03%

3.2 Methods

3.2.1 Disparity Measurement Approach

African-Americans and Hispanics tend to live in different areas than the non-Hispanic white population. Hence, geographical disparities in the level and quality of offered health care can drive reported racial/ethnic disparities (Baicker *et al.* 2004). However, most studies used national samples which masked geographic variations in health. Identification of geographic disparities in late-stage diagnosis and mortality of breast cancer provides quantitative assessment that helps health policy-makers to determine the effectiveness of intervention programs and health resources at the local level.

According to a report from the Surveillance Research Program (SPR) and the Applied Research Program (ARP) of the Division of Cancer Control and Population Sciences of the National Cancer Institute (Harper and Lynch 2006), pair-wise absolute

and relative comparisons for cancer data suffice for the comparison of specific groups. Health differences can be measured at relative and absolute scales, which are the primary means to enumerate effect sizes across diverse social and racial groups. A relative measure articulates the rate differences against a reference point. An absolute measure provides a simple arithmetic difference between a target group and a reference group (Keppel *et al.* 2005). Thus, relative and absolute measures provide fundamentally different information with the possibility of arriving at different conclusions. Relative disparity cannot reflect the variation patterns in health in absolute measure. So the rate difference (absolute measurement) and rate ratio (relative measurement) were employed to investigate racial disparities in breast cancer in the study. Rate difference provides arithmetic difference in health outcomes between two racial or social groups and rate ratio is simply a ratio between health outcomes measured on target (e.g. minority) and reference groups (Harper *et al.* 2008). The non-Hispanic white, the most favorable group is commonly used as a reference point which all groups are desired to achieve.

Accounting for population size in the computation of the disparity statistics allows one to allocate greater weights to racial or socioeconomic groups with larger population. In addition, a population-weighted scheme corrects for the small number problem that is often observed for minority groups and smaller geographic units. Fleiss (1981) and Lachin (2000) proposed four different population-weighted statistics in absolute scales and two different ones in relative scales. Goovaerts *et al.* (2007) assessed the six statistics through a simulation approach mimicking different scenarios in terms of the magnitude and frequency of disparities. They identified two test statistics (Equations 1 and 3 below) that had higher power and created fewer false positives using prostate and

lung cancer mortality datasets for the period of 1970-1994 in the Southeastern part of the United States. The power measures the probability of correctly detecting significant disparities, while a false positive corresponds to the situation where a racial disparity is wrongly declared significant. The statistic to measure absolute differences between two racial groups is as follows:

$$Disp_{Abs}(u_i) = \frac{r_2(u_i) - r_1(u_i)}{\sqrt{\bar{r}(u_i)(1 - \bar{r}(u_i)) \left[\frac{1}{p_1(u_i)} + \frac{1}{p_2(u_i)} \right]}} \quad (1)$$

where $\bar{r}(u_i)$ is the population-weighted average of cancer rates computed as:

$$\bar{r}(u_i) = \frac{p_1(u_i)r_1(u_i) + p_2(u_i)r_2(u_i)}{p_1(u_i) + p_2(u_i)} \quad (2)$$

The statistic to measure relative differences between two racial groups is as follows:

$$Disp_{Rel}(u_i) = \frac{\log\left(\frac{r_2(u_i)}{r_1(u_i)}\right)}{\sqrt{\frac{(1 - r_2(u_i))}{p_2(u_i)r_2(u_i)} + \frac{(1 - r_1(u_i))}{p_1(u_i)r_1(u_i)}}} \quad (3)$$

In the above expressions, $p_1(u_i)$ and $r_1(u_i)$ denote the population size and mortality rate of the reference group in region u_i (i.e. non-Hispanic white population), while $p_2(u_i)$ and $r_2(u_i)$ are the same quantities measured for the disadvantaged racial groups including African-American and Hispanic. Region u_i represents any geographic units of census tracts. The two statistics in equations (1) and (3) assume a normally distributed dataset.

The null and alternative hypotheses for testing whether the difference in health outcomes between two racial groups is significant are as follows:

The hypotheses for absolute difference

$$H_{1A} \quad H_0 : |RD(u_i)| = |Disp_{Abs}(u_i)| = 0$$

$$H_{1B} \quad H_1 : |RD(u_i)| = |Disp_{Abs}(u_i)| \neq 0$$

The hypotheses for relative difference

$$H_{2A} \quad H_0 : |RR(u_i)| = |Disp_{Rel}(u_i)| = 1$$

$$H_{2B} \quad H_1 : |RR(u_i)| = |Disp_{Rel}(u_i)| \neq 1$$

The significance (p) of the above two test statistics can be assessed by comparing the test statistic against its expected distribution under the null hypothesis of equality of rates among all race/ethnicity. However, hundreds of individual tests might need to be conducted, particularly when small geographic units are analyzed. Correction for multiple testing is thus essential to avoid overestimating the proportion of significant disparities (i.e. high likelihood of erroneous alarms). The procedure of false discovery rate (FDR), which controls the expected amount of factual null hypotheses out of the total number of rejections, was implemented because of its less restriction and more power than other correction methods (Benjamini and Hochberg 1995).

A two-tailed test was performed with a critical α level equal to 0.05 and FDR correction using the Space-Time Intelligence System (Avruskin *et al.* 2004). Non-Hispanic white population was used as the reference population, which means that a positive rate difference (RD) indicates higher cancer rates for the African-American and Hispanic women. Similarly, if rate ratio (RR) exceeds 1, the minority women experience worse health outcomes than their non-Hispanic white counterparts. This study reports all

geographic units where absolute and relative racial disparities in late-stage diagnosis and mortality of breast cancer that were tested significant. The geographic unit is excluded from the RD/RR statistics and subsequent analysis if the population of any ethnic group is equal to 0. The geographic units were not considered as well when the number of cases is zero for both ethnic groups when measured in RD and for either ethnic groups when measured in RR.

3.2.2 Regression Methods

Multivariate regression method is the most common approach exploited to assess a dependent variable in relation to a set of explanatory variables (i.e. the independent variables) (Devore 2003; Zar 1999). Different approaches can be used to derive the best fitting regression model. Linear and logistic regressions were adopted to analyze the spatial connection of racial disparities of breast cancer between late-stage diagnosis and mortality rates as well as to investigate the significance of racial disparities in breast cancer mortality using other predictors. A linear regression model predicts the quantity of a dependent variable based on a number of independent variables. There are three primary assumptions for linear regression: 1) all observations must be independent, 2) all variables must be normally distributed, and 3) variances must be similar across both dependent and independent variables (Reminton and Schork 1985). A simple linear regression can be established as follows:

$$y_i = \sum_{j=0}^M \beta_j x_{u_{i,j}} + \varepsilon_i \quad (4)$$

Where u_i is observation at that location i (i.e. census tract), j is the number of independent variables. β_j is the coefficient of the independent variable j in the regression. ε_i is the residual variable at location u_i . The regression parameters were estimated using a maximum likelihood approach (Aldrich 1997). The significance of the overall model was assessed by comparing the test statistic against the F distribution (Kutner *et al.* 2004).

Aspatial regressions assign equal weight to all the data regardless their geographical locations and lead to a single regression model for the study area. However, traditional regression models are not sufficient for analyzing geographic datasets because they ignore the dependence of observations in space. Waldo Tobler formulated the first law of geography as “everything is related to everything else, but near things are more related than distant things” (Tobler 1970, 236). Geographical weighting of the data is required to consider the spatial dependency by fitting a single intercept and slope across multiple overlapping neighborhoods in a spatial dataset (Fotheringham *et al.* 2002).

The geographically weighted regression (GWR) was proposed by Brunson *et al.* (1996) and modified by Stewart Fotheringham and Martin Charlton (2002). In a GWR regression, each observation is assigned a weight based on its proximity to the center of the local overlapping windows centered on the location u_i . In this study, a linear regression is constructed for each focal point of census tracts by weighting the neighborhoods. Neighbors in a spatial weight can be defined by the number of nearest neighbors and distance range. Weighting schemes used in the study is the Gaussian kernel function, which allows flexible spatial contexts according to the number of neighborhood specified. When the spatial features are denser, the spatial context is

smaller; vice versa. The Gaussian kernel bandwidth varies across space with the decay function of weighting neighborhoods. The weight is diminishing as the distance increases from focal points to neighborhoods. The spatial weighting algorithm of Gaussian function is formed in equation (5).

$$w_{ij} = \exp(-d_{ij}^2 / b^2) \quad (5)$$

Where w_{ij} is the weight assigned at location i based on neighborhoods j , d_{ij} is referred as the distance from the local point of i to neighborhood of j , and b represents the kernel bandwidth. The optimal bandwidth can be determined by minimizing Akaike Information Criterion (AICc). The detailed description AICc can be found in the book written by Sakkamoto and colleagues (1986).

This research utilized the traditional and GWR linear regressions to assess the spatial connections of racial disparities of breast cancer between late-stage diagnosis and mortality rates. The normal-transformation was applied to the test statistics of the two types of racial disparities, which was obtained from the racial disparity analysis. The normal-transformation was done by ranking dataset values from lowest to highest and assigning the values of normal distribution according to their ranks (Goovaerts and Jacquez 2004). The determination coefficient (R^2) in the regression analysis indicates how much variance in the dependent variable could be explained by the independent variables. The correlate coefficient is another indicator of the association of dependent and independent variables.

A logistic regression has been developed for predicting the binary dependent variable according to independent variables which could be categorical or numerical data

(Zar 1999). Logistic regression has two assumptions of independent observations and linear relationships between independent variables and the log odds of the dependent variable. In this study, significance of a census tract in racial disparities of breast cancer mortality can be assessed using a bivariate or multivariate logistic regression.

Significance test determines if a census tract is significant in racial disparities of breast cancer in late-stage diagnosis and mortality rates. The logistic regression can also be applied to evaluate the association of significance of racial disparities with the socioeconomic level for each census tract. In a logistic regression, the probability of dependent variable coded as 1 could be established by the following function.

$$p_{u_i} = \frac{1}{1 + e^{-\sum_{j=0}^M \beta_j x_{u_i j}}} \quad (6)$$

In the above function, p_{u_i} is the probability of a "1" at location u_i . β_j is the coefficient of the variable j in the regression. e is the natural logarithm. The parameters are estimated with the goal of maximizing the log-likelihood estimation, which is a repetitive process. The significance of the full model and individual terms are assessed using a chi-square distribution (Devore 2003). Logistic regression provides the results of the parameter estimates, the significance of parameters, and odds ratio along with confidence interval. The odds ratio can be computed by the power of parameter estimates with the base of natural log. Two R-squares of Cox and Snell and Nagelkerke are calculated in a logistic regression model in STIS. Nagelkerke R square is similar to the adjusted R square in a linear regression. The Gaussian weight method allows logistic regression to account for the spatial dependency for geographical datasets. However, the GWR logistic regression has the possibility with failure of convergence in the

determination of model coefficients using the estimation of Maximum Likelihood Approach. In particular, logistic regression could not converge if there are a larger number of missing values and smaller count of “1s” for binary dependent variable.

3.2.3 Local Indicator of Spatial Autocorrelation

To detect if the residuals in either linear regressions or logistic regressions have any spatial autocorrelation, local indicator of spatial autocorrelation (LISA) statistics was adopted in the study. LISA statistics, known as Local Moran’s I (Anselin 1995), were utilized to identify the adjacent areas with similar values (clusters) or dissimilarity of neighbour values (outliers). Local Moran’s I analysis provides an index of I and Z score for each spatial feature at local scale by decomposing the global Moran’s I (Moran 1950). Index I indicates the strength of the similarity of each feature with its surrounding features, while Z score represents the statistical significance for each I value at local scale. Local Moran’s I can be calculated as follows.

$$I_{u_i} = (x_{u_i} - \bar{x}_{u_i}) \sum_{j=1}^{j=n} w_{u_{ij}} (x_{u_j} - \bar{x}_{u_i}) \quad (7)$$

where I_{u_i} is the local Moran’s index denoting the strength of similarity at location u_i .

x_{u_i} and x_{u_j} is a pair of values tested for location u_i and u_j which includes all locations falling within a searching window radiating from a focal point of location u_i .

\bar{x}_{u_i} represents average of all values within the searching window of location u_i . In addition, $w_{u_{ij}}$ denotes the spatial weight signifying the strength of connection between locations u_i and u_j . The null hypothesis for Local Moran’s I is stated as follows.

Null hypothesis: there is no association between the value observed at a location and the values around its surroundings.

$$H_0 : I_{u_i} = 0 \quad (8)$$

Index I_{u_i} ranges from -1 to +1 in the equation 8. Large positive index represents that the region I_{u_i} have stronger similar values as its near neighbors, while negative values suggests the dissimilarity of values at location u_i and its surroundings. The Moran's I was performed in STIS to evaluate the significance of spatial autocorrelation with Monte Carlo randomizations.

3.2.4 Factor Analysis

Factor analysis is an approach to identify the composite factors underlying all observed variables. This approach is typically used to reduce the multicollinearity and the number of variables by making possible combinations among the observed variables. In this study, principal component analysis (PCA) was utilized to extract the common factors underlying the independent variables before performing regression (Jolliffe 2002). Principal component analysis was developed by Karl Pearson in 1901 as a standard tool of revealing the unobserved composition using minimum dimensions to explain the most variance within a dataset. Principle component analysis has stringent assumptions including a linear combination of its basis vector to account for the variance within the original database.

For a matrix of X_{mn} , m is the number of measurement types and n is the number of samples. Each row of a matrix X_{mn} corresponds to observations at all locations for a

particular type of measurement and each column represents all types of measurements at a particular observed location. The covariance for the matrix C_X is defined as a measurement of the linear relationship between the m measurement types. The covariance of matrix C_X is expressed as a dot product matrix computation with the written form as follows.

$$C_X \equiv \frac{1}{n} XX^T \quad (10)$$

X^T is the transpose matrix of X . The principle algorithm then applies an orthogonal or oblique rotation matrix (P) on the base vector to maximize its variance in m -dimension space. The principle components in matrix P are order-ranked based on the total variance explained by each component. The matrix P is equal to the eigenvectors of C_X (Shlens 2005).

This study applied an exploratory factor analysis which assumes no a priori framework in grouping the variables. PCA was conducted to reduce the correlation among the independent variables for subsequent regression analysis. Because the PCA is based on the linear relationships, screening the data is critical step to identify the need of data transformation before performing PCA. This study adopted the normal transformation by ranking the data from minimum to maximum and matching them to the value in the normal transformation. In this study, factor loading greater than 0.6 was utilized to extract the underlying constructs in addition to the light of theory (Hair *et al.* 1998). The detailed description of PCA can be found in the book written by Jolliffe (2002).

CHAPTER IV

SIGNIFICANCE TESTS AND SPATIAL RELATIONSHIP OF RACIAL DISPARITIES IN LATE-STAGE DIAGNOSIS AND BREAST CANCER MORTALITY

This chapter provides the results of significance tests of racial disparities in late-stage diagnosis and breast cancer mortality at the census tract level for both African-American and Hispanic women. This part intends to answer the research questions I, II and the first part of research question III that are elaborated in Chapter I. Research questions I and II attempt to identify the census tracts with significant racial disparities of late-stage diagnosis and breast cancer mortality. The first part of research question III is to examine the spatial relationship of racial disparities between late-stage diagnosis and breast cancer mortality. After mapping all late-stage incidence and mortality cases using the software package of ArcInfo 9.3, the age-adjusted rates were calculated for each census tract in SPSS 16.0 using the population of 2000 Census. Non-Hispanic whites served as a reference point to compute the statistics of rate difference (RD) and rate ratio (RR) for African-American and Hispanic women at each census tract. Significant tests were corrected with a false discovery rate (FDR) approach in the software package STIS 1.8.

The spatial relationship of the two racial disparities in breast cancer is explored as well in this chapter. The visualization of census tracts significant in both racial disparities indicated that the number and locations of these two disparities were not consistent. The confirmatory analysis based on the normalized statistics of racial disparities suggested the strong association occurred in the linear and logistic regression. Local indicators of spatial autocorrelation found that the residual of these regression models had spatial non-stationary characteristics. The geographically weighted regression improved the performance of the regression models and revealed the spatial variance underlying the connection of racial disparities in late-stage diagnosis and breast cancer mortality.

This chapter is organized as follows. Sections 4.1 and 4.2 describe the results of significance tests of racial disparities in late-stage diagnosis and breast cancer mortality among African-American and Hispanic women. Section 4.3 analyzes the spatial relationship of these two disparities using confirmatory data analysis approach.

4.1 Significance Tests of Racial Disparities in Late-stage Diagnosis of Breast Cancer

4.1.1 Late-stage Diagnosis Rates of Breast Cancer

As the result of an uneven distribution of population and cases among the three racial groups, not all census tracts in Texas (4,388) had residents and breast cancer cases for each race. Table 4.1 lists the statistics of census tracts with non-zero population and late-stage breast cancer cases at the census tract level. The ratio of number of census tracts with non-zero population to late-stage diagnosis was expressed as a percentage. In Texas, about 90% of census tracts with non-Hispanic white residents reported at least one

late-stage diagnosis case, while 42% and 56% were observed for African-American and Hispanic women respectively. These results indicated that the much smaller proportion of census tract units had at least one late-stage diagnosis for African-American and Hispanic females. Moreover, more census tracts were observed to have at least one late-stage diagnosis of breast cancer for Hispanics than for African-Americans.

Table 4.1 Number of census tracts with non-zero population and non-zero cases of breast cancer late-stage diagnosis by race

		Non-Hispanic white	African- American	Hispanic
Number of Census Tracts (n=4,388)	Population	4,382	4,322	4,375
	Late-Stage	3,925	1,818	2,447
	Percentage	89.57%	42.06%	55.93%

All cases from 1995-2005 were aggregated for each race to determine late-stage incidence of breast cancer in Texas. Late-stage diagnosis refers to distant and regional stages of breast cancer among women. The 2000 population obtained from the U.S. Census Bureau was utilized as the denominator of cancer rates which were adjusted based on the 2000 US Standard Million Population. Although the 2000 population may introduce inaccuracy of breast cancer rates for the study period of 1995-2005, over/under estimation of population before and after 2000 may improve the accuracy of rates calculation for the whole study period by combining all cases together. The rate and proportion of late-stage diagnosis in Texas were computed for the 18 age groups of 0-4, 5-9, 10-14....80-84, and 85+ for non-Hispanic white, African-American and Hispanic women.

Figure 4.1a displays late-stage diagnosis rates by age and racial groups. All women experienced close to zero late-stage incidence rates in their middle twenties. Then late-stage rates of breast cancer had an increase trend by age groups among non-Hispanic white, African-American, and Hispanic females. However the rates would never cross with each other with the highest rate for African-American women followed by non-Hispanic whites and Hispanics. On the other hand, non-Hispanic whites and African-Americans reached the peak of late-stage incidence rates at the age 70-74. More specifically, non-Hispanic whites had the late-stage rates of 128.60 per 100, 000 while African-Americans experienced the rates of 140.08 per 100, 000 at the age group of 70-74. Furthermore, Hispanic women displayed slightly oscillating late-stage rates after the age group of 55-59 and approached the peak of 102.39 per 100,000 women at the age 80-84.

The proportion of late-stage diagnosis was computed for each race as the total number of late-stage cases in breast cancer for each age group divided by the total number of late-stage cases for all age groups. Figure 4.1b illustrates the proportion of late-stage diagnosis by age group and race. Interestingly, African-Americans and Hispanics showed very close patterns of proportion of late-stage diagnosis by age group, although overall Hispanics had the lowest late-stage incidence rates. At the age 40-44, both African-Americans and Hispanics reached their proportion peak of late-stage diagnosis of 14.20% and 14.31% respectively. Smaller proportion of non-Hispanic white women experienced late-stage diagnosis at the age younger than 50-54, while more cases for non-Hispanic whites were diagnosed late-stage at the age group older than 55-59.

These results showed that African-American and Hispanic females were diagnosed in a more severe stage of breast cancer among the younger age group.

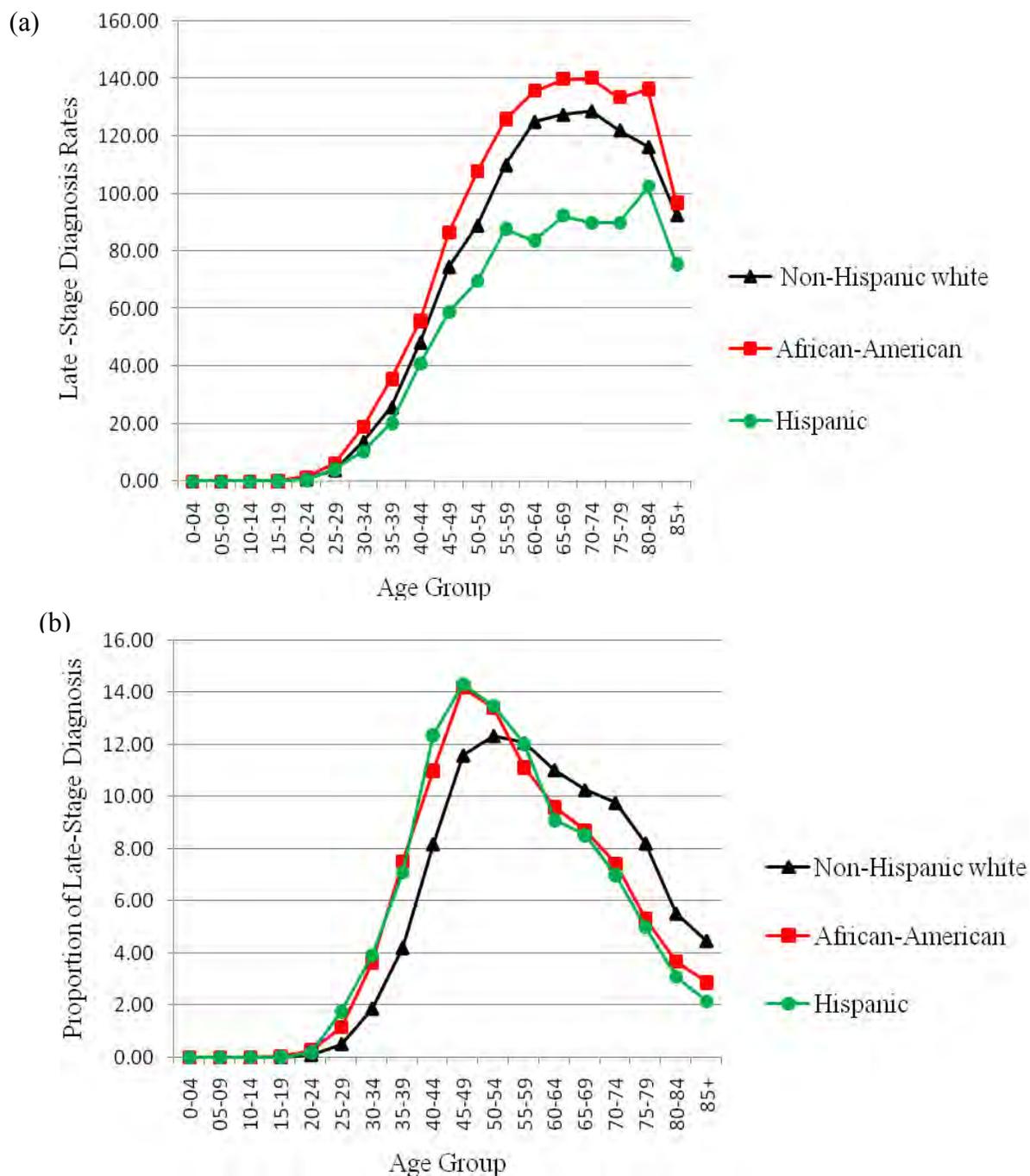


Figure 4.1 Late-stage specific incidence rates per 100,000 (a) and proportion of late-stage cases (b) for breast cancer by race and age group.

4.1.2 Racial Disparities of Breast Cancer Late-stage Diagnosis among African-American Women

The methodology to identify significance of racial disparities was discussed in Chapter III. The late-stage incidence rates were computed for all the racial groups under study based on the adjustment of five age groups of 0-14, 15-24, 25-44, 45-64, and 65+ and 2000 US Million Population. These age groups were selected for age-adjustment rates so as to minimize the small number issue caused by calculating the cancer rates at such a small geographical level of census tract.

Table 4.2 lists the number of geographic units identified as experiencing significant higher late-stage diagnosis for African-American women in comparison with non-Hispanic whites using RD measurement. If RD is greater than 0, it indicates that minority has significant higher rates compared to non-Hispanic whites, otherwise significantly lower rates. Thus, 188 census tracts out of 4388 exhibited significant racial disparities in terms of rate difference (RD) statistic (Figure 4.2 a). Most census tracts that showed significant in racial disparities of late-stage diagnosis were observed in the downtown areas of Houston, Dallas, and Austin-San Antonio. When zoomed, Houston and Dallas had more census tracts testing significant in racial disparities of late-stage diagnosis than the Austin-San Antonio area.

Geographic units were further classified according to their level of socioeconomic status (SES). The poverty line is defined by combining household income and family size. For instance, the poverty line was defined as \$17,463 for a family size of four, including two children under 18 years in 2000 (U.S.Census Bureau 2009). Low SES is defined to have more than 20 percent of population living under poverty line, middle SES and high SES are referred to have 10%-20% and less than 10% population living below

the federal poverty line. The association between low SES and significant racial disparities of late-stage diagnosis was more apparent in the inner center of the three metropolitan areas. About 79% of census tracts (149 out of 188) tested significant in racial disparities of late-stage diagnosis to have a designated low SES.

When the RR measurement was utilized, nine census tracts were identified as having significant racial disparities in late-stage diagnosis (Table 4.2). RR greater than 1 suggests that the target groups (i.e. the minorities) were found to have significantly higher cancer rates compared to the reference group (i.e. non-Hispanic white). In contrast to the RD, six out of nine census tracts having significant racial disparities showed as the high SES level.

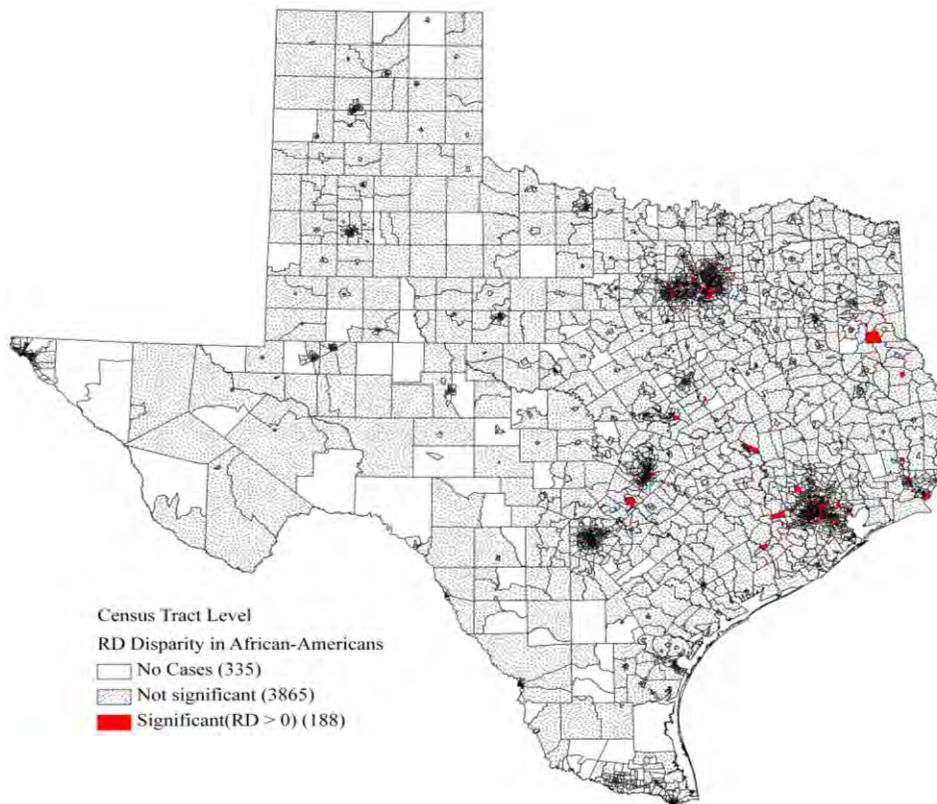
Table 4.2 Number of census tracts with significant racial disparities in breast cancer late-stage diagnosis for African-American women

Poverty Level		High SES (0-9.99%)	Middle SES (10.00-19.99%)	Low SES (≥20.00%)	Total
African-Americans	RD < 0	0	0	0	0
	RD > 0	13	26	149	188
	RR < 1	0	0	0	0
	RR > 1	6	2	1	9

Figure 4.2b illustrates these census tract locations on map. The outskirts of Houston were found to have three census tracts with significant racial disparities and low SES level. The Dallas metropolitan area had four census tracts where African-Americans experienced significantly higher late-stage diagnosis of breast cancer. In addition, one census tract was located in Hays County with low SES and the other one was found in Tom Green County with the middle SES, which located in the central Texas. One explanation is that these census tracts have a smaller number of African-American

residences and one or two late-stage cases could inflate the incidence rates dramatically. In contrast to non-Hispanic white women, none of the census tracts displayed significantly less late-stage diagnosis for African-American women because no census tracts were found either RD less than 0 or RR smaller than 1 (Table 4.2).

(a)



(b)

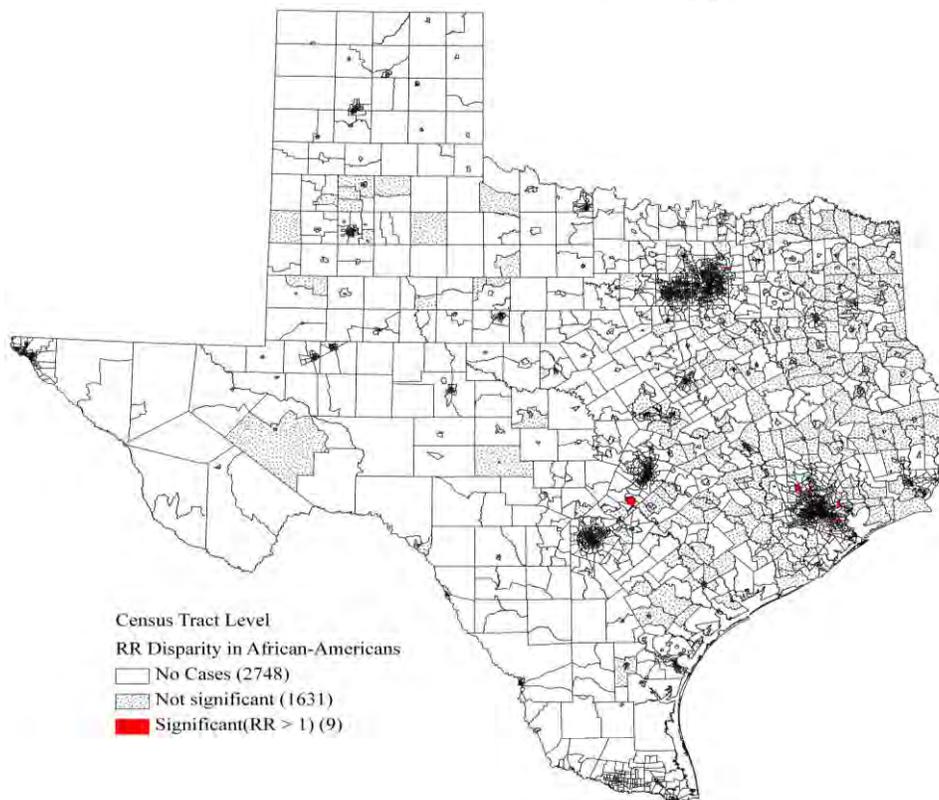


Figure 4.2 African-American breast cancer late-stage diagnosis: significant racial disparities according to the RD (a) and RR (b) statistics.

4.1.3 Racial Disparities of Breast Cancer Late-stage Diagnosis among Hispanic Women

Table 4.3 summarizes the number of census tracts that tested significant in racial disparities of late-stage diagnosis for Hispanic women compared to non-Hispanic whites. The geographic units were classified into low, middle, and high SES as well based on the percentage of population living under the poverty line. Hispanics were found to have significantly higher late-stage diagnosis than non-Hispanic white females within 266 census tracts based on the rate difference (RD) statistics. These census tracts were not only located in the Metropolitan areas of Houston, Dallas, and Austin-San Antonio, but also were found along the Southwest border of Texas (Figure 4.3a). Hispanic populations were concentrated on the above areas and had low socioeconomic status. About 88.34% (235 out of 266) of these census tracts tested significantly higher late-stage diagnosis for Hispanics and had low SES level of more than 20.00% population living under the federal poverty line. In terms of the RR statistics, only two census tracts were detected as displaying significant racial disparities (Figure 4.3b). Both of the two census tracts have a poverty level greater than 10.00%. Hispanics were not observed to have significantly lower late-stage diagnosis for both RD and RR measurements (Table 4.3).

Table 4.3 Number of census tracts with significant racial disparities in breast cancer late-stage diagnosis for Hispanic women

Poverty Level		High SES (0-9.99%)	Middle SES (10.00-19.99%)	Low SES (\geq 20.00%)	Total
Hispanics	RD < 0	0	0	0	0
	RD > 0	8	23	235	266
	RR < 1	0	0	0	0
	RR > 1	0	1	1	2

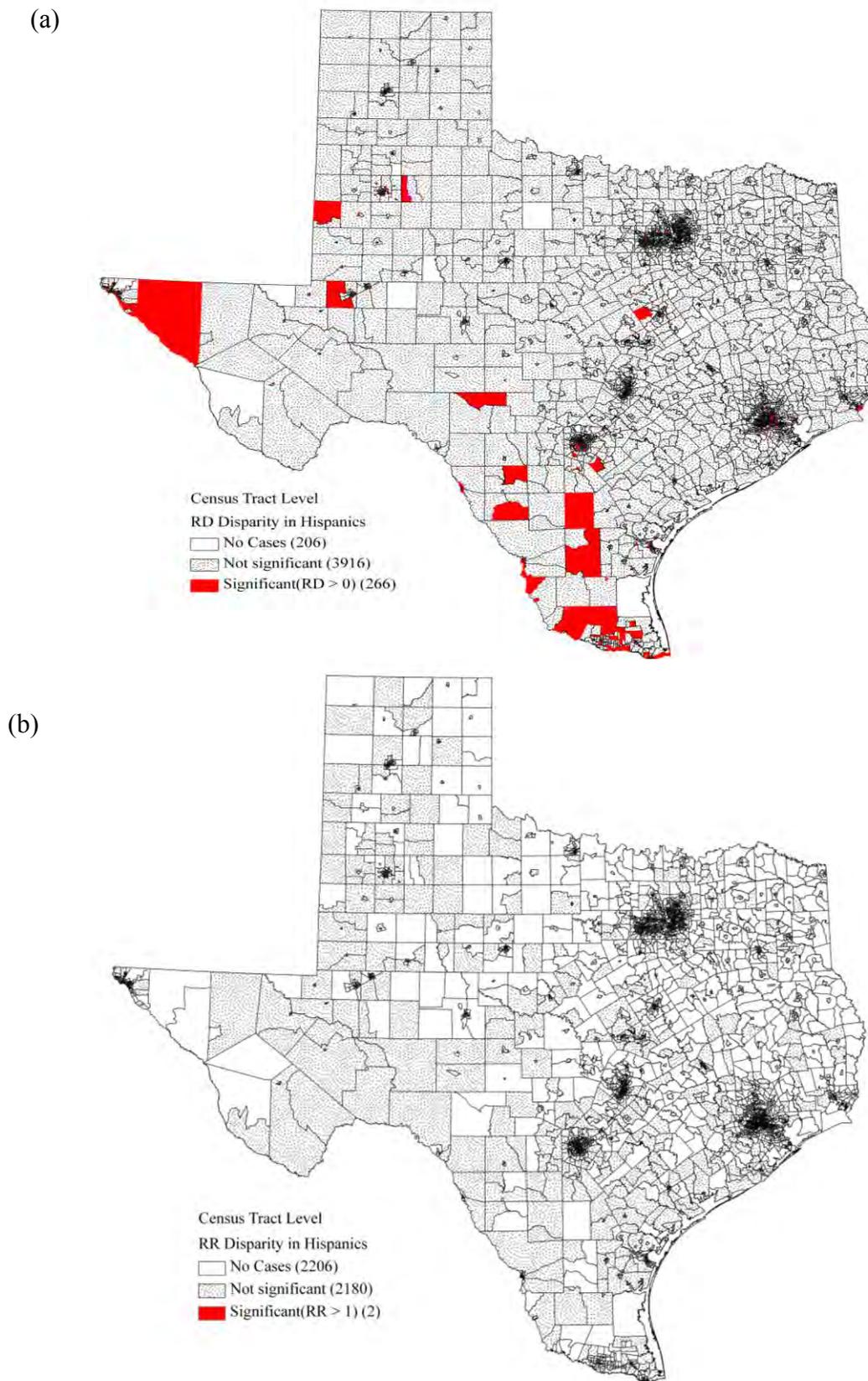


Figure 4.3 Hispanic breast cancer late-stage diagnosis: significant racial disparities according to the RD (a) and RR (b) statistics.

4.1.4 Discussion on Racial Disparities of Breast Cancer Late-stage Diagnosis

In the study, the proportion of stage by age and race illustrates that women in the younger age group experienced more late-stage breast cancer for both African-Americans and Hispanics than the older age group (65 and older). The phenomenon could be explained by the proposition that younger women are inclined to develop more aggressive cancer lumps and are less likely to be screened (Chu *et al.* 2001; Russell *et al.* 2003; Wells and Roetzheim 2007). On the other hand, women older than 65 have better access to health care through Medicare program and the higher utilization of mammograms, which leads to a decrease in late-stage diagnosis for older women (Potosky *et al.* 1993).

This study identified the number and locations of census tracts that tested significant in racial disparities of breast cancer late-stage diagnosis for both African-American and Hispanic women. The results were not consistent based on the RD and RR statistics. Many more census tracts were identified significant using RD statistics than RR in that the number of census tracts with missing values differed considerably for the two measurements due to their measurement essences. RR statistic generally leads one to discard more geographic units because the measurement assigns any census tracts as missing values if they have no population or no cases for either target or reference groups.

A large number of census tracts were found to have significantly higher late-stage diagnosis rates for African-American and Hispanic women in comparison with non-Hispanic white females. Richardson and colleagues (1992) found that African-Americans and Hispanics were 1.29 and 1.32 folds greater risk to diagnose in the advance stage of

breast cancer in Los Angeles, California than White non-Hispanics, when controlled for socioeconomic status. Hispanics are often reported to have lower incidence rates than White non-Hispanics (Risser *et al.* 2009). However, the incidence rate ratios for Hispanics were observed significantly lower in the early detection of breast cancer than non-Hispanic whites, especially within the age group under 50 (Bentley *et al.* 1998).

This study concluded that more census tracts had significantly higher late-stage diagnosis for Hispanics than for African-American women compared with non-Hispanic whites in Texas from 1995-2005. This could be explained by the higher percentage and wider spatial distribution of Hispanic populations in the Metropolitan areas of Houston, Dallas and San Antonio as well as the Southwest border of Texas. Significantly higher proportions of census tracts tested significant in racial disparities of late-stage diagnosis were fallen into the low SES level of more than 20% population living under the poverty line. The strong connection of racial disparities in late-stage diagnosis and low SES status may reflect the obstacles to accessing limited health resources and lack of financial support for minorities, especially in the impoverished areas.

4.2 Significance Tests of Racial Disparities in Breast Cancer Mortality

4.2.1 Breast Cancer Mortality Rates

Non-zero death of breast cancer cases was observed in merely a smaller proportion of census tracts for minority groups as compared with non-Hispanic white whites in virtue of small population size for minorities within such a small geographical scale of census tract. Table 4.4 provides the information of the number of census tracts

with non-zero population and mortality data for the three racial groups. The ratio of row 1 by row 2 was referred to as percentage. Only 36% of census tracts with Hispanic residents reported at least one death from breast cancer during the eleven-year period of 1995-2005. A similar proportion (31%) was observed for African-Americans. In comparison, non-Hispanic white females had a much larger percentage (80.28%) at the census tract level. The much smaller proportion of units which had at least one death was observed for minority populations. In addition, as census tracts generally include similar population sizes, they are smaller within metropolitan urban areas such as Houston, Dallas, and the Austin-San Antonio area.

Table 4.4 Number of census tracts with non-zero population and non-zero mortality of breast cancer by race

		Non-Hispanic white	African- Americans	Hispanics
Census Tracts (n=4,388)	Population	4,382	4,322	4,375
	Mortality	3,518	1,341	1,592
	Percentage	80.28%	31.03%	36.39%

Figure 4.4 demonstrates the distribution of mortality rates and proportion by age and race. Women of each race before their mid twenties had close to zero mortality rates (Figure 4.4a). Afterwards, breast cancer mortality increased with age for African-American, Hispanic, and non-Hispanic women. The mortality curves for all three racial groups did not overlap and displayed similar ranking across age groups. African-Americans had the highest mortality while Hispanics had the lowest, with intermediate rates for non-Hispanic whites. There was a stronger slope to the curve for African-Americans around age 70-74, with a lag of about five years for non-Hispanic whites and

Hispanics. The proportion of mortality cases by age and race is shown in Figure 4.4b.

Younger African-American and Hispanic women had the higher proportion of mortality among the younger age group than non-Hispanic women, and vice versa among the older age group of 70-74. For African-American and Hispanic women, the mortality peaked five years earlier than non-Hispanic white females who approached their mortality crest at age 55-59.

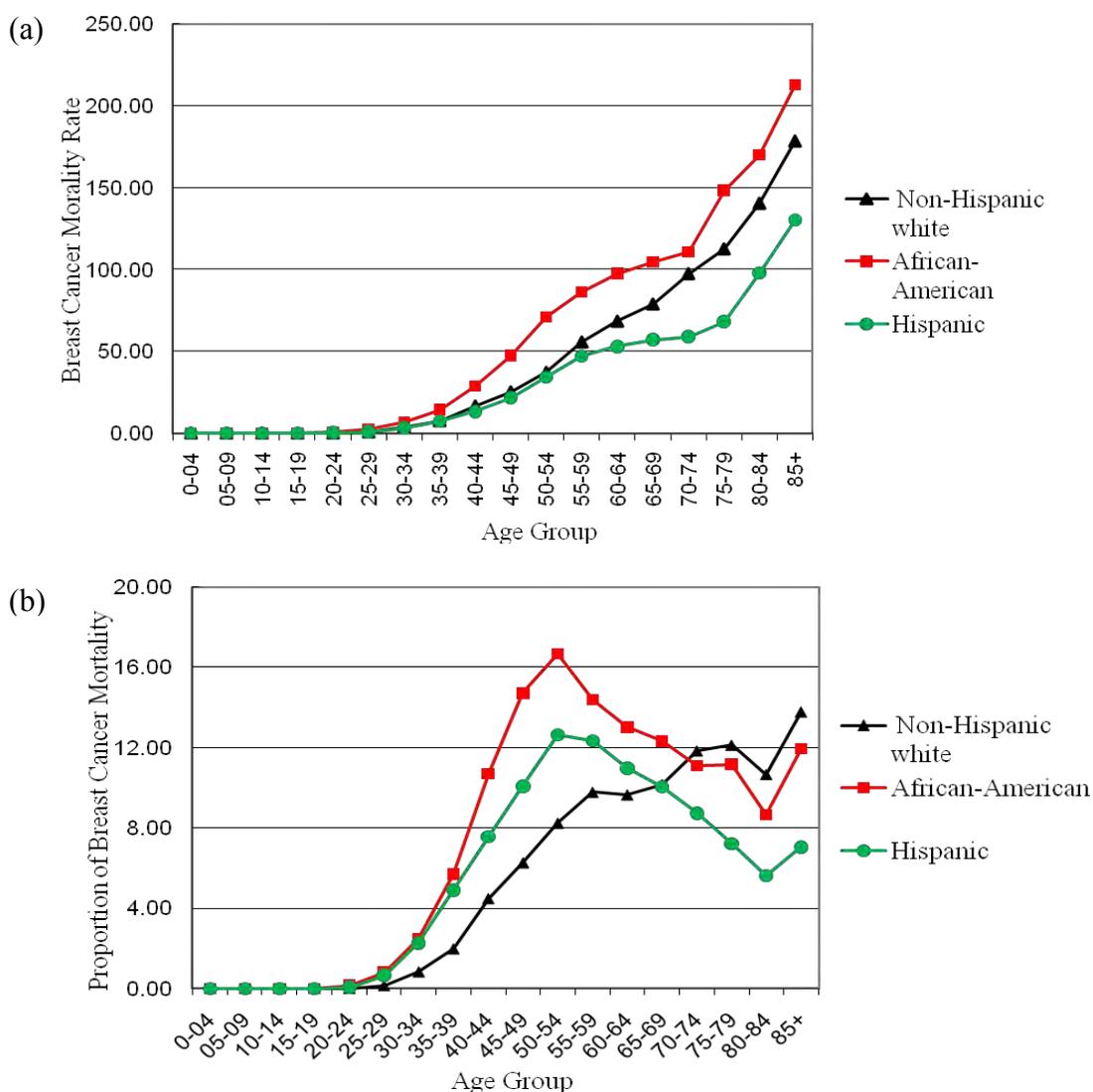


Figure 4.4 Age-specific mortality rates per 100,000 (a) and proportion of mortality cases (b) for breast cancer by race and age group.

4.2.2 Racial Disparities of Breast Cancer Mortality Rates among African-American Women

Table 4.5 shows all census tract units where absolute and relative racial disparities in breast cancer mortality tested significant. The rate difference (RD) measurement is an absolute measure of breast cancer mortality rates for minority women compared to the non-Hispanic white. The rate ratio (RR) measurement is a relative measure of breast cancer mortality rates for minority groups compared to non-Hispanic whites. A positive rate difference (RD) indicates higher breast cancer mortality rates for minority women. Similarly, if the rate ratio (RR) exceeds 1.0 the minority women experienced significantly higher mortality rates than their non-Hispanic white counterpart. The number of census tract units was classified further according to low, middle, and high socioeconomic status (SES) as measured by the percentage of the population living below the poverty threshold.

The statistic was not computed and labeled as “no cases” in the legend of Figure 4.5 whenever populations were zero for either reference or target groups using both measurements. In addition, census tracts were not considered as well if the number of cases was zero for one of the two ethnic groups using RR measurement or zero for the reference group (non-Hispanic white) using RD measurement. About 278 out of the total 4,388 census tracts (6.30%) displayed significantly higher mortality for African-American women compared to the non-Hispanic whites in terms of rate difference (RD) measurement (Table 4.5). Among these census tracts with significantly higher mortality rates for African-American women, a large proportion (87.77%) had a poverty rate greater than 10.00%. Most of the significant disparities for African-American women were found within the metropolitan areas of the Houston, Dallas, and Austin-San Antonio

(Figure 4.5a). On the other hand, two census tracts located within the Houston metropolitan areas displayed significantly higher mortality rates for non-Hispanic white women. The small number issue was responsible for the observation. A closer examination indicated that the two census tracts had either one or two deaths for a population of twelve non-Hispanic white females in the past eleven years, which contributed to the exceptionally high mortality rates. These two census tracts had a low SES level with more than 20% population living under the poverty line. No census tract tested for significant relative racial disparities (RR) (Figure 4.5b).

Table 4.5 Number of census tracts with significant racial disparities in breast cancer mortality for African-American women using RD and RR measurements

Poverty Level		High SES (0-9.99%)	Middle SES (10.00-19.99%)	Low SES (≥20.00%)	Total
	RD < 0	0	0	2	2
Census Tract (n=4,388)	RD > 0	34	54	190	278
	RR < 1	0	0	0	0
	RR > 1	0	0	0	0

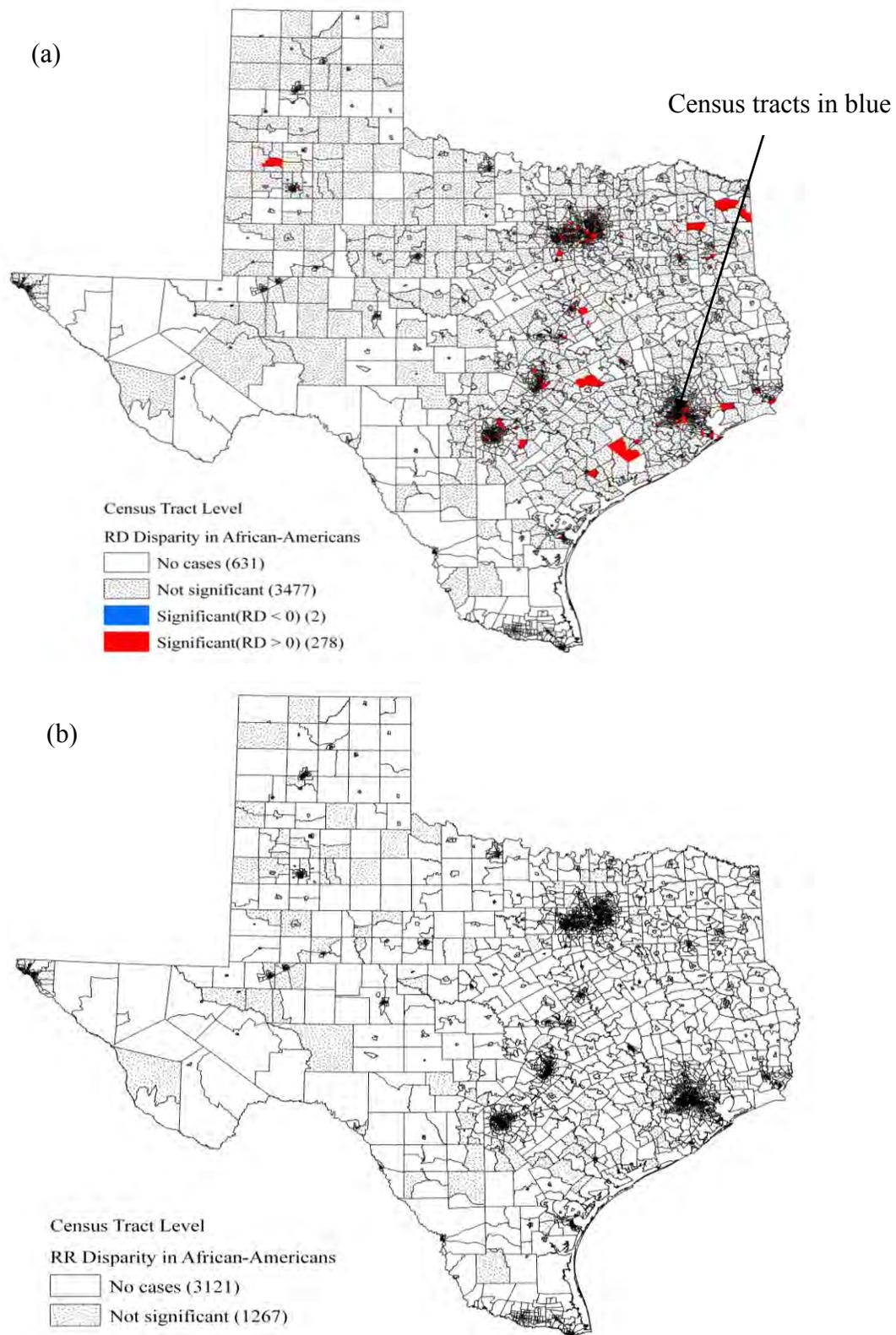


Figure 4.5 African-American breast cancer mortality: significant racial disparities according to the RD (a) and RR (b) statistics.

4.2.3 Racial Disparities of Breast Cancer Mortality Rates among Hispanic Women

Table 4.6 shows the number of census tracts significant in racial disparities of breast cancer mortality for Hispanic women at various SES levels. Low, middle, and high SES levels were referred as 0-9.99%, 10.00%-19.99%, and 20.00%+ population living the federal poverty line (Chu *et al.* 2007). Out of 4,388 census tracts in total, 328 had significantly higher mortality rates for Hispanic women in comparison with non-Hispanic white using RD statistics, while significantly higher mortality rate for non-Hispanic whites were detected for only two census tracts (Table 4.6). These 328 census tracts were primarily found within the three metropolitan areas and the Southwest border of Texas (Figure 4.6a). The majority (81.11%) of census tracts with significantly higher mortality rates of breast cancer among Hispanic women had low SES with more than 10.00% population living under the poverty line. However, two census tracts had significantly higher mortality rates for non-Hispanic white with a low SES level. No significant racial disparity was found when using the RR statistic (Figure 4.6b).

Table 4.6 Number of census tracts with significant racial disparities in breast cancer mortality for Hispanic women using RD and RR measurements

Poverty Level		High SES (0-9.99%)	Medium SES (10.00-19.99%)	Low SES (≥20.00%)	Total
Census Tract (n=4,388)	RD < 0	0	0	2	2
	RD > 0	20	42	266	328
	RR < 1	0	0	0	0
	RR > 1	0	0	0	0

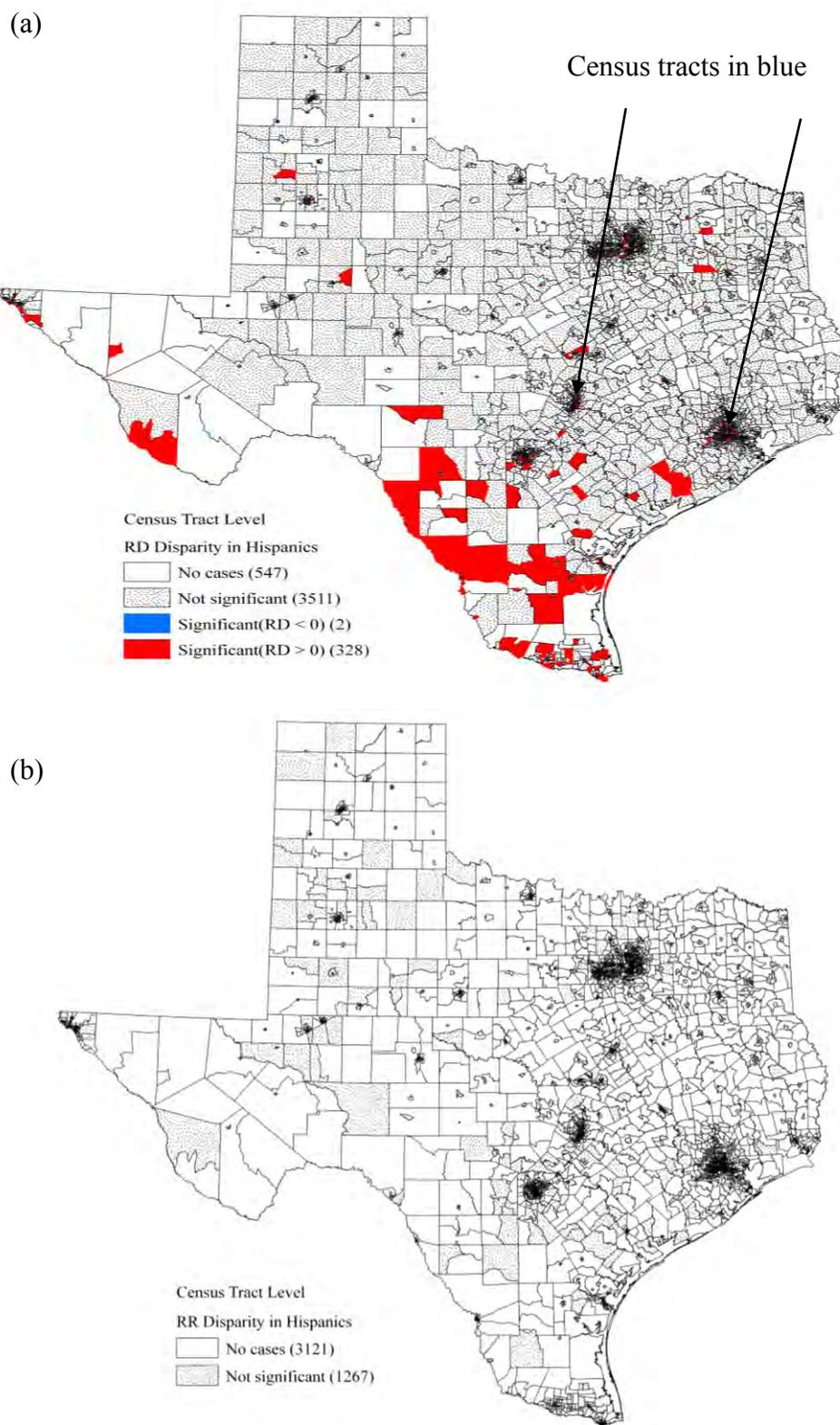


Figure 4.6 Hispanic breast cancer mortality: significant racial disparities according to the RD (a) and RR (b) statistics.

4.2.4 Discussion on Racial Disparities of Breast Cancer Mortality

This study found that the proportion and location of census tracts tested significant in racial disparities of breast cancer mortality changed depending on the type of statistic (absolute versus relative). The application of the RD statistic to census tract data resulted in the detection of a larger proportion of significant racial disparities. None of census tracts was identified significant through the RR statistic. Statistically, more differences are tested significant when measured in absolute terms (i.e. rate difference RD) than relative terms (i.e. rate ratio RR). Sometimes, absolute and relative measures can lead to opposite conclusions on health disparities (Harper and Lynch 2005). Thus, the Centers for Disease Control recommend utilizing both absolute and relative measurement in order to fully understand the magnitude and direction of health differences, especially across geographic areas and populations (Keppel *et al.* 2005). So both relative and absolute measurements should be employed when assessing racial disparities in order to clarify where significant racial disparities occur.

African-American and Hispanic women experienced significantly higher mortality rates than non-Hispanic whites in the Southeast metropolitan areas and the Southwest border of Texas. Non-Hispanic whites generally displayed the lower mortality rates of breast cancer. Although a few census tracts had significantly better mortality outcome for African-Americans and Hispanics, it is likely an artifact of the analysis caused by too small population sizes following the partitioning into such a small geographic unit (census tract) and different age groups. For example, the two census tracts with significantly larger mortality for non-Hispanic whites within the Houston Metropolitan area contained only twelve non-Hispanic white females and one or two

deaths in the study period of 1995-2005, which could cause the dramatically higher mortality for non-Hispanic whites.

More units were found significant for Hispanics than for African-Americans, which could be explained by the wider geographical distribution of Hispanics relative to African-Americans in Texas. Significantly higher mortality rates for African-American and Hispanic women occurred in more impoverished areas of the Southeast metropolitan areas and Southwest border with Mexico. Regions of lower socio-economic status were found to be associated with more substantial racial disparities for African-Americans and Hispanics. A plausible explanation might be that disadvantaged minorities living in underserved areas could not access health care resources as conveniently as other races in higher SES areas.

It should be noted that impoverished neighborhoods are typically characterized by lack of sufficient health care facilities, physicians, and even appropriate cancer treatments (Bradley *et al.* 2002; Farley and Flannery 1989; Heck *et al.* 1997). Although non-Hispanic whites live in the same disadvantaged regions, these females may be able to overcome transportation barriers, have better financial support, and access to more health care in affluent neighborhoods (Blackman and Masi 2006; Wang *et al.* 2008). Late-stage diagnosis may happen more frequently in minorities and ultimately widens racial disparities in mortality rates within the same geographic regions (Adams *et al.* 2006; Baquet *et al.* 2008; Merkin *et al.* 2002). On the other hand, socio-economic status is a composite statistic that mixes all races, which could be misleading in that a region dominated by minorities may have lower SES overall that does not reflect the status of non-Hispanic whites in that region (Goovaerts 2010).

4.3 Spatial Relationship of Racial Disparities in Late-stage Diagnosis and Breast Cancer Mortality

4.3.1 Visualization of Spatial Locations of Significant Racial Disparities

To examine the spatial association of the two racial disparities in both late-stage diagnosis and breast cancer mortality, Table 4.7 shows the number of census tracts that tested significant in the two types of racial disparities. These geographic units were further sorted based on the low, middle, and high SES measured in terms of percentage of population living below the poverty line. For both African-Americans and Hispanics, more census tracts were identified as experiencing significant racial disparities in breast cancer mortality than late-stage diagnosis. For the absolute statistic RD, it was observed that about 188 census tracts had African-American women with significantly higher late-stage diagnosis and 278 census tracts with significantly higher breast cancer mortality relative to their non-Hispanic white counterparts. A similar pattern was observed for Hispanics who experienced significantly higher rates of late-stage diagnosis and breast cancer mortality in 266 and 328 census tracts, respectively.

The significance test of the two racial disparities was not consistent using RD and RR measurements (Table 4.7). When measured using RD, none of the census tracts reported significantly lower late-stage diagnosis for the two minority groups (i.e. the RD statistic is never significantly less than 0). However, a couple of census tracts had significantly lower mortality rates of breast cancer for minority groups. This could be explained by the small number issue. For the relative statistic RR, neither African-Americans nor Hispanics had significantly lower rates of either late-stage breast cancer diagnosis or breast cancer mortality because RR was never observed significant less than

1.0. A few census tracts were tested to have significantly higher rates of late-stage diagnosis for both African-Americans and Hispanics. But, no significant racial disparities in mortality were detected using RR measurement.

Table 4.7 Number of census tracts with significant racial disparities in late-stage diagnosis and breast cancer mortality using RD and RR measurements

			HighSES (0-9.99%)	Middle SES (10.00-19.99%)	Low SES (≥20.00%)	Total
African- Americans	Mortality	RD < 0	0	0	2	2
		RD > 0	34	54	190	278
	Late-State Diagnosis	RD > 0	13	26	149	188
		RR > 1	6	2	1	9
Hispanics	Mortality	RD < 0	0	0	2	2
		RD > 0	20	42	266	328
	Late-State Diagnosis	RD > 0	8	23	235	266
		RR > 1	0	1	1	2

Census tracts that reported significantly higher rates of late-stage diagnosis and breast cancer mortality often had the low SES level. A logistic regression analysis was performed with the significance test of racial disparities as the dependent variable and the level of SES as the independent variable. The analysis was conducted only for the RD statistic because of the greater number of census tracts reported as significant. Census tracts with RD less than 0 are coded as missing values/no data. The census tracts were dichotomized as either 1' (significant) or 0' (not significant) based on the significance test of racial disparities. SES was re-coded as numeric dummy variables of 3' (low), 2' (middle), and 1' (high). The relationships between racial disparities and SES levels (dependent variable) were assessed for both late-stage diagnosis and breast cancer

mortality (independent variable). Table 4.8 summarizes the results of the logistic regression with odds ratio of the parameters and significance of the model (p). The relationship between the two racial disparities and SES levels were tested significant with the p values less than 0.05. For African-American women, census tracts classified as middle and low SES were 2.27 and 18.35 times more likely to report significant racial disparities in late-stage diagnosis, while the ratio was 1.86 and 10.19 for breast cancer mortality. Within census tracts of middle and low SES level, Hispanics were found to be 3.24 and 46.89 times more likely to be associated with significantly higher rates of late-stage diagnosis than non-Hispanic white females. These ratios were 2.51 and 24.29 for breast cancer mortality. Furthermore, the impacts of SES level on racial disparities were greater for late-stage diagnosis than for mortality as well as for Hispanic women than for African-Americans.

Table 4.8 Results of the logistic regression: odds ratios (ORs) and p values with the significance of racial disparities in late-stage diagnosis/breast cancer mortality (dependent variable) and SES (independent variable) using RD measurement

	SES	African-Americans OR (95% CI)	p	Hispanics OR (95% CI)	p
Late-stage Diagnosis	High	1 (reference)	<0.001	1 (reference)	<0.001
	Middle	2.27 (1.16 - 4.44)	0.016	3.24 (1.45 - 7.27)	0.004
	Low	18.35 (10.35 - 32.52)	<0.001	46.89 (23.07 - 95.30)	<0.001
Mortality	High	1 (reference)	<0.001	1 (reference)	<0.001
	Middle	1.86 (1.20 - 2.89)	0.005	2.51 (1.46 - 4.29)	0.001
	Low	10.19 (7.00 - 14.82)	<0.001	24.29 (15.30 - 38.57)	<0.001

However, the census tracts that tested significant for racial disparities both in late-stage diagnosis and breast cancer mortality did not fully overlap across space (Figure 4.7). The number and geographical locations of census tracts that had significant absolute

disparities were shown for both African-Americans (Figure 4.7a) and Hispanics (Figure 4.7b). The results of significance tests were not shown in figures for RR statistics due to the much smaller number of census tracts significant in the two racial disparities. Among African-Americans, 109 census tracts tested significant for both late-stage diagnosis (188) and breast cancer mortality (278). These significant census tracts were primarily found within the three metropolitan areas. This result can be explained by the larger population of African-American women resided in urban areas. Among Hispanics, the 130 census tracts were tested significant in both late-stage diagnosis (266) and breast cancer mortality (328) (Figure 4.7b). These tracts were located in the metropolitan areas and along the Southwest border of Texas where the Hispanic population was larger. The downtown regions of these areas displayed significant higher late-stage diagnosis and breast cancer mortality rates for Hispanic women.

In summary, the dichotomous results from the significance test indicate that about 40% (109/278 or 130/328) of the census tracts with significant racial disparities in breast cancer mortality also displayed significant disparities in late-stage diagnosis for both African-Americans and Hispanics. On the other hand, 58% (49%) of the census tracts with significant racial disparities for late-stage diagnosis tested significant for African-Americans (Hispanics) for breast cancer mortality. More census tracts were identified as having significant racial disparities in breast cancer mortality than late-stage diagnosis, which might indicate that some other confounders impact mortality. There were more census tracts with significant results for Hispanics than for African-Americans due to the much higher Hispanic population resided in Texas (Guzman 2001).

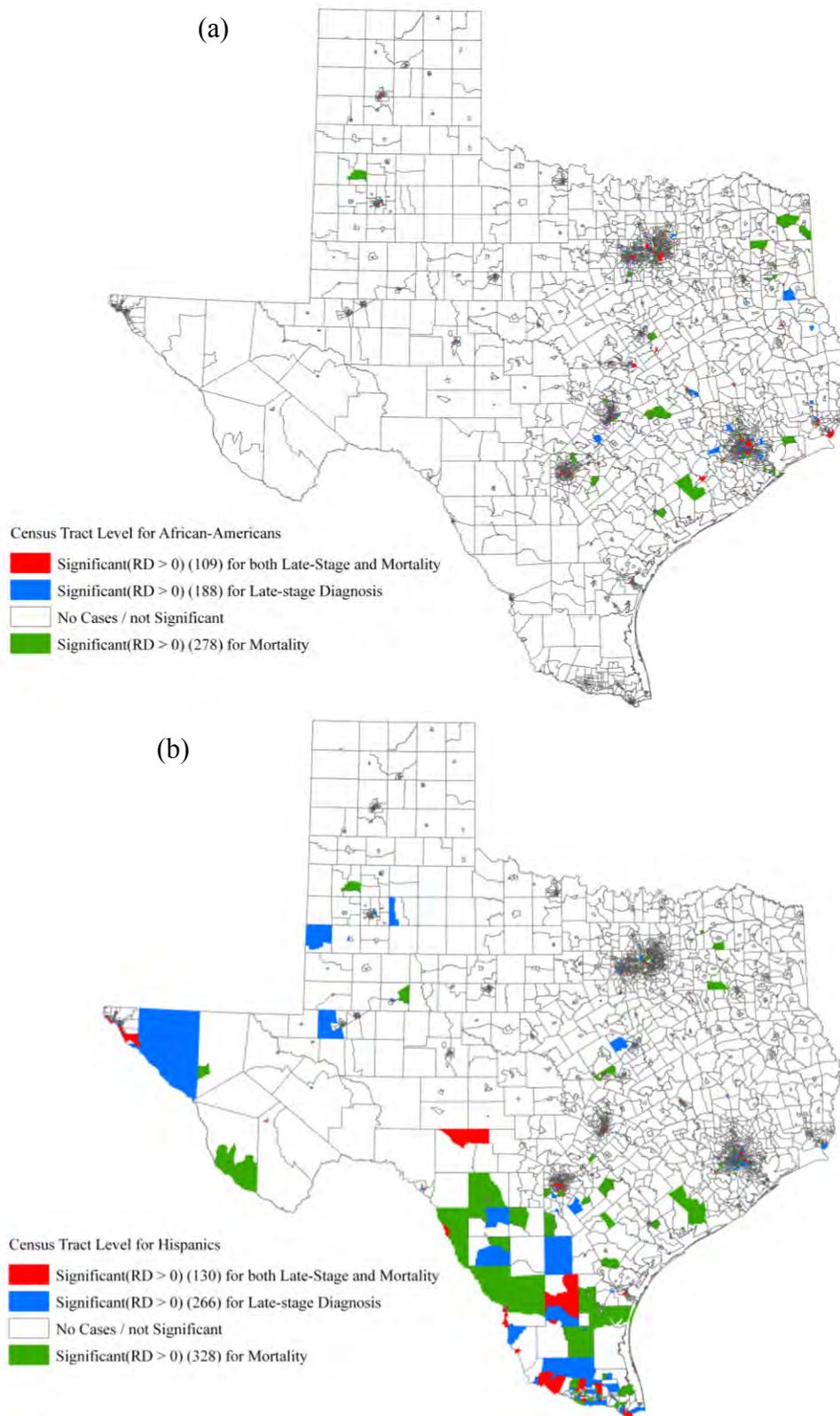


Figure 4.7 Geographic distributions of census tracts with significant racial disparities of late-stage diagnosis and breast cancer mortality (RD) for both African-American (a) and Hispanic women (b).

4.3.2 Descriptive Statistics of Racial Disparities in Late-stage Diagnosis and Breast Cancer Mortality

Spatial relationships of racial disparities between late-stage diagnosis and breast cancer mortality need to be assessed in a regression analysis. The regression analysis can be very sensitive to the presence of extreme values for the dependent and independent variables. Therefore, examining the statistical distribution of RD and RR statistics becomes a critical step before constructing any regression models. A .dbf file with all the RD and RR statistics for both African-American and Hispanic women was exported from STIS, which prepared for the data-process in the software package SPSS.

Table 4.9 summarizes the descriptive information of RD and RR statistics, which includes minimum, maximum, mean, standard deviation, skewness, and kurtosis for racial disparities of late-stage diagnosis and breast cancer mortality for both African-American and Hispanic women. The disparity statistics were not computed for census tracts with either zero population or zero cases for minority groups and the reference group (i.e. white). N in Table 4.9 represents the number of census tracts with either RD or RR statistics. Min, max, mean, and standard deviation provided quantitative information about the data distribution of RD and RR measurements in both late-stage diagnosis and breast cancer mortality. Skewness measures the lack of symmetry of a distribution, while kurtosis informs on the degree of peak or flat relative to a normal distribution. Positive skewness implies the distribution has a long upper tail while positive kurtosis indicates the peak of distribution is sharper than the normal distribution. A normal distribution has skewness and kurtosis ranging from -3 to 3 (Rogerson 2010).

Table 4.9 Descriptive statistics of RD and RR racial disparities in late-stage diagnosis and breast cancer mortality for both African-American and Hispanic women

			N	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
African-Americans	RD	Mortality	3757	-10.24	20.50	0.00	0.65	13.88	404.30
		Late-stage Diagnosis	4053	-2.70	4.75	-.00	0.47	2.91	19.90
	RR	Mortality	1076	-3.57	5.58	0.26	0.52	1.70	18.67
		Late-stage Diagnosis	1640	-1.92	2.68	0.19	0.49	0.88	3.64
Hispanics	RD	Mortality	3841	-3.07	3.23	-0.16	0.39	1.59	14.72
		Late-stage Diagnosis	4182	-2.61	5.31	-0.14	0.42	1.64	16.44
	RR	Mortality	1267	-1.43	1.99	0.03	0.43	0.89	2.51
		Late-stage Diagnosis	2182	-1.51	3.07	-0.0	0.41	0.93	3.99

A histogram is an effective visual approach to examine the skewness and kurtosis of datasets. The shape of these histograms and the statistics, shown in Figures 4.8-4.9 and Table 4.9 convey the information on skewness and kurtosis in another mean. For African-American women, the RD statistics showed the leftmost concentrated peak with quite a few outliers in both late-stage diagnosis (Figure 4.8a) and breast cancer mortality (Figure 4.8b). The histograms of the RR statistics illustrated the data distribution slightly away from the normal distribution for both late-stage diagnosis (Figure 4.8c) and breast cancer mortality (Figure 4.8d). For Hispanic women, both RD and RR statistics illustrated positive skewness and kurtosis for late-stage diagnosis and breast cancer mortality (Figure 4.9). Moreover, racial disparity statistics for Hispanics were more likely to conform the normality than for African-American women. In addition, the disparity statistics in late-stage diagnosis (Figures 4.7 a, 4.7c, 4.8a, and 4.8c) exhibited a better

normal distribution shape than in breast cancer mortality (Figures 4.7b, 4.7d, 4.8b, and 4.8d)

A box plot is a convenient approach to depict numerical data with easy identification of outliers and extreme values. Provided with their histograms, all the box plots have circle indicating the outliers and asterisk illustrating the extreme values. Figures 4.7 and 4.8 indicate the RD and RR statistics does not follow the normal distribution because of a number of the outliers and extreme values within datasets. Regression analysis requires avoiding extreme and outlier values so that it is necessary to transform the datasets into normal distributions. Score transformation aims to give meaning to raw score and allow a direct comparison of two scores (Bartlett 1947). The normal score transformation was achieved by ranking the dataset from the minimum to maximum values and matching these ranks to their corresponding ranks in a normal distribution. The procedure of normal-score transformation was performed in STIS.

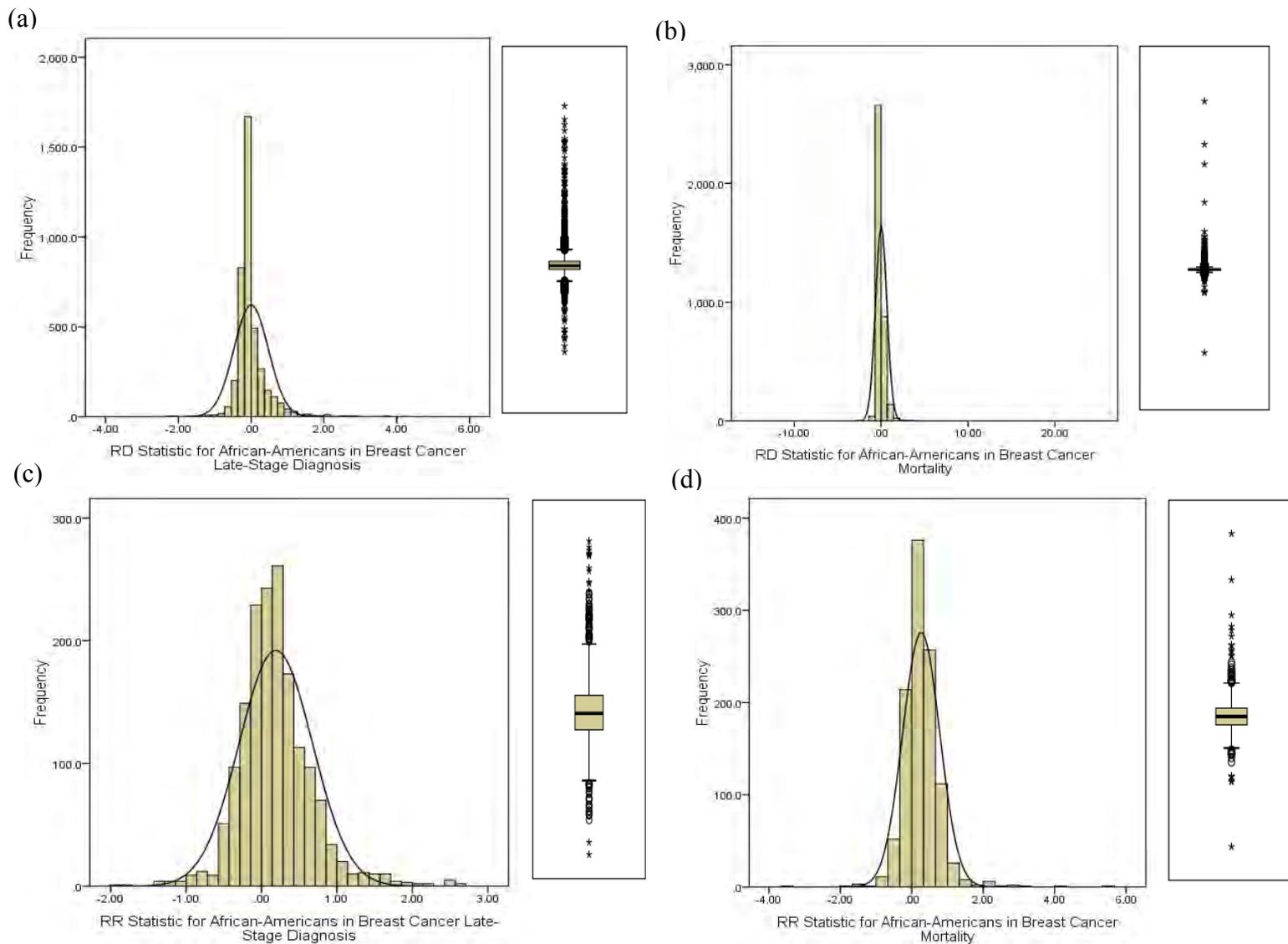


Figure 4.8 African-American women: histograms and box plots of RD (a and b) and RR (c and d) in late-stage diagnosis and breast cancer mortality.

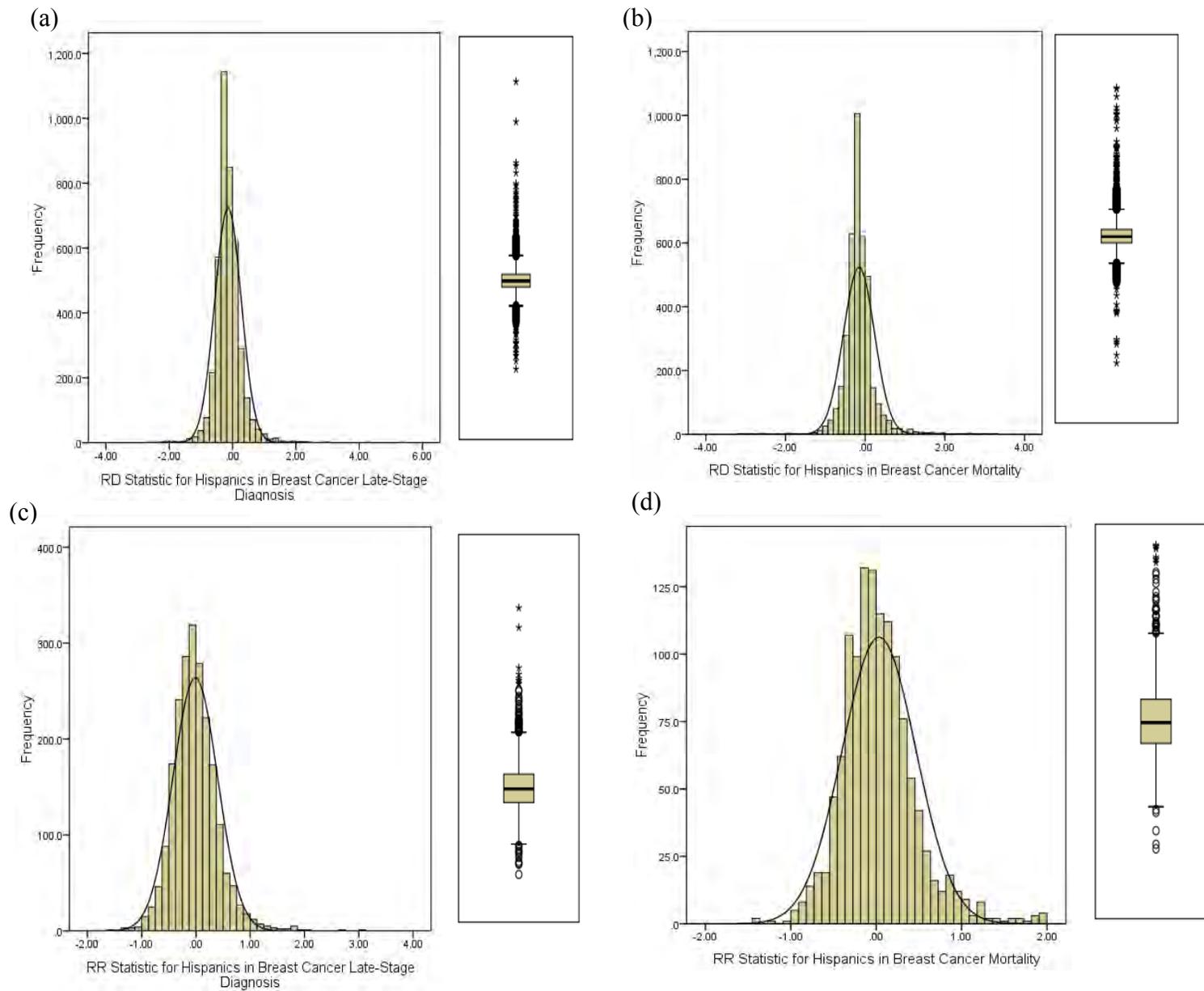


Figure 4.9 Hispanic women: histograms and box plots of RD (a and b) and RR (c and d) in late-stage diagnosis and breast cancer mortality.

4.3.3 Linear Regressions, Spatial Autocorrelation Diagnosis, and Geographically Weighted Regressions

The spatial relationships of the two racial disparities in late-stage diagnosis and breast cancer mortality can be assessed using confirmatory data analysis such as linear regression. In the linear regression models, normalized RD or RR statistics in breast cancer mortality was utilized as dependent variables and normalized RD/RR in late-stage diagnosis as independent variables. The linear regression analysis was implemented in STIS.

The basic assumption for the conventional linear regression is the absence of spatial autocorrelation among observations. However, this is often not the case when dealing with spatial datasets. Local indicators of spatial autocorrelation (LISA) were employed to investigate the spatial correlation of residuals from linear regressions. The method of LISA statistics were discussed in details in Chapter III. If spatial dependency is significant within the residuals, it indicates that the spatial factor is not considered in regression analysis and geographically weighted regression is necessary to fully explain the relationships of the two racial disparities between late-stage diagnosis and breast cancer mortality. The LISA statistic, often referred to as the local Moran's I , have the capability to detect clusters or outliers. A cluster is defined as a subset of data which shares a similar values in magnitude, while an outlier refers to an observation that distinctly different from the surrounding areas around focal points.

Table 4.10 shows the results of linear regression using the racial disparity statistics of breast cancer mortality as dependent variable and late-stage diagnosis as independent variable. First, all the regression models were significant with the p values less than 0.001, which implied that the independent variable of racial disparities in late-

stage diagnosis was a significant predictor. R square in a regression model refers to the percentage of variance in the dependent variable that can be explained by the independent variables. The R square was as low as 0.053 using the RD measurement and the regression with the RR measurement had much higher R square of 0.235 for African-American women and 0.197 for Hispanic women. RR statistics lead to the higher model fit than RD statistics.

Since R square is the square of the correlation coefficient, similar observations were found for the correlation coefficient, which measured the strength of the association between dependent variable and explanatory predictors. For both African-American and Hispanic women, the coefficients of correlation were 0.23 for normalized RD statistics, while they were almost two times higher for RR statistic. More missing values were reported for the RR statistic due to its measurement nature, which could increase the correlation coefficient in a regression model. Based on the R square and correlate coefficient, the relationships of racial disparities in both late-stage diagnosis and breast cancer mortality were tested significant.

Table 4.10 Linear regression results on normalized RD/RR statistics with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis for both African-American and Hispanic women

	Racial Disparities	Dependent variable	Independent variable	R Square	Coefficient	<i>p</i>
African-Americans	RD	Breast cancer mortality	Late-stage diagnosis	0.053	0.230	<0.001
	RR			0.235	0.486	<0.001
Hispanics	RD	Breast cancer mortality	Late-stage diagnosis	0.053	0.230	<0.001
	RR			0.197	0.445	<0.001

Spatial autocorrelation tested significant within the residuals of the linear regression based on the Moran's I . The minimum p value of 0.037 indicated that the spatial autocorrelation was significant for the residual of linear regression, as shown in Table 4.11. Moran's I is a mean to detect the outlier and cluster through reflecting the spatial dependency of neighborhoods. The low values signified that the spatial dependency was weak, but significant. The positive values indicated the existence of spatial clustering characteristics that nearby things were more similar with each other.

Table 4.11 Moran's I results on residuals of the linear regression analysis with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis for both African-American and Hispanic women

	Racial Disparities	Moran's I	p
African-Americans	RD	0.040	0.001
	RR	0.055	0.037
Hispanics	RD	0.040	0.001
	RR	0.100	0.001

Four types of significant spatial autocorrelation were provided in Figures 4.9 including high-high (hot cluster: high and neighbors are also high), low-low (cold cluster: low and neighbors are also low), high-low (outliers among low neighbors), and low-high (low outliers among high neighbors). For African-American women, 77 census tracts were identified to have positive autocorrelation with concentrated high values and 83 census tracts with low values together (Figure 4.10a). Fewer census tracts were found as clusters for the residuals of linear regression in terms of RR measurements (Figure 4.10b). This might result from a large number of census tracts having missing values for the RR disparity statistic. Hispanic women had 69 census tracts with high values surrounding each other and 94 census tracts with low values together using the RD measurement in residuals of linear regressions (Figure 4.11a). Much smaller portion of census tracts was tested as cluster using the RR measurement (4.11b). More census tracts were identified as clusters for Hispanic women than for African-Americans.

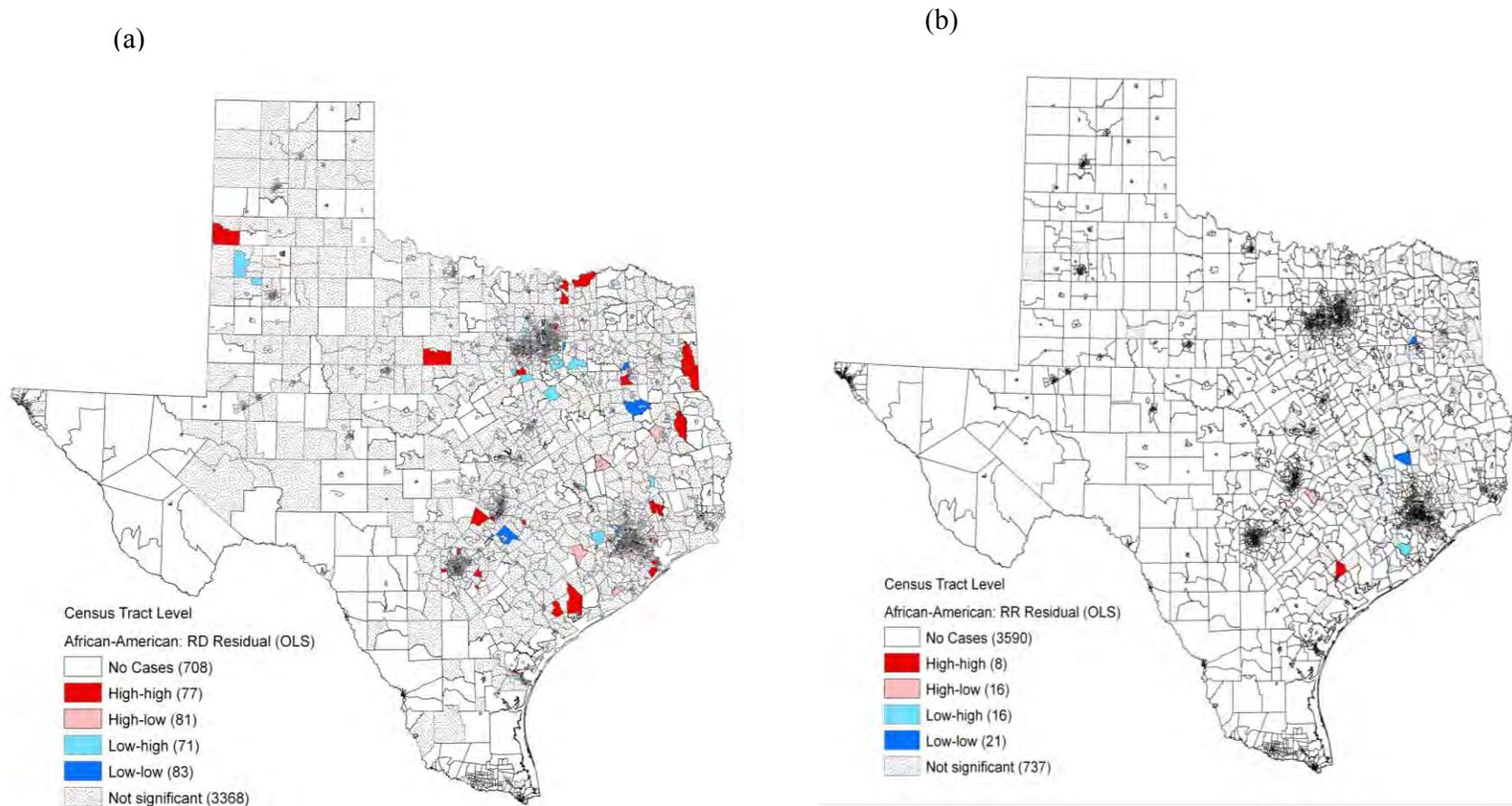


Figure 4.10 LISA statistical results on the residuals of the linear regression analysis with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis: RD (a) and RR (b) for African-American women.

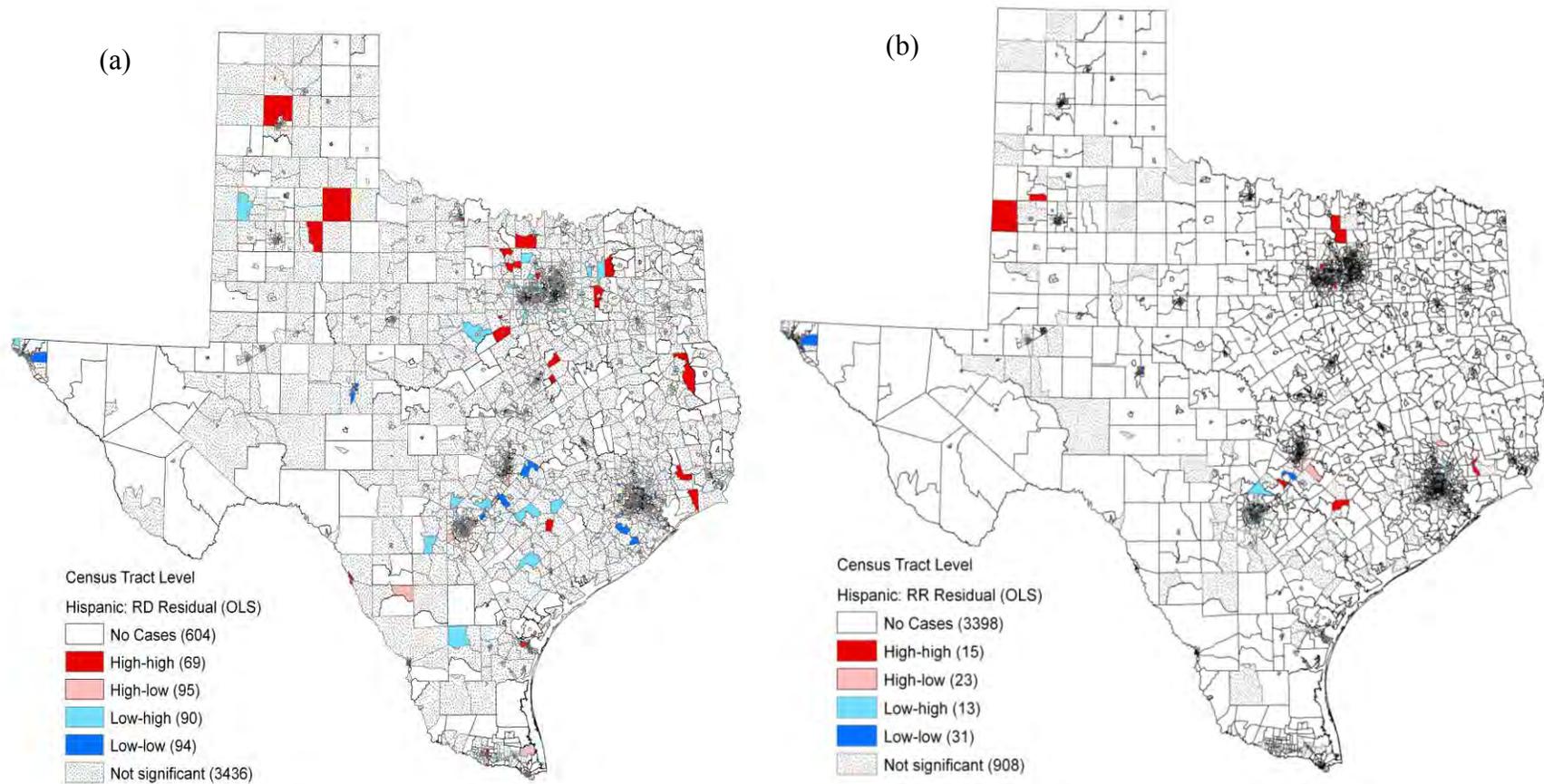


Figure 4.11 LISA statistical results on the residuals of the linear regression analysis with the dependent variable of racial disparities in breast cancer mortality and independent variable of racial disparities in late-stage diagnosis: RD (a) and RR (b) for Hispanic women.

The presence of spatial autocorrelation within the residuals of the linear regression suggests that the spatial relationships of the two racial disparities should be examined further by taking into consideration local neighbors effects on the focal census tracts. Geographically weighted regression (GWR) provides R squares and correlation coefficients for each census tract at local scale instead of a global assessment. A local linear regression was implemented to a subset of spatial data with weight proportional to proximity to the center of the focal point for each census tract. The GWR has the same dependent variable of racial disparities of breast cancer mortality and the independent variable of racial disparities of late-stage diagnosis.

The Gaussian kernel was applied to compute the weight of the neighborhoods in the GWR analysis, which was performed in ArcInfo 9.3. The bandwidth of the kernel was determined using Akaike Information Criterion (AIC). Kernel type was set up as adaptive due to the missing values of census tracts across space. A count of 30 neighbors was defined as the size of the neighborhood window for the analysis. The detailed discussion pertaining to the geographically weighted regression can be found in Chapter III. The geographically weighted linear regression provides not only a global adjusted R squares, but also the local R squares. The adjusted global R squares were 0.072 and 0.250 using RD and RR measurement respectively for African-American women and 0.084 and 0.269 for Hispanic women. The model fit of the linear regression was slightly improved by accounting for the local neighborhoods.

The correlation coefficient varied across space from the linear geographically-weighted regression (GWR) by taking into account the surrounding census tracts (Figures 12 and 14). The positive relationship was primarily observed between racial disparities of

late-stage diagnosis and breast cancer mortality. The correlate coefficient values fluctuated from -0.04 to 0.46 and 0.33-0.58 in terms of RD and RR measurements respectively for African-Americans (Figure 12a and Figure 12b), while they ranged from -0.01 to 0.51 and -0.01 to 0.73 in RD and RR respectively for Hispanic women (Figure 13a and Figure 13b). Further, the stronger correlate coefficients were found in the Southeast for both African-American and Hispanic women.

Figures 4.13 and 4.15 display the distribution of local R square with the RD and RR measurements for both African-American and Hispanic women. The spatial relationships of racial disparities of late-stage diagnosis and breast cancer mortality vary across space. Local R squares were observed higher using RR measurement than RD measurement. Among African-American women, the maximum local R square was 0.17 in terms of RD statistic (Figure 4.13a) and 0.35 in terms of RR statistic (Figure 4.13b). Hispanic women were found to have the maximum local R square of 0.24 using RD measurement (Figure 4.15a) and R square of 0.53 when using RR measurement (Figure 4.15b). Hispanics had higher spatial variation of R square than African-Americans. Higher R square values for African-Americans in absolute disparity were found within the metropolitan areas of Dallas and Austin-San Antonio. The R squares were higher for Hispanic women in the Southeast region.

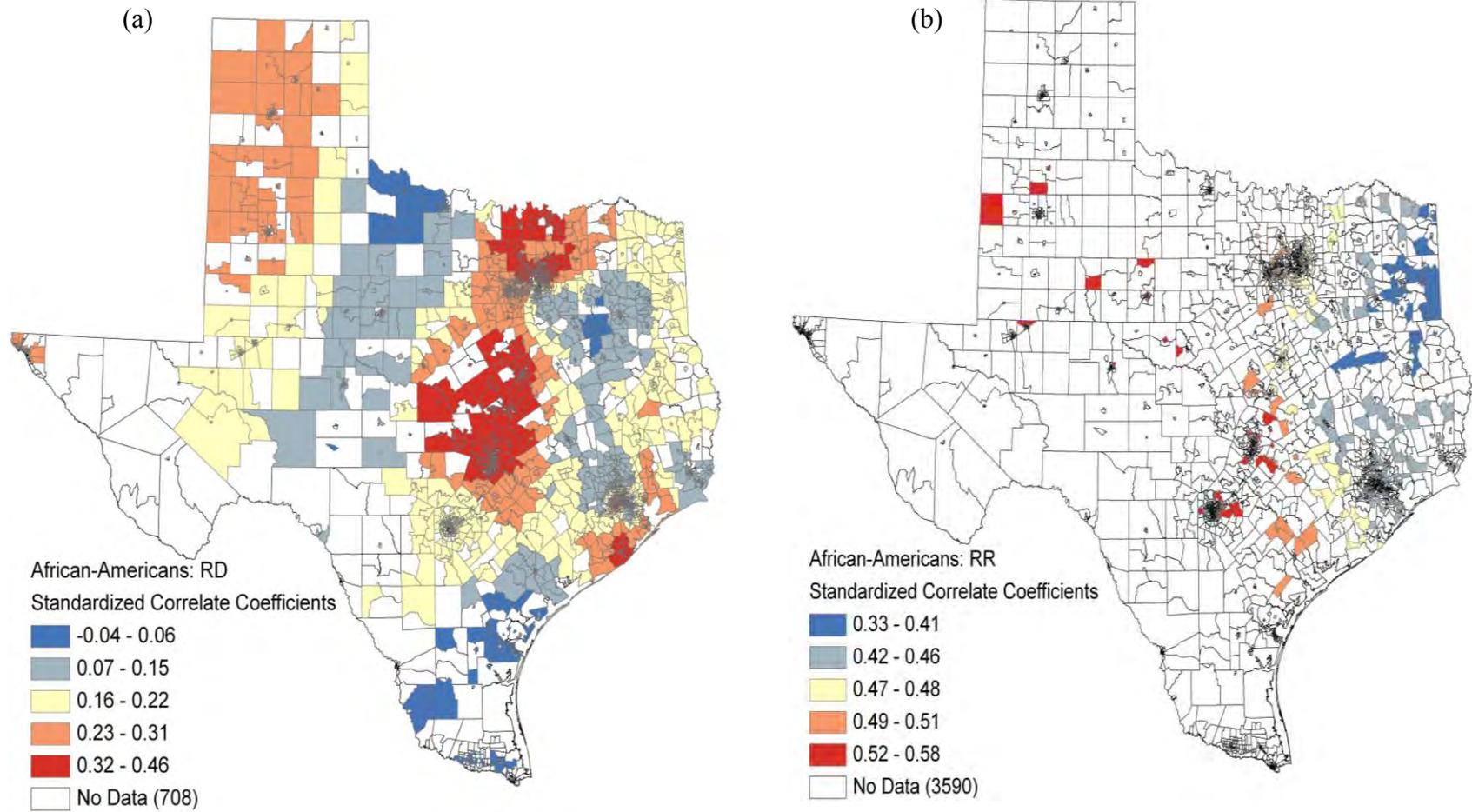


Figure 4.12 Standardized correlate coefficients results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) statistics for African-American women.

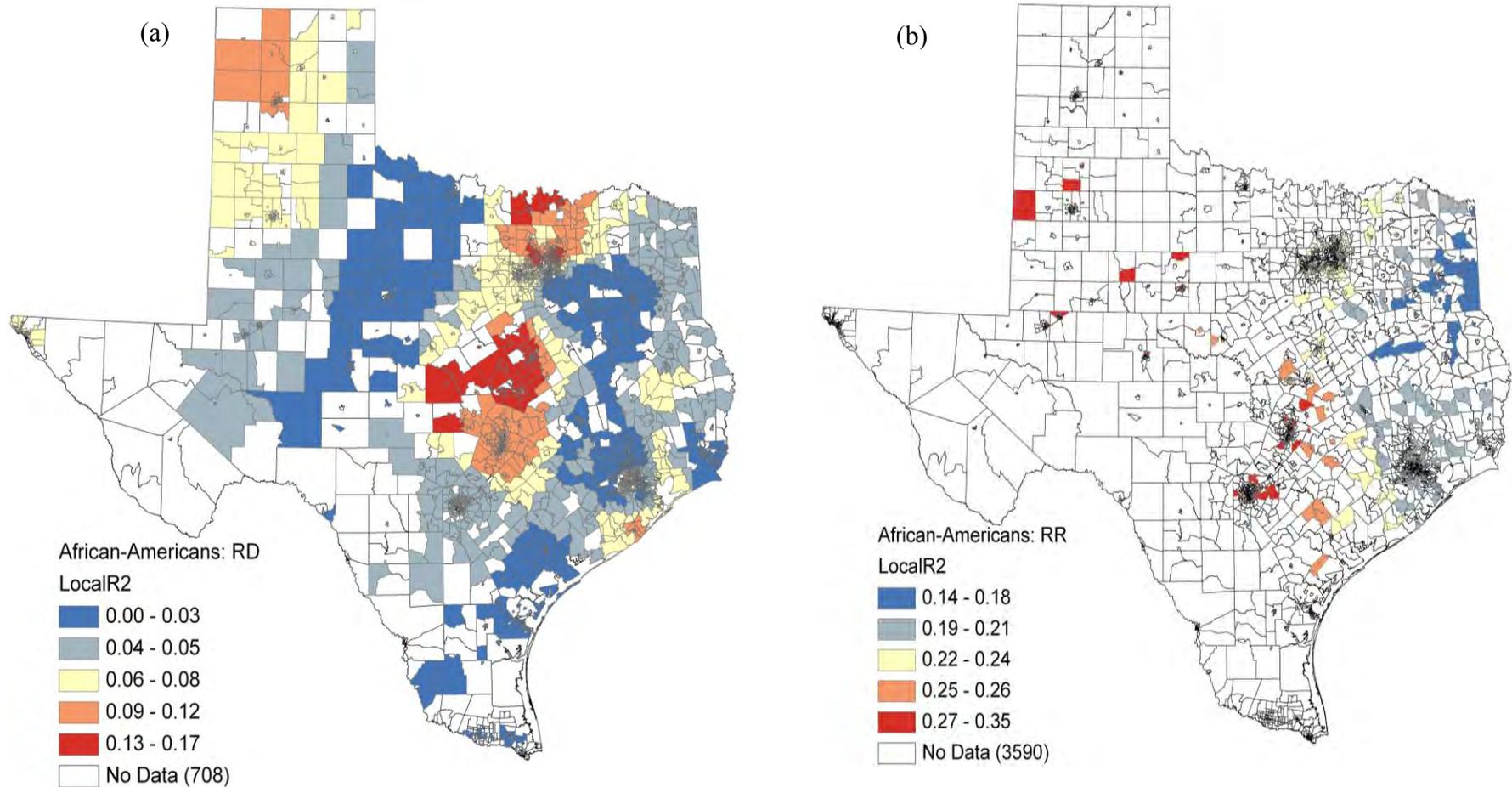


Figure 4.13 Local R square results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) statistics for African-American women.

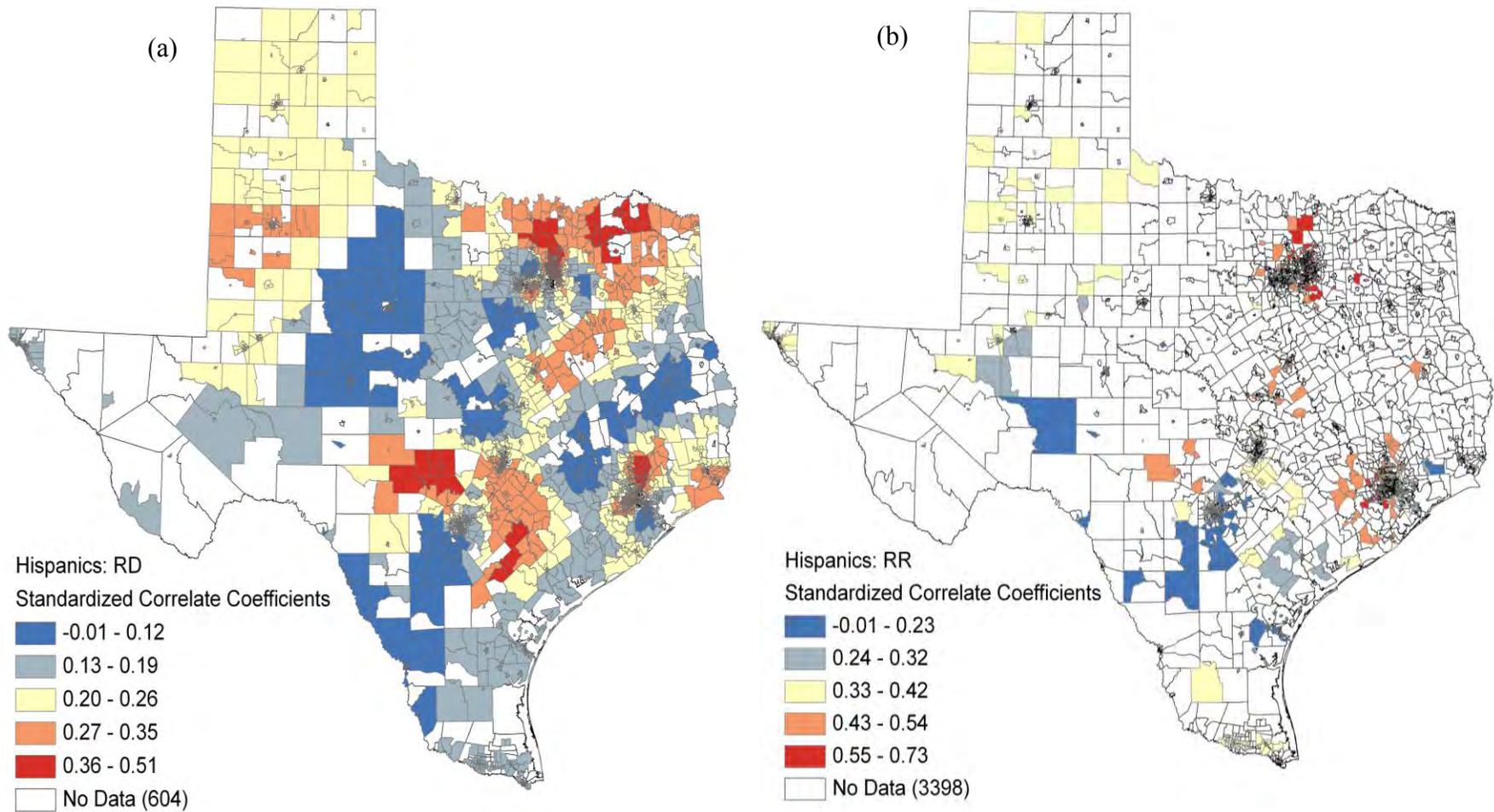


Figure 4.14 Standardized correlate coefficients results on the GWR analysis in racial disparities of breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) statistics for Hispanic women.

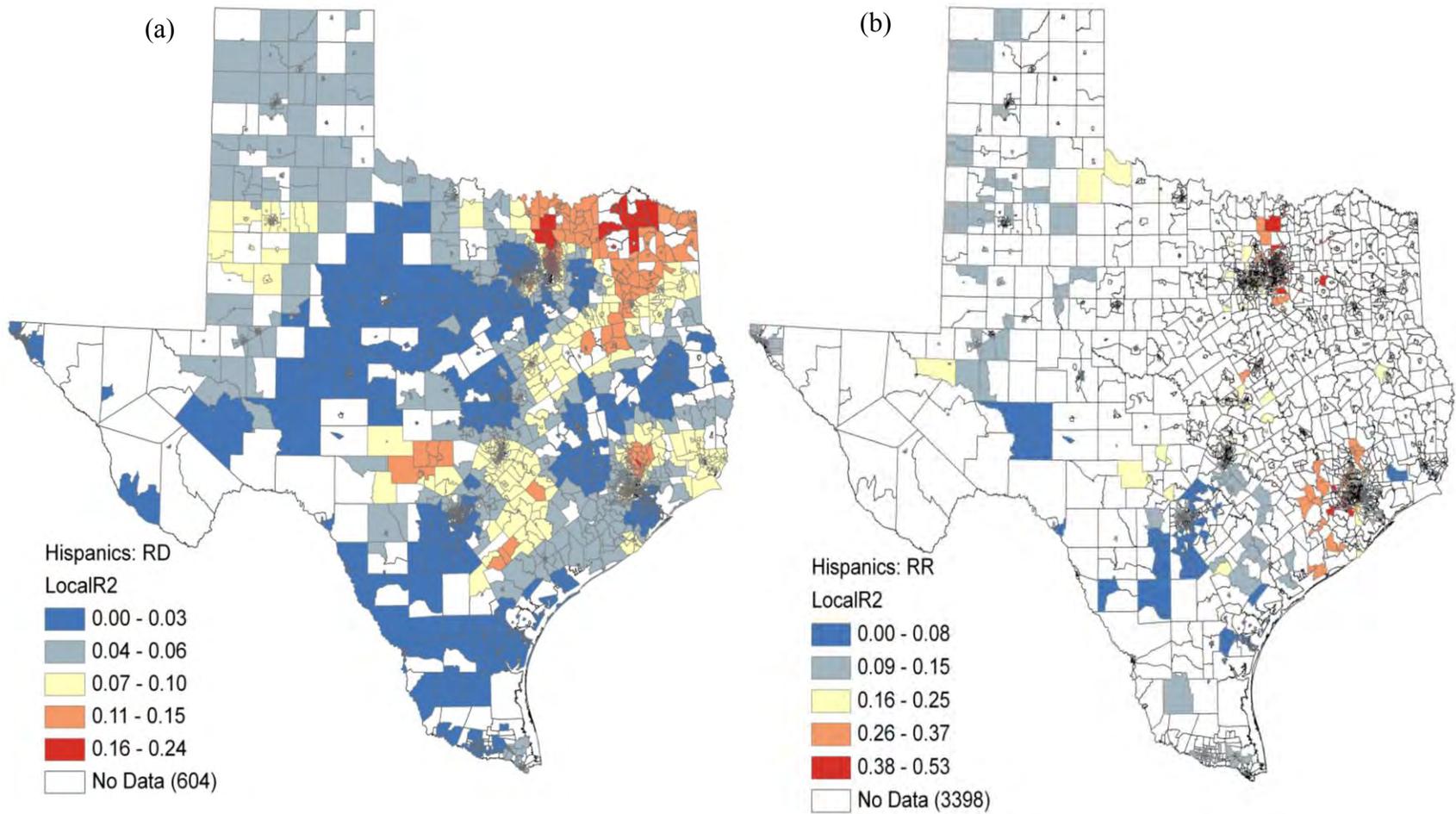


Figure 4.15 Local R square results on the GWR analysis with racial disparities in breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable): RD (a) and RR (b) statistics for Hispanic women.

4.3.4 Logistic Regression, Spatial Autocorrelation Diagnosis, and Geographically Weighted Regression

The linear regression model reported the low R squares which were utilized to evaluate the spatial relationship of racial disparities between late-stage diagnosis and breast cancer mortality. A few census tracts were identified significant in the two racial disparities when using RR statistic. Thus, the logistic regression is not suitable to assess the spatial relationship of the two racial disparities for the RR measurement. In this study, the logistic regression was constructed for RD measurement with the significance of racial disparities in mortality as independent variable (binary: significant = 1/non significant = 0) and significance of racial disparities in late-stage diagnosis as dependent variable (binary: significant = 1/non significant = 0). The logistic regression was discussed in detail in Chapter III.

Table 4.12 shows the results of logistic regression with respect to the spatial relationships of the two racial disparities using RD measurement. The full models and individual model parameters were tested significant with the p value less than 0.000001. Cox & Snell and Nagelkerke R squares were 0.10 and 0.25 respectively for both African-American and Hispanic women. Higher Nagelkerke R square indicated the better model fit. The odds ratio refers to the significance probability of racial disparities in breast cancer mortality. Significantly higher odds ratio reflected the strong relationship of the two racial disparities. For example, if a census tract was identified to have significant racial disparities of late-stage diagnosis, the census tracts were 33.76 times more likely to be significant in racial disparities of breast cancer mortality for African-Americans and 30.39 times for Hispanics.

Table 4.12 Results of logistic regression for racial disparities of breast cancer mortality (dependent variable) and late-stage diagnosis (independent variable) using RD measurement for African-American and Hispanic women

	R square (Cox & Snell)	R square (Nagelkerke)	Odds ratio	Odds ratio C.I. (95%)	<i>p</i> value
African- Americans	0.10	0.25	33.76	23.96-47.57	<0.000001
Hispanics	0.11	0.25	30.39	22.09-41.82	<0.000001

Moran's *I* was applied to the residuals of logistic regression to detect any spatial autocorrelation within the dataset, which intended to verify if the spatial dependency should be concerned in the regression model. For African-American women, Moran's *I* index was reported as 0.08 with *p* value of 0.001, while for Hispanic women, Moran's *I* had index value of 0.139 with the significance level equal to 0.001. Moran's *I* suggested the existence of spatial dependency within the residual of logistic regression. However, the geographically weighted logistic regression failed to converge for both African-American and Hispanic women.

4.3.5 Discussion of Spatial Relationships of Racial Disparities of Late-stage Diagnosis and Breast Cancer Mortality

The spatial relationships of racial disparities between late-stage diagnosis and breast cancer mortality were complex and non-stationary across Texas from 1995-2005. The results of the study suggested that 40% of the census tracts with significant absolute disparity for breast cancer mortality were found to be significant as well in late-stage diagnosis for both African-American and Hispanic women. This study confirmed the strong and positive spatial relationships between late-stage diagnosis and breast cancer mortality established in the literature review. For instance, using SEER data, Li and colleagues (2003) found that African-Americans were 2.3 and 2.5 times more likely to

have the elevated risk of Stage III and Stage IV breast cancer compared with non-Hispanic whites. On the other hand, Hispanic women were tested 1.8-fold in diagnoses at the late stage of breast cancer. Furthermore, African-Americans and Hispanics were found to have 1.5 and 1.1 greater in mortality risk, even after adjusting age, SEER registry, stage, ER and PR status, surgical treatment, and radiation therapy than whites (Li *et al.* 2003).

Lower socioeconomic status (SES) women have less likelihood to be diagnosed in early-stage breast cancer and consequently have higher mortality rates (Farley and Flannery 1998; Wells and Horns 1992). African-American women are reported to have higher mortality rates in breast cancer relative to white women. It can be partly explained by the racial difference in late-stage presentation of breast cancer. Lannin *et al.* (1998) found that being African-American and having low income were more than three times more likely to be diagnosed in advanced-stage breast cancer. In addition, no private health insurance, delaying seeing doctors, lack of transportation, less utilization of mammography, and cultural beliefs are significant predictors for late-stage diagnosis of breast cancer (Ayanian *et al.* 1993; Lannin *et al.* 1998; McCarthy *et al.* 1998). Although Hispanic women have lower breast cancer incidence than non-Hispanic whites, advanced stage of breast cancer is more likely found among Hispanic women than non-Hispanic white due to their lower SES (Bentley *et al.* 1998).

More census tracts tested significant in racial disparities of breast cancer mortality than in late-stage diagnosis. This implies that different underlying factors impact the two racial disparities. Not only the stage of breast cancer at diagnosis plays an important role in determining the survival of breast cancer, but also other factors such as treatment

options and financial support are critical in ultimately saving women's lives. The lower R squares further demonstrates that stage of diagnosis is a significant predictor, but not the only one for racial disparities in breast cancer mortality.

The results of the bivariate linear regression indicated that the spatial relationships of racial disparities between late-stage diagnosis and breast cancer mortality were stronger for RR statistic than for RD statistic in that more census tracts were not taken into account using RR statistic due to the missing values. The weak and significant spatial autocorrelations were detected within the residual of linear regressions using the LISA (Moran's *I*) statistics. A geographically-weighted regression (GWR) was adopted to investigate the goodness-of-fit of the model at the local scale. The geographical weight was implemented into the linear regression and the local spatial connections of racial disparities between late-stage diagnosis and breast cancer mortality varied across space.

A strong spatial connection was detected for racial disparities between late-stage diagnosis and breast cancer mortality in the logistic regression. Moran's *I* tested the significant spatial autocorrelation within the residuals of the logistic regression. However, the logistic regression with the geographical weight cannot converge due to a number of missing values and smaller number of census tracts with significant value. Although the R squares was improved using the GWR analysis compared with the ones in the linear regression, the lower values of the R square call for further investigation of other risk factors in order to fully explain racial disparities in breast cancer mortality.

CHAPTER V

RACIAL DISPARITIES OF BREAST CANCER MORTALITY: RISK FACTORS

The results in Chapter IV indicate that racial disparity in late-stage diagnosis of breast cancer is a significant predictor for racial disparities in breast cancer mortality at the census tract level. However, the values of R^2 in both linear and logistic regression analyses suggest that other potential factors may also play a role in the higher mortality rates recorded among African-American and Hispanic women. This chapter investigates whether covariates such as demographic characteristics, socioeconomic status, and lack of spatial accessibility can explain some of the significant racial disparities in breast cancer mortality in African- American and Hispanic women.

Since a few census tracts were identified as having significant racial disparities in breast cancer mortality in terms of rate ratio (RR) measurement, only racial disparities measured with rate difference (RD) was considered to investigate the risk factors responsible for racial disparities of breast cancer mortality. Significance of a census tract was utilized as the dependent variable in a logistic regression that used predictors including socioeconomic status, demographic composition, and lack of spatial accessibility to mammography facilities. In a broader context, the exploration of additional variables helps understand the mechanism of racial disparities in breast cancer

mortality from community environmental perspectives and can be used to identify potential modifiable risk factors to guide future development of intervention programs.

This chapter contains five sections. Section 5.1 presents the descriptive statistics of demographic and socioeconomic factors. Section 5.2 illustrates the measurements and statistics of lack of spatial access to mammography facilities. Section 5.3 describes the results of a factor analysis that identifies the common underlying components that is used as covariates for the logistic regression in Section 5.4 to lessen the multicollinearity within a dataset. Section 5.5 concludes the chapter with a thorough discussion and outlines the implications of these analyses.

5.1 Descriptive Statistics of Demographic and Socioeconomic Factors

The demographic and socioeconomic information was retrieved from 2000 Census Summary File 3 (SF3). Histogram and box plot were utilized as exploratory data analysis tools to examine the statistical distribution of variables and help identify outliers in a number of datasets. Asterisks in a box plot denote the outliers in a dataset.

Demographic variables include percentage of African-Americans in a census tract, percentage of Hispanics, percentage of minority, and population density for each census tract. Percentage of minorities was calculated by adding the percentage of African-Americans and Hispanics together.

Table 5.1 lists the descriptive statistics of the four demographic variables. Both the percentages of African-Americans and Hispanics were calculated by the number of their population by the total population for each census tract. The average percentages of African-Americans and Hispanics were 12% and 31% for all census tracts in Texas. Most

census tracts had more Hispanic population than African-Americans. The percentage of African-Americans is positively skewed with a number of outliers and extreme values, as shown in Figure 5.1a. The distribution of percentage of Hispanic population is less skewed than percentage of African-Americans (Figure 5.1b). The mean of average percentage of minority is 43% of Hispanics and African-Americans for all census tracts in Texas. Figure 5.1c shows the percentage of minority follows the normal distribution. Population density per square km was computed by dividing the total population by the area of each census tract. The distribution of population density is positively skewed with numerous outliers (Figure 5.1d). In summary, the histograms and box plots in Figure 5.1 indicate that percentage of African-American and population density does not obey the normal distribution while percentages of Hispanic and minority conform to the shape of a normal distribution.

Table 5.1 Descriptive statistics of individual demographic variables

Variable	Description	Mean	Std. Deviation	Skewness	Kurtosis
Percentage of African-Americans	Total number of African-American by total population	.12	.19	2.58	6.71
Percentage of Hispanics	Total number of Hispanic by total population	.31	.28	1.04	-.14
Percentage of minority	Total population of African-American and Hispanics by the total population	.43	.30	.45	-1.16
Population density	Total population by the area (sq km)	1,079.26	1,219.53	3.17	30.01

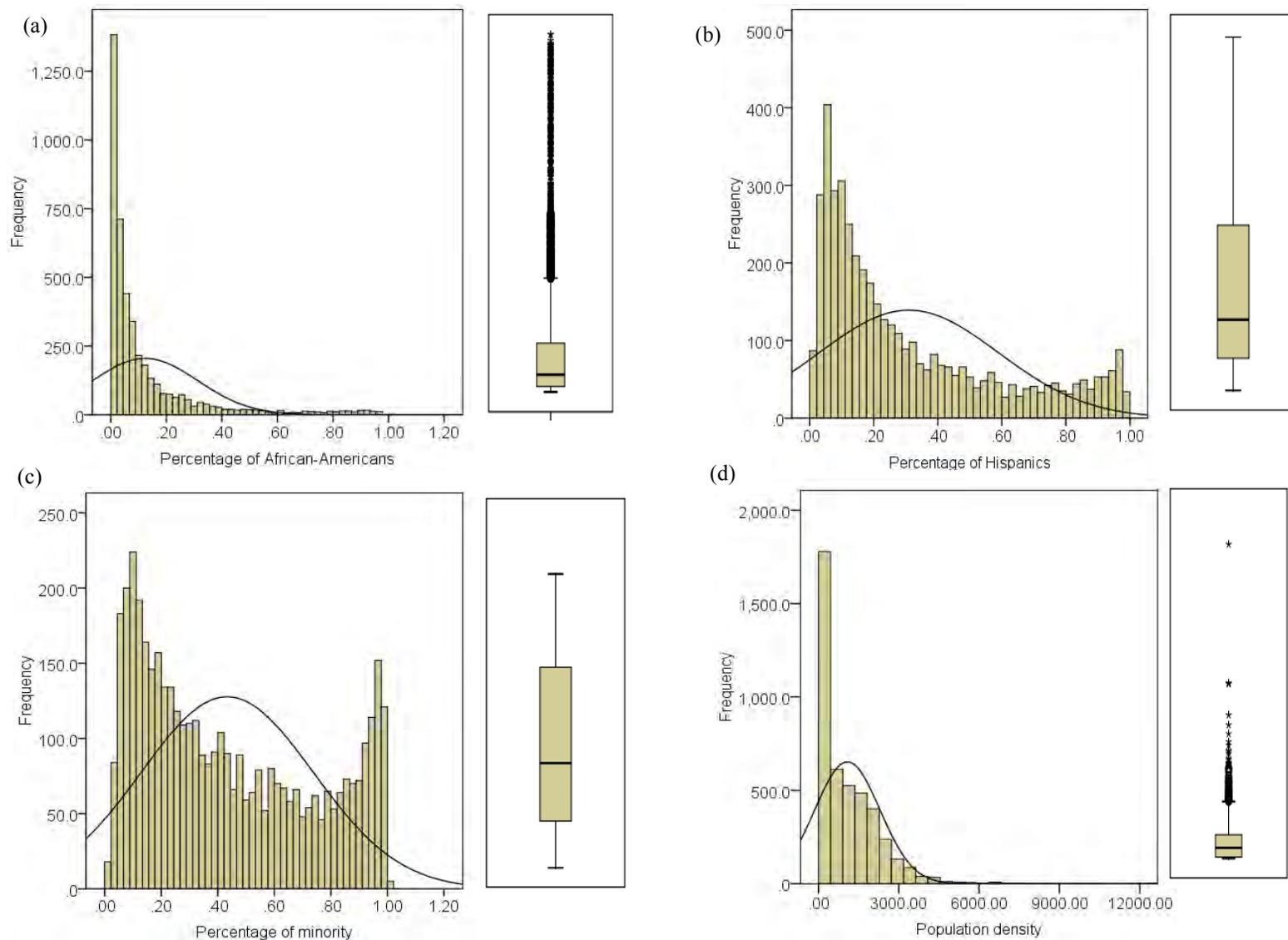


Figure 5.1 Histograms and box plots of demographic variables: percentage of African-Americans (a), percentage of Hispanics (b), percentage of minority (c), and population density (d).

Socioeconomic variables include rural/urban residence, education attainment, unemployment rate, median household income, and poverty level. The descriptive statistics for the five individual socioeconomic variables are summarized in Table 5.2. Percentage of rural population was obtained as the ratio of the population living in rural areas over the total population at each census tract. About 62% of the census tracts (2,721 out of 4,388) have no population living within rural areas. About 12% of the census tracts (532 out of 4,388) have their entire population living within rural areas. An average 20.32% population was living in rural areas for all census tracts in Texas. Percentage of rural population is positively skewed with a number of outliers on both ends (Figure 5.2a).

Percentage of population with less than a college education was calculated by dividing the total number of females without a college degree by the total female population at each census tract. The mean of the education attainment index is 52.22% females with less than a college education. The histogram and box plot illustrate that the education attainment follows a reasonably symmetric normal distribution (Figure 5.2b). An average 7.11% of census tract population did not have a job in 2000. The data distribution of percentage of unemployment is positively skewed and has a sharp peak relative to the normal distribution (Figure 5.2c).

The percentage of population living under the poverty line was calculated as the ratio of the population living under the federal poverty line in 2000 over the total population at each census tract. The average population living under the poverty line was 16.22% for a census tract in Texas. The histogram of percentage of population living under poverty line has skewness of 1.254 and Kurtosis of 2.054 with a larger number of

census tracts with more than 20% population living under poverty line (Figure 5.2d). The average of median household income was \$41,184 for a census tract with a positive skewness and peak. A number of outliers were observed in the box plot of average median household income (Figure 5.2e). The descriptive statistics complement the histograms and box plots and indicate that percentage of population with unemployment and median household income do not conform to a normal distribution. To be consistent, a normal-score transformation was performed for all the five SES measurements. Normal distribution is required for the subsequent factor and regression analyses.

Table 5.2 Descriptive statistics of individual socioeconomic variables

Variable	Description	Mean	Std. Deviation	Skewness	Kurtosis
Percentage of rural population	Total population living in rural areas by the total population at each census tract	20.32	36.00	1.52	.56
Percentage of population with less than a college education	Total females with less than a college education by the total female population at each census tract	52.22	20.43	-.30	-.66
Percentage of population under unemployment	Population under unemployment by the total population in labor	7.11	6.11	3.77	33.05
Percentage of population under the poverty line	Population living federal poverty line by the total population at each census tract	16.22	12.13	1.25	2.04
Median household income	Median household income in 1999	41,184.49	21,489.25	2.16	8.15

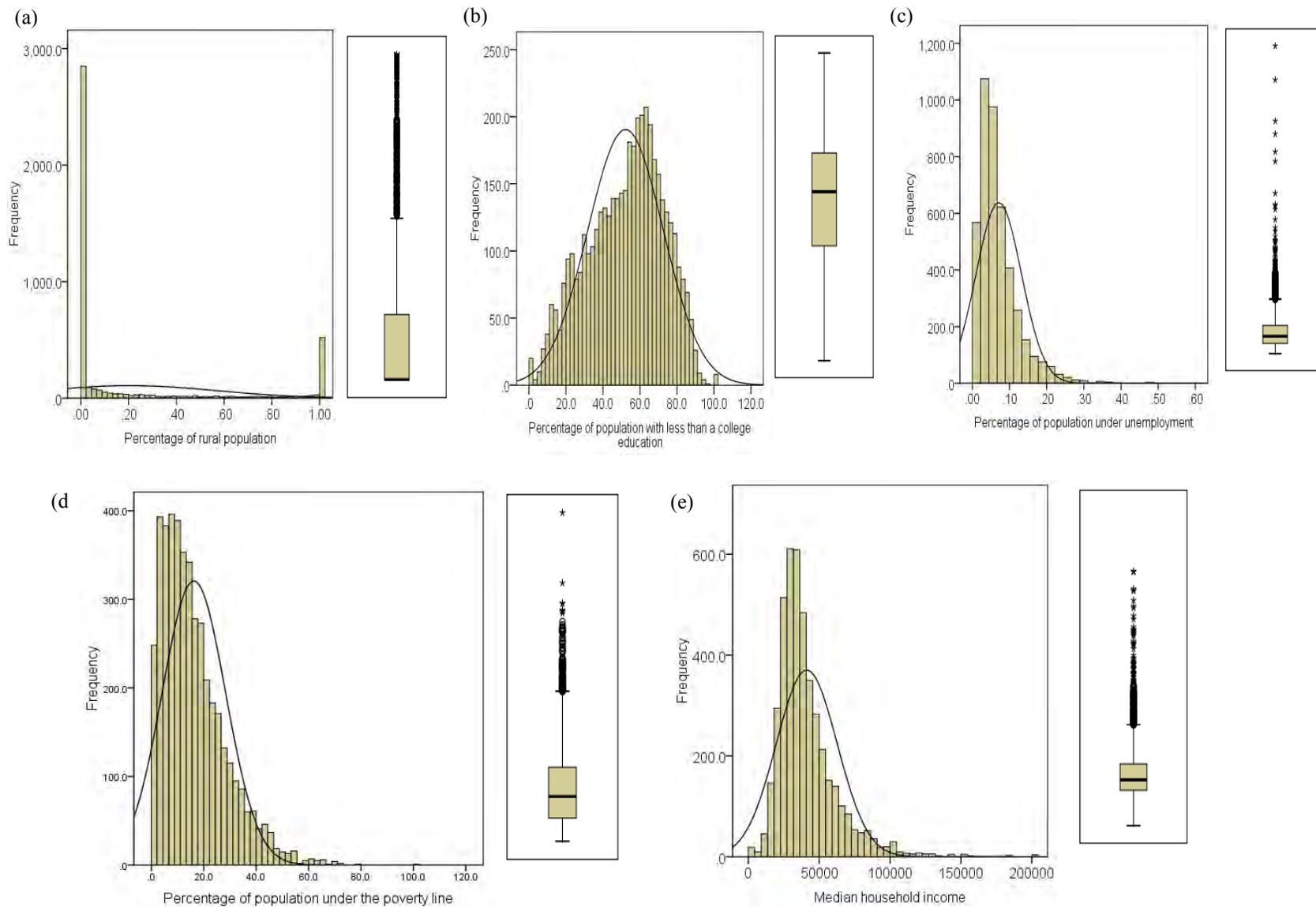


Figure 5.2 Histograms and box plots of socioeconomic variables: percentage of rural population (a), percentage of population with less than a college education (b), percentage of population under unemployment (c), percentage of population under the poverty line (d), median household income (e).

5.2 Spatial Accessibility of Mammography Facilities

Accessibility to health care is a major concern for health planners and health policy-makers who want to address health inequity. Poor access to health services results in more advanced diseases, which cost even more to treat (Haynes *et al.* 1999). The concept of accessibility is complicated by the interactions of geographic, financial, cultural, and functional components. The mammography facilities data were obtained from the Texas Mammography Accreditation Program, Texas Department of State Health Services.

The facilities operating in 2000 were gathered to analyze the spatial access to mammography facilities. There were 605 mammography facilities that were run in 2000 with the Federal Drug Administration (FDA) approval. The mammography facilities were geocoded using a customized program that fed each address record into Google map database and retrieved the latitude and longitude as a result. Mammography locations were then mapped based on the latitude and longitude information. This study did not utilize the conventional means to address-match all facility records to the U.S. Census Bureau Tiger/Line Files, because the latitude and longitude identification provides more accurate locations than geocoding based on street reference layer. Four mammography facilities with US highway address did not come up with the correct latitude and longitude which were manually retrieved.

Previous studies often utilized Euclidean distance from the centroid of population center to the nearest mammography facilities to measure lack of spatial access to health care facilities (Gumpertz *et al.* 2006; Meersman *et al.* 2009). GIS techniques can be used to model the Euclidean distance using the near command or buffer function. However,

Euclidean distance only provides a crude measure of physical accessibility to health care facilities. Road networks based distance is a more accurate measure of separation between a person's residence and the location of a health care facility.

Therefore, this research applied network distance and travel time from the centroid of 4,388 census tracts to the closest mammography facilities as a measure of lack of spatial accessibility. In addition, the average network distance and travel time to the five closest mammography facilities were derived to represent the choices available to a patient in choosing the preferred facility. This set of five choices takes into account potential constraints because different facilities may require certain types of health insurances, may have different operation hours, and may be convenient based on their office locations (Tarlov *et al.* 2009).

Street network data with detailed information of road length and speed limit was obtained from the Environmental Systems Research Institute (ESRI). The average network travel distance and travel time were calculated using the Network Analysis Extension of ArcInfo 9.3. Moreover, mammography density was calculated as well by the number of mammography facilities within a 30-mile buffer of each census tract centroid per 1,000 females, which provides an estimate of greater choice. Distance and density theoretically represent different angles of potential health care accessibility and contribute to health outcomes in a distinctive way. Table 5.3 summarizes the five variables of lack of spatial accessibility to mammography facilities under consideration.

Table 5.3 Descriptive statistics of lack of spatial accessibility to mammography facilities

Variable	Description	Std.	Skewness	Kurtosis	
		Mean	Deviation		
Density of mammography facilities	Number of mammography facilities per 1000 females within 30 miles buffer of each census tract centroid	69.45	1,480.71	64.08	4,189.24
Travel distance to the closest facility	Travel distance from centroid of each census tract to the closest facility on the transportation network	7.29	12.22	6.54	71.09
Average travel distance to the five closest facilities	Average of travel distance from centroid of each census tract to the five closest facilities on the transportation network	13.18	16.15	3.68	25.00
Travel time to the closest facility	Travel time from centroid of each census tract to the closest facility on the transportation network	11.97	17.47	5.63	56.84
Average travel time to the five closest facilities	Average of travel time from centroid of each census tract to the five closest facilities on the transportation network	19.82	22.82	3.39	21.26

Detailed statistics were computed for the five measurements of lack of spatial accessibility to mammography facilities in Texas. The spatial distribution, histograms, and box plots of the five variables were shown in Figures 5.3 through 5.5. The mammography density had a wide spread distribution with an average of about 70 mammography facilities for 1,000 females within 30-mile buffer from the centroid of each census tract. Some census tracts had a great number of mammography facilities within the 30-mile buffer with only a few females (Figure 5.3). The spatial distribution map of mammography density illustrates that most mammography facilities were concentrated in metropolitan areas of Houston, Dallas, and Austin-San Antonio (Figure 5.3).

The mean of network travel distance from a centroid of the census tracts to the closest mammography facility was 7.29 miles, while that of average network travel distance to the five closest mammography facilities was 13.18 miles (Figures 5.4a and 5.4b). The average travel time from a centroid of the census tracts to the closest mammography facilities was 11.97 minutes and mean of average travel time to the five closest facilities was 19.82 minutes (Figures 5.5a and 5.5b). The spatial distribution of the five measurements illustrates that longer distance and more driving time are required to reach mammography facilities in the Southwest of Texas with the longest travel distance of 218.15 miles to the closest mammography facility (a five-hour trip). Metropolitan areas, including Houston, Dallas, and Austin-San Antonio, had shorter driving distance and less travel time because most mammography facilities were located within these urban areas.

An examination of the maps of driving distance and time to the mammography facilities reveals the existence of edge effect along the border of Texas and other states, since the study only included the mammography facilities within the state of Texas. For example, in the Northeast Texas between the border of Texas and Oklahoma, longer average driving distance and travel time were observed because some facilities located in the state of Arkansas were not considered in the analysis. Women living along the border may have the tendency to have mammogram in the adjacent state. The histogram and box plots for the five measurements of lack of spatial accessibility to mammography facilities were strongly skewed with numerous outliers (Figures 5.3 through 5.5). Thus, the measurements of lack of spatial accessibility of mammography facilities were first normalized in order to investigate how these predictors impact the significance of racial disparities in breast cancer mortality.

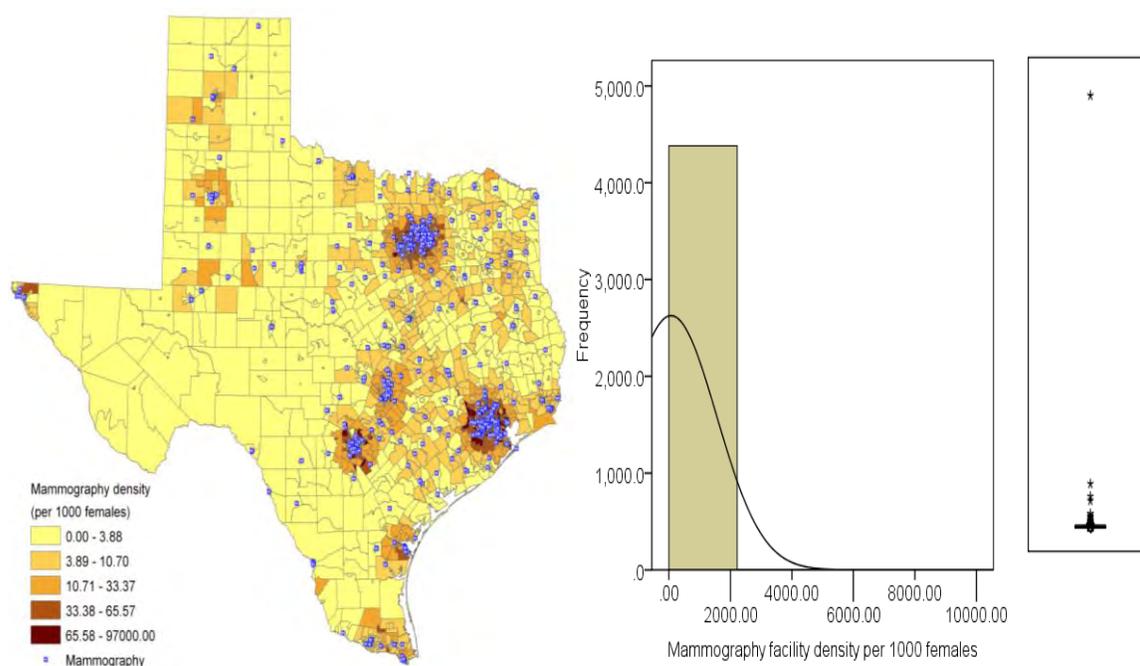


Figure 5.3 Spatial distribution, histogram and box plot of mammography facility density per 1,000 females.

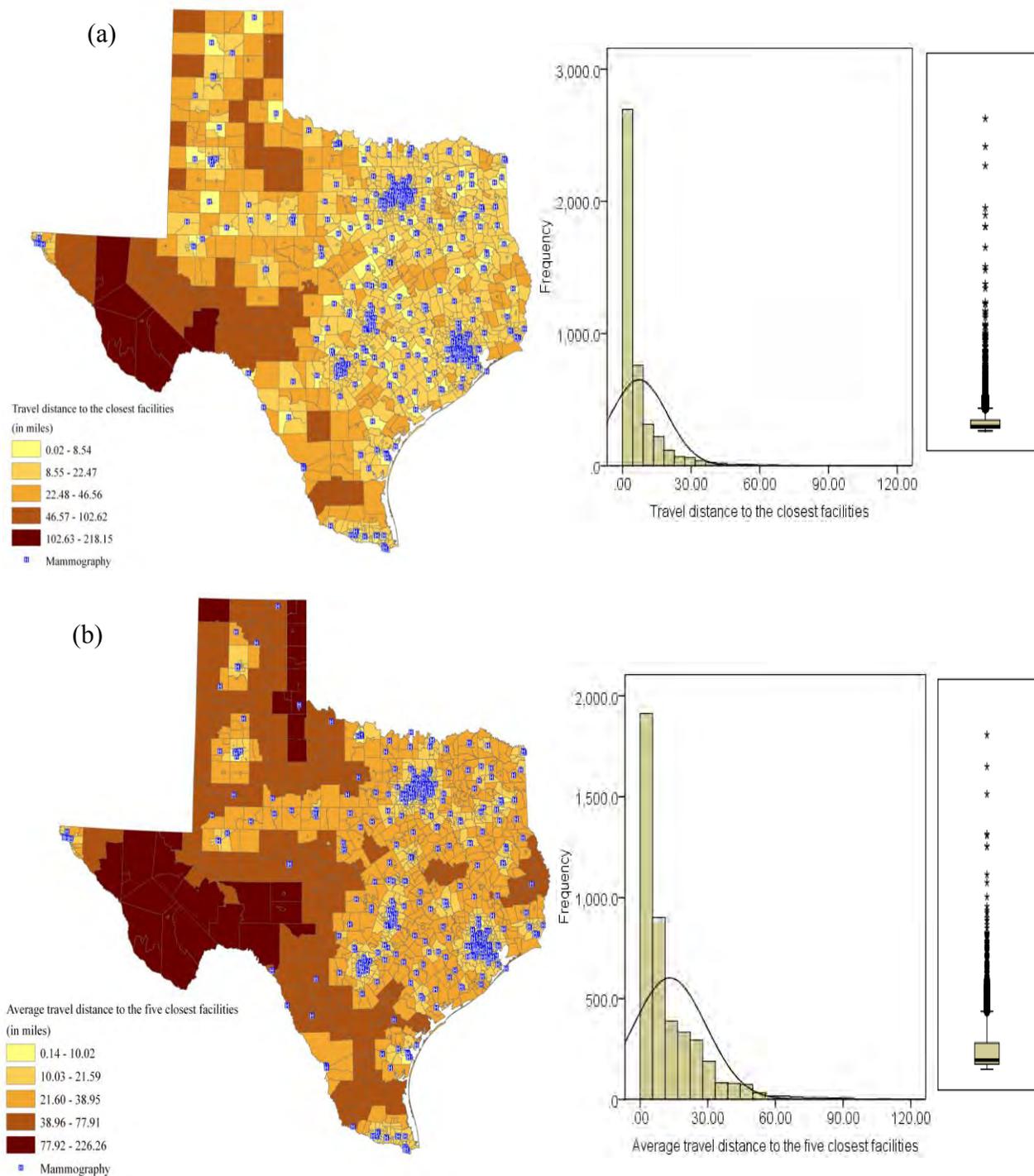


Figure 5.4 Spatial distributions, histograms, and box plots: (a) travel distance to the closest facility, and (b) average travel distance to the five closest facilities.

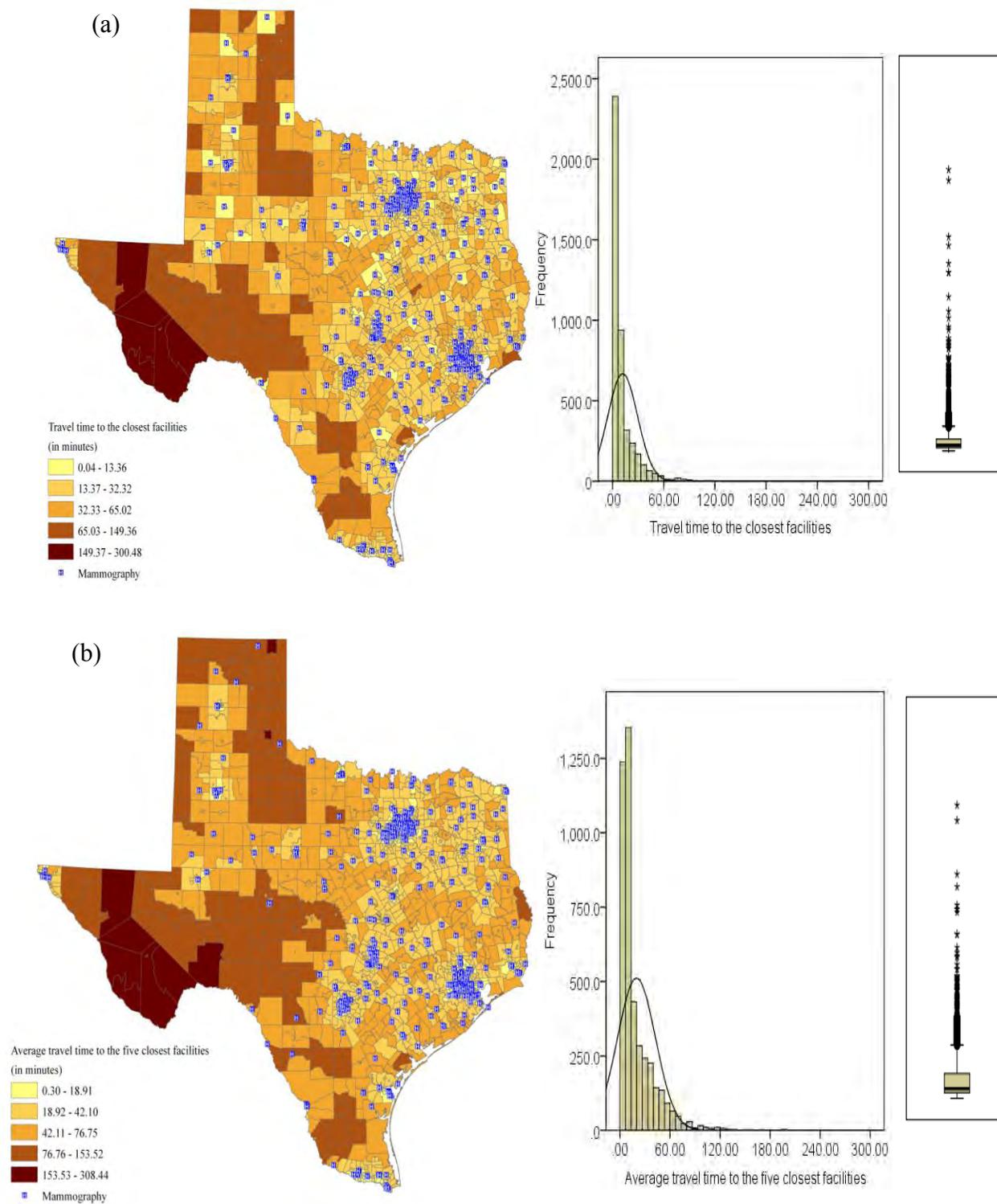


Figure 5.5 Spatial distributions, histograms and box plots: (a) travel time to the closest facility, and (b) average travel time to the five closest facilities.

5.3 Results of Factor Analysis

Multicollinearity is a frequent issue when building multivariate regression models. It occurs when two or more predictor variables are highly correlated. Factor analysis can assess the inter-correlations within an array of variables for the purpose of revealing the common structures underlying all variables under consideration (Harman 1976). Principle component extraction method in factor analysis leads to the identification of a manageable subset of factors, often referred to as components. Factors with eigenvalues greater than 1 are considered as valid components in the study. Details about the factor analysis used in this section can be found in Chapter III as well as in the paper by Thurstone (1930).

Table 5.4 lists all fourteen predicting variables reflecting socioeconomic status, demographic characteristics, and lack of spatial accessibility to mammography facilities. The amount of total variance that can be accounted for by each variable is also shown in Table 5.4. Communality is defined as the percentage of variance in a particular variable explained by the common underlying factors. The initial communalities were set to 1.0 and extraction communalities represented the estimates of common variance of each variable which could be accounted for by the unobserved components. High values of extraction communalities in Table 5.4 indicate that the extracted components reflect all the variables well in terms of explained variance.

According to the Kaiser Criterion (1960), only components with an arbitrary eigenvalue greater than 1.0 were retained for the factor analysis. Three components fulfill this condition and are listed in Table 5.5. The cumulative variance in Column 3 of Table 5.5 indicates that the first three components can explain 75% of the total variance in the

dataset. The scree plot is a graphical approach to extract the principal components. Scree is a geological term referring to collected debris on the lower part of a hilly slope. Only the components with sharp slopes are identified in the scree plot. Figure 5.6 illustrates that the smooth decrease occurs on component 3 and then levels off. Scree plot indicates there are three principal components underlying all the fourteen variables measuring socioeconomic status, demographic characteristics, and lack of spatial accessibility to mammography facilities.

Table 5.4 Variables and communalities in the factor analysis

	Initial	Extraction
Percentage of minorities	1.000	.789
Percentage of Hispanics	1.000	.753
Percentage of population less than a college education	1.000	.759
Percentage of population living under the poverty line	1.000	.821
Median household income	1.000	.748
Percentage of population with unemployment	1.000	.622
Mammography density	1.000	.545
Population density	1.000	.723
Percentage of rural population	1.000	.568
Travel distance to the closest facilities	1.000	.777
Average travel distance to the five closest facilities	1.000	.862
Travel time to the closest facilities	1.000	.780
Average travel time to the five closest facilities	1.000	.875
Percentage of African-Americans	1.000	.910

Table 5.5 The variance accounted for by the successive components in the factor analysis

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.281	37.723	37.723	5.136	36.686	36.686
2	4.106	29.329	67.052	4.190	29.926	66.612
3	1.147	8.190	75.242	1.208	8.630	75.242
4	.874	6.240	81.482			
5	.660	4.717	86.199			
6	.484	3.461	89.659			
7	.395	2.822	92.482			
8	.341	2.438	94.919			
9	.278	1.987	96.906			
10	.196	1.402	98.307			
11	.131	.937	99.244			
12	.084	.600	99.844			
13	.018	.126	99.970			
14	.004	.030	100.000			

Scree Plot

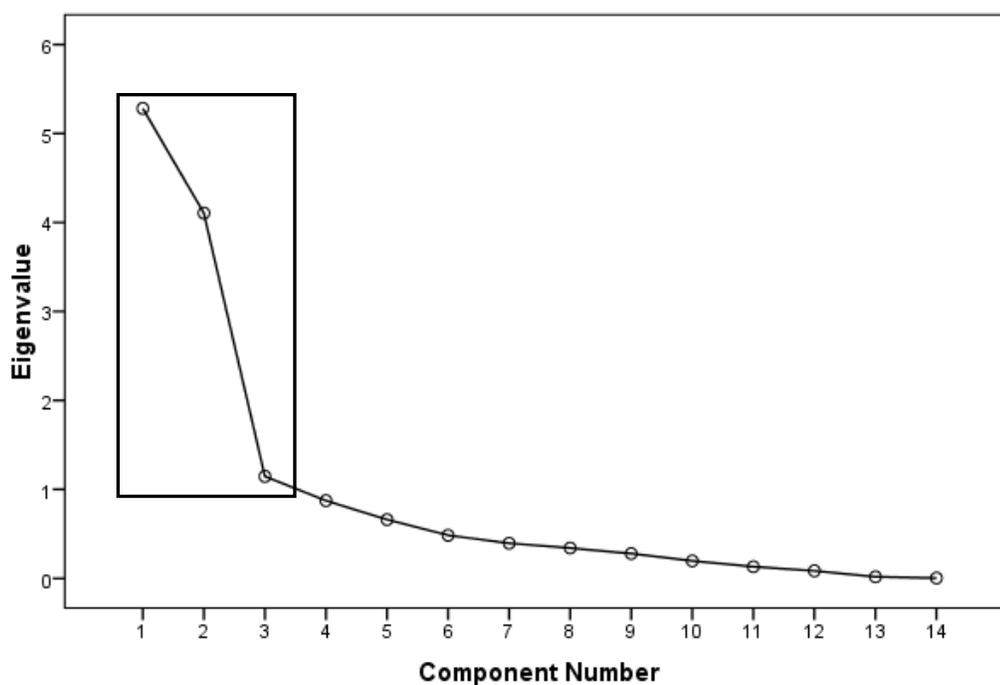


Figure 5.6 Scree plot of eigenvalues by the component number in the factor analysis.

Table 5.6 lists the rotated component matrix, which provides information about which individual variables are contained within each component. The first component included seven variables including mammography density, population density as well as percentage of rural population, travel distance, and travel time to the closest facilities as well as average travel distance and time to the five closest facilities. All the variables were highly correlated to the first component with negative correlation with mammography and population densities and positive correlation with travel distances and time to the mammography facilities. The census tracts with more mammography facilities within a 30-mile buffer generally had shorter driving distance and less travel time, which explained the negative associations between these variables. The first component represents a comprehensive measure of the lack of spatial accessibility to mammography facilities since this component is positively correlated with distance to mammography clinics.

The second component contained six variables including percentage of minorities, percentage of Hispanics, and percentage of population with less than college education, percentage of population living under the poverty line, median household income as well as percentage of population that were unemployed. The second component revealed the construct of the socioeconomic factor underlying these variables. All the individual variables within the second component are highly correlated with a minimum coefficient of 0.721. Median household income has a negative correlation of -0.857 to the second component, but all other variables have positive correlations because higher median income corresponds to the lower values of other socioeconomic variables including unemployment rate and poverty level.

The third component included a single variable which is percentage of African-Americans with correlation coefficient of 0.924. It is interesting that percentage of African-Americans stood out by itself and has a negatively weak correlation with the variable of percentage of Hispanic. Table 5.7 summarizes the component coefficients for all the fourteen variables. The factors scores were calculated by multiplying the variable values with its corresponding component coefficients. The three components are representative for all the fourteen variables with less than 25% variance unaccounted for. Thus, the first three factors were utilized in the multiple logistic regression analysis to embrace all the information represented by the fourteen variables measuring demographic characteristics, socioeconomic status, and lack of spatial accessibility to mammography facilities.

Table 5.6 Rotated component matrix in the factor analysis

	Component		
	1	2	3
Percentage of minority	-.244	.843	.140
Percentage of Hispanic	-.242	.721	-.419
Percentage of population less than a college education	.166	.854	-.045
Percentage of population living under the poverty line	.008	.906	.022
Median household income	-.101	-.857	-.062
Percentage of population with unemployment	-.046	.777	.125
Mammography density	-.640	-.107	.351
Population density	-.837	.141	.052
Percentage of rural population	.745	-.105	-.046
Travel distance to the closest facilities	.881	-.010	.020
Average travel distance to the five closest facilities	.925	.060	-.061
Travel time to the closest facilities	.882	-.036	.014
Average travel time to the five closest facilities	.933	.023	-.064
Percentage of African-American	-.180	.152	.924

Table 5.7 Component score coefficient matrix in the factor analysis

	Component		
	1	2	3
Percentage of minority	-.027	.198	.095
Percentage of Hispanic	-.074	.172	-.385
Percentage of population with less than a college education	.041	.207	-.030
Percentage of population living under poverty line	.016	.217	.014
Median household income	-.037	-.206	-.057
Percentage of population with unemployment	.011	.185	.100
Mammography density	-.103	-.036	.251
Population density	-.164	.022	-.025
Percentage of rural population	.146	-.015	.023
Travel distance to the closest facilities	.181	.009	.090
Average travel distance to the five closest facilities	.184	.027	.023
Travel time to the closest facilities	.180	.003	.086
Average travel time to the five closest facilities	.185	.018	.022
Percentage of African-American	.042	.029	.781

5.4 Results of Multivariate Logistic Regression Analysis among African-American Women

A detailed description of the logistic regression model is provided in Chapter IV.

In a regression model, the null hypothesis for each predictor is that the regression coefficient is equal to zero. The significance level was set to 0.05 for the logistic regression analysis. If the p value for each individual predictor is greater than 0.05, then it means that the coefficient for that independent variable is not significant and the variable could be dropped without affecting the performance of the full model. The significance of logistic regression model was assessed using a Chi-square distribution.

A logistic regression model was constructed to assess the significance of racial disparities in breast cancer mortality using as predictors including the significance of late-stage diagnosis and three components obtained in the factor analysis. The significance of

racial disparities in breast cancer mortality was identified for each census tract in terms of rate difference (RD) measurement in chapter IV. A total of 278 census tracts were identified to have significantly higher mortality rates for African-American women in comparison with non-Hispanic whites, while 328 census tracts were identified for Hispanic women.

In the logistic regression, the census tracts were coded as dichotomous variable with a value of either $_0$ (not significant) or $_1$ (significant). Significance of racial disparities in late-stage diagnosis of breast cancer is also an important predictor for racial disparities in breast cancer mortality. Thus, census tracts were also coded as binary value based on the significance of racial disparities in late-stage diagnosis in terms of RD measurements. The reference cell value was set as $_0$ (not significant) in the logistic regression model. The three comprehensive factors extracted from the factor analysis included demographic characteristics, socioeconomic status, and lack of spatial accessibility and were utilized as the continuous independent variables in the logistic regression model.

R^2 values of Cox and Snell (0.19) and Nagelkerke (0.47) can be used to quantify the strength of a set of independent variables as predictors of the dependent variable in a logistic regression model. Nagelkerke R^2 is similar to the adjusted R^2 in a linear regression. The adjusted R^2 value of 0.47 indicates that the model with the set of independent variable did a reasonably good job to predict the possibility that a census tract displays significant racial disparities in breast cancer mortality.

Table 5.8 summarizes the logistic regression results with estimated parameter values, significance of each individual independent variable, odds ratio, and confidence

interval of odds ratio. If the odds ratio is greater than 1 for an individual term in a logistic regression model, it indicates that a census tract is more likely to have significant racial disparities of breast cancer mortality. Table 5.9 lists the correlation coefficients for all the four independent variables established in a logistic regression model. Existence of multicollinearity was ruled out using principle component analysis.

Table 5.8 indicates that all the independent variables were significant in the logistic model with p value less than 0.000006. If a census tract has significant racial disparities in late-stage breast cancer diagnosis, that census tract is 4.08 times more likely to display significant racial disparities in breast cancer mortality among African-American women. The lack of spatial accessibility factor has an odds ratio of 0.62, which suggests that a census tract with less physical access to mammography facilities is less likely to display significant racial disparities. The percentage of African-Americans (demographic factor) played an important role in determining the possibility that a census tract tested significant for racial disparities in breast cancer mortality (OR=3.45, 2.89-4.13).

The global Moran's I was applied to the residuals of logistic regression model constructed for African-American women. The spatial autocorrelation was very weak with a Moran's I value of -0.00026. There was no significant spatial dependency after accounting for these four independent variables because the p value was 0.492.

Therefore, further geographically-weighted logistic regression is unnecessary to explore the local relationship between racial disparities in breast cancer mortality and other predictors.

Table 5.8 Logistic regression results for African-American women: racial disparities in breast cancer mortality as dependent variable (RD) and racial disparities in late-stage diagnosis (RD), demographic characteristics, socioeconomic status, and lack of spatial accessibility as independent variables

Independent variables	Parameter estimate	Standard error	<i>p</i> value	Odds ratio	Odds ratio C.I. (95%)
Significance of racial disparities in late-stage diagnosis	1.41	0.13	< 0.000001	4.08	2.61 - 6.39
Lack of spatial accessibility factor	-0.48	0.23	< 0.000001	0.62	0.50 - 0.76
Socioeconomic factor	0.89	0.11	0.000006	2.43	1.95 - 3.04
Demographic factor	1.24	0.11	< 0.000001	3.45	2.89 - 4.13

Table 5.9 Correlation coefficients of the logistic regression model for African-American women

Independent variables	Significance of racial disparities in late-stage diagnosis	Lack of spatial accessibility factor	Socioeconomic factor	Demographic factor
Significance of racial disparities in late-stage diagnosis	1.0	-0.09	0.43	-0.25
Lack of spatial accessibility factor	-0.09	1.0	0.004	0.07
Socioeconomic factor	0.43	0.004	1.0	-0.11
Demographic factor	-0.25	0.07	-0.11	1.0

5.5 Results of Multivariate Logistic Regression Analysis among Hispanic Women

The Cox and Snell R^2 was 0.20 and Nagelkerke (adjusted) R^2 was 0.46 in the logistic regression analysis conducted for Hispanic women. The R^2 indicated that about 46% of the variance within the dependent variable could be explained by the set of predictor variables including racial disparities in late-stage diagnosis, demographic characteristics, socioeconomic status, and lack spatial accessibility to mammography facilities.

The logistic regression results are summarized in Table 5.10 with significance level and odds ratio for each independent variable. Table 5.11 shows the correlation coefficients between the four independent variables and smaller coefficients indicate that the independent variables are not highly correlated and not suitable in the logistic regression model. The lack of spatial accessibility was not a significant term in the logistic regression model with the p value of 0.28, while all other factors tested significant including significance of racial disparities in late-stage diagnosis, demographic characteristics, and socioeconomic status.

Odds ratio is a measurement of the probability of occurrence for the event of $_1$ of the dependent variable. The regression model for Hispanic women indicated that if a census tract was tested significant for racial disparities in late-stage diagnosis, this tract had 4.42-folds likelihood to display significant racial disparities for breast cancer mortality. The factor of lack of spatial accessibility did not have a significant impact; the p value was greater than the significance level (0.05) and its odds ratio had a 0.80-1.07 (95% confidence interval). Therefore, dropping the spatial accessibility factor in the

analysis won't have a significant effect on the prediction capability of the logistic regression model established for Hispanic women.

Census tracts that tested significant in racial disparities of breast cancer mortality were found within the metropolitan areas of Houston, Dallas, and Austin-San Antonio as well as along the Southwest border for Hispanic women. Examining the spatial distribution of five measurements of lack of spatial accessibility to mammography facilities showed that it took shorter time and travel distance to the available mammography facilities within the urban areas relatively to the area along the Southwest border of Texas. As a consequence, lack of spatial accessibility mitigated its influence on the significance of racial disparities in breast cancer mortality constructed in the logistic regression model across space.

The odds ratio was 5.30 with the confidence interval of 4.26-6.59 for socioeconomic factor, which played an important role in predicting whether a census tract had significant racial disparities in breast cancer mortality. The percentage of African-Americans as the demographic factor was a significant predictor in the logistic regression model. Its odds ratio was 0.73 with a confidence interval of 0.63-0.84. This indicated that a census tract with a smaller percentage of African-Americans was more likely to test significant in racial disparities of breast cancer mortality for Hispanic females.

The global Moran's I was used to test whether the residuals of the logistic regression model established for Hispanic women had any spatial autocorrelation. The Moran's I index was 0.034 indicating a weakly but significant ($p=0.001$) spatial dependency for the residuals. Geographic weight can be applied in the logistic regression

to take the neighborhood effect into consideration and improve the performance of the regression model. However, the geographically-weighted logistic regression could not converge because there were quite a few missing values and not enough counts of '1' value for the independent variable. Failure to converge often happens to the logistic regression model. So the results could not be provided for the geographically-weighted logistic regression model.

Table 5.10 Logistic regression results for Hispanic women: racial disparities in breast cancer mortality as dependent variable (RD) and racial disparities in late-stage diagnosis (RD), demographic characteristics, socioeconomic status, and lack of spatial accessibility as independent variables

Independent variables	Parameter estimate	Standard error	P value	Odds ratio	Odds ratio C.I. (95%)
Significance of racial disparities in late-stage diagnosis	1.49	0.20	<0.000001	4.42	3.01 - 6.51
Factor of lack of spatial accessibility	-0.08	0.07	0.279867	0.92	0.80 - 1.07
Socioeconomic factor	1.67	0.11	<0.000001	5.30	4.26 - 6.59
Demographic factor	-0.31	0.07	0.000017	0.73	0.63 - 0.84

Table 5.11 Correlation coefficients of logistic regression model for Hispanic women

Independent variables	Significance of racial disparities in late-stage diagnosis	Factor of lack of spatial accessibility	Socioeconomic factor	Demographic factor
Significance of racial disparities in late-stage diagnosis	1.0	-0.09	0.43	-0.25
Factor of lack spatial accessibility	-0.09	1.0	0.004	0.07
Socioeconomic factor	0.43	0.004	1.0	-0.11
Demographic factor	-0.25	0.07	-0.11	1.0

5.6 Conclusions

This study assessed the association between significance of racial disparities in breast cancer mortality and predictors including the significance of racial disparities in late-stage diagnosis, demographic characteristics, socioeconomic status, and lack of spatial accessibility factors. All fourteen variables include a number of outliers; hence a normal-score transformation was required before the analysis could proceed.

A factor analysis was performed in order to reduce the high correlation among the predictor variables. In this study, the factor analysis concluded that the percentage of population was an important predictor of lack of spatial accessibility to mammography facilities. Thus, percentage of rural population was grouped into the factor of lack of spatial accessibility (Table 5.6). The rural and urban areas often show distinctive accessibility to health care facilities, which impacts the breast cancer survival and mortality rates. Breast cancer incidences are reported to be different between the rural and urban areas. For example, urban areas were reported to have 1.06-1.83 higher breast cancer incidence than rural areas in the United States (Hall 2005; Gregorio *et al.* 2002; Prehn and West 1998; Weiss *et al.* 1997). Remote areas were often reported to lack access to health care (Girgis *et al.* 2000).

As expected, the travel distance and driving time were shorter within urban areas due to the concentration of health care facilities and physicians within urban areas which leads to easier geographical access to health care facilities (Jordan *et al.* 2004). The mammography services were not well-developed in the remote areas, particularly along the Southwest border of Texas with a large Hispanic population. Moreover, Hispanic percentage was identified as a measurement of socioeconomic factor in that Hispanic

population is frequently reported to have lower socioeconomic status, lagging behind non-Hispanic whites in education attainment and median household income (Chapa and Rosa 2004; Chapa and Valencia 1993). The percentage of African-Americans was the only variable referred to the demographic factor in the factor analysis and impacted the significance of racial disparities in breast cancer mortality for both African-American and Hispanic women.

The logistic regression model was constructed to investigate the significance of racial disparities in breast cancer mortality. The predictors included the significance of racial disparities in late-stage diagnosis of breast cancer as well as demographic characteristics, socioeconomic status, and lack of spatial accessibility which were obtained in the factor analysis. All the independent variables tested not highly correlated with each other in the regression models.

First, the logistic regression concluded that a census tract was more than four times more likely to have significant racial disparities in breast cancer mortality if the census tract displayed significant racial disparities in late-stage diagnosis. African-Americans and Hispanics are frequently reported to have a higher likelihood of being diagnosed late, which leads to lower survival and higher mortality rates for the minority groups in addition to cancer treatment difference across racial groups (Boyer-Chammard *et al.* 1999; Chevarley and White 1997; Joslyn and West 2000). For example, African-Americans and Mexicans had 1.4-3.6-folds likelihood to be diagnosed with stage IV breast cancer and were less likely to undertake surgical treatment required by the 2000 National Comprehensive Cancer Network Standards. The treatment difference was responsible for 1.2-3.0 fold greater risk of dying from breast cancer (Li *et al.* 2003).

Second, the odds ratio of spatial accessibility factor in the logistic regression indicated that easier access to mammography facilities did not help alleviate the additional burdens of breast cancer mortality carried by African-American women. The spatial accessibility to mammography facilities had no significant influence on higher breast cancer mortality for Hispanic females compared to non-Hispanic whites. The study presents another piece of evidence on the controversial influence of geographical obstruction on health care utilization in the current literature studies (Arcury *et al.* 2005; Athas *et al.* 2000; Baldwin *et al.* 2004; Nattinger *et al.* 2001; Piette and Moos 1996).

The current study is consistent with the conclusions drawn by Tarlov and colleague in 2009 that the mean network distance from mammography facilities to the residence of breast cancer patients in Chicago had no significant influence on stage diagnosis of breast cancer. However, another study found that the longer driving distance measured using the census tract centroid weighted by its population to the closest mammography facility was a risk factor of advanced stage diagnosis in breast cancer for Hispanic and white women in Los Angeles County (Gumpertz 2006). Meersman and his colleague reinforced Gumpertz' statement pertaining to the effect of distance to mammography facilities in addition to an unexpected finding that lack of English proficiency was a protective factor to have a recent mammogram, which could be justified by the effective outreach program of the Every Woman Counts in low-income Latina communities.

The unique characteristic of the current research is to investigate how spatial accessibility to mammography facilities had an influence on the significance of racial disparities of breast cancer. The census tracts with significant racial disparities in breast

cancer mortality were found within the metropolitan areas of Houston, Dallas, and Austin-San Antonio for both African-Americans and Hispanics. Health care services are generally denser and take less time to drive to, but the results in the study indicated that other more important factors determined the health outcomes among minority groups instead of geographical access to mammography facilities. Besides, the emphasis should be placed on the Southwest border of Texas to reduce the breast cancer mortality for Hispanic women as well. The intervention program could be planned to locate more mammography facilities to improve the physical access to health care.

Third, this study suggests that socioeconomic factor played a critical role in determining if a census tract had significantly higher death rates of breast cancer for African-Americans compared with non-Hispanic white women. Lower socioeconomic level had substantially negative influence on racial disparities of breast cancer mortality. For instance, the study concluded that a census tract with lower socioeconomic status was about five times more likely to experience significant racial disparities in breast cancer mortality.

In geographic regions with lower SES, the vulnerability for minorities such as African-American and Hispanic women increases more than for their white counterparts. The socioeconomic status is an essential driving force for the deterioration of health outcomes experienced by the minority and disadvantaged groups. African-American and Hispanic women cannot receive the equal benefit from current medical advancement and intervention programs of early detection (Chu *et al.* 1999). Consequently, socioeconomic status impacts people's health through multi-facet pathways ranging from affordability of

health insurance, knowledge of health issues, perceptions of early detection, nutrition, and life-style behavior (Baldwin *et al.* 2004; Goodman 1999).

Cost concern is more related to mammography use. For example, high SES individuals among Hispanic women were reported to have higher utilization of mammography facilities (Stein *et al.* 1991). Population-based SES measured at the aggregate level can provide information not only on individual health, but also on the effects of community characteristics on individual health. Community socioeconomic characteristics shape the individual health by making its unique contribution of the neighborhood effect and are highly associated with individual SES (Robert 1998).

Fourth, the percentage of African-Americans had different effects on the significance of racial disparities in breast cancer mortality. An examination of the dataset indicated that African-American and Hispanic women had different geographic distributions in the state of Texas. By taking population distributions into consideration, this study points out the different targeted regions in order to reduce racial disparities of breast cancer mortality for African-Americans and Hispanic women.

Augmenting the physical access to health care facilities through building more health facilities couldn't be an effective approach to reduce racial disparities of breast cancer. Instead, health-policy makers should focus on offering the subsidized health insurance in a short term and improving the socioeconomic status in a long run for minority groups so as to meet the goal of Healthy People 2010. The limitation of this study is to assume that all populations were living in the centroid of each census tract, which could lead to different conclusion with respect to the spatial accessibility in comparison with the individual case.

CHAPTER VI

CONCLUSIONS, INTERVENTIONS, AND FUTURE WORK

This chapter summarizes and concludes the study. The first section recaps the research findings and provides the conclusions. The second section gives a thorough discussion of potential intervention programs based on the assessment of risk factors for racial disparities of breast cancer mortality that was presented in Chapter V. The third section outlines contributions of this research to the literature of cancer disparities. The fourth section points out limitations of the research and suggests a number of emerging research opportunities in breast cancer disparities.

6.1 Conclusions

As stated in Chapter I, the objectives of this research were to: 1) identify the census tracts in Texas where both late-stage diagnosis and breast cancer mortality for African-Americans and Hispanics are significant high compared with non-Hispanic whites, 2) examine the spatial relationship of the significant test of racial disparities between late-stage diagnosis and breast cancer mortality, and 3) investigate the risk factors responsible for significant racial disparities in breast cancer mortality.

In achieving the first objective, this study examined whether geographic and socio-demographic disparities existed in the occurrence of late-stage breast cancer

diagnosis and mortality rates at the census tract level across the state of Texas from 1995-2005. The number and locations of census tracts that exhibited significant racial disparities of breast cancer late-stage diagnosis were not consistent based on results from the rate difference (RD) and rate ratio (RR) statistics. Many more census tracts with significant racial disparities were identified using the RD statistic than RR in that the number of census tracts with missing values varied greatly for the two statistics due to their measurement characteristics. With the RD measurement, 188 census tracts were found to have significantly higher late-stage diagnosis for African-American in comparison with non-Hispanic whites, while only 9 census tracts were significant in RR. Most of these census tracts were located within the metropolitan areas of Houston, Dallas, and Austin-San Antonio, where African-American residents were concentrated. About 80% of the census tracts that were significant in racial disparities of late-stage diagnosis had more than 20% population living under the poverty line. Hispanics were found to have significantly higher late-stage diagnosis within 266 census tracts. Besides the metropolitan areas stated above, the census tracts with significant racial disparities included the area along the Southwest border of Texas for Hispanics. Among these census tracts, about 88% in these tracts had a low SES status because that more than 20% population lived under the poverty line. The strong connection of racial disparities in late-stage diagnosis and low SES status may reflect the obstacles to accessing limited health resources and lack of financial support for minorities, in more impoverished areas.

More census tracts were identified as significant in racial disparities for breast cancer mortality than for late-stage diagnosis. Using a RD measurement, African-Americans were found to have had 278 out of 4,388 census tracts (6.30%) that tested for

significantly higher breast cancer mortality rates and 328 census tracts (7.47%) for Hispanic women. No census tracts were significant using the RR statistic. Most of these census tracts with significant racial disparities were located in the metropolitan areas of Houston, Dallas, and Austin-San Antonio. For Hispanics, the Southwest border of Texas was also found to have a great number of census tracts with significant racial disparities in breast cancer mortality. Most of these identified census tracts had a low socioeconomic status of more than 10.00% population living under poverty line.

The second objective of the study was to compare the spatial patterns of racial disparities in late-stage diagnosis rates and breast cancer mortality rates. The map overlay illustrated that the census tracts that tested significant in racial disparities of both late-stage diagnosis and breast cancer mortality did not fully overlap. African-American women had 109 census tracts tested significant in racial disparities of both breast cancer mortality and late-stage diagnosis, while Hispanics had 130 census tracts tested significant in both racial disparities. The confirmatory data analysis of logistic regression models was utilized to assess the relationship between racial disparities and socioeconomic status. As a result, census tracts with the middle and low SES were 2.27 and 18.35 times more likely to report significant racial disparities in late-stage diagnosis among African-Americans, while the odds ratios were 1.86 and 10.19 for breast cancer mortality respectively. Within census tracts of middle and low SES level, Hispanics were found to be 3.24 and 46.89 times more likely to be associated with significantly higher rates of late-stage diagnosis than non-Hispanic white females. Additionally, these ratios were 2.51 and 24.29 for breast cancer mortality among Hispanics. The poverty level

impacted the significance of these two racial disparities to a much greater extent for Hispanics than for African-Americans.

The spatial relationship of the two racial disparities based RD and RR statistics was assessed using linear regression models. Using the normalized RD measurements for both racial disparities, the linear regression demonstrated that 23% of the variance in racial disparities of breast cancer mortality could be explained by the disparities of late-stage diagnosis for both African-Americans and Hispanics. On the other hand, more than 45% of the variance of racial disparities in breast cancer mortality was contributed by racial disparities in late-stage diagnosis if measured in RR. LISA statistics were employed to the residuals of linear regression and found a significant spatial autocorrelation within these residuals. Therefore, geographically-weighted regressions were conducted and had slightly improved R squares by taking the neighborhood effect into account. Furthermore, the logistic regression models found that if a census tract identified significant racial disparities in late-stage diagnosis, the census tract was 33.76 times more likely tested significant in racial disparities of breast cancer mortality for African-Americans and 30.39 times more likely for Hispanics. The strong association of these two disparities was found and varied across the state of Texas.

The third objective was to investigate other potential risk factors responsible for significant racial disparities in breast cancer mortality since the regression analyses had low R^2 when only racial disparities in late-stage diagnosis were considered as the independent variable. A factor analysis was performed for fourteen normalized predictors. The factor analysis revealed that three factors accounted for 75% variance underlying the dataset. These three factors are demographic characteristics,

socioeconomic status, and lack of spatial accessibility to mammography facilities. Subsequent logistic regression analysis found the significant contribution of these predictors. For instance, if a census tract is tested significant in racial disparities of late-stage diagnosis, then the census tract was four times more likely to be significant in racial disparities of breast cancer mortality. The spatial accessibility factor had the odds ratio less than one for African-American women, which indicated that close proximity to mammography facilities was not a determining factor in the significance of racial disparities in breast cancer mortality. The spatial accessibility factor was not a significant predictor in the logistic model established for Hispanics.

The socioeconomic factor was recognized as a critical predictor to determine if a census tract was significant in racial disparities of breast cancer mortality. For example, an impoverished census tract was 2.43 times and 5.30 times more likely to be significant in racial disparities of breast cancer mortality for African-Americans and for Hispanics respectively. Moreover, the socioeconomic factor plays an important role in determining if people have health insurance and financial support to afford the costs associated with cancer screening for an early detection and adequate treatments. This study pointed out that racial disparities in breast cancer were not merely driven by deficient supply of health care. Therefore, establishing more health care facilities in the poor areas with more than 20% population living under the poverty line may not be an effective way to reduce cancer disparities.

6.2 Intervention Programs

This research uncovered the geographic regions in Texas that exhibited significant racial disparities in late-stage diagnosis and breast cancer mortality among African-American and Hispanic women from 1995-2005. An effective intervention program should target needed subpopulations in the prioritized geographic regions. This study revealed the important role of socioeconomic factor and significance of racial disparities in late-stage breast cancer diagnosis, which determined if a census tract was significant in racial disparities of breast cancer mortality. As a result, low-income and uninsured women were frequently confronted with the multiple burdens of having limited access to early detection care and lower access to post-diagnosis treatments. Thus, the subsidized health insurances and free mammograms for disadvantaged African-Americans and Hispanics could be applied at local communities to reduce racial disparities in breast cancer. The long term goal to improve the African-American and Hispanic women's health is to boost their income and enhance their social status through higher educational attainment.

Community-based intervention programs proved to be a potentially effective approach to reduce health disparities and improve health standing of underserved populations. The campaign program called –Cancer Detection Programs: Every Woman Counts (CDP: EWC), has been successful in promoting mammography utilization. A recent study found that women who were not proficient in English reported more mammogram screening than women with better English proficiency in Los Angeles County (Meersman *et al.* 2009). The CDP: EWC program provided a subsidized screening and diagnostic service to qualified women in California and referred 62%

screened women among the age group of 40-64 for further evaluation (Bhaskara *et al.* 2008).

Program similar to CDP: EWC could be replicated in Texas to reach out underserved Hispanic women based on the culturally-sensitive model, particularly in the metropolitan areas and Southwest border identified in this study. Computerized tracking systems can be effective in reminding women due for their screening and follow-up checks (Ruffin *et al.* 2000). For some time we have known that the cancer health care faced by the Hispanic population can be improved through offering comprehensive financial service of health insurance coverage, increasing the level of access to health care, reinforcing health care delivery by shortening the waiting time, and boosting the social and education levels of minorities in the long run (Andersen *et al.* 1981). More recently, mobile mammography units have been developed, and these units could be deployed in the remote areas of the Southwest border of Texas to increase the accessibility of people in these areas to health care.

The Texas Cancer Council funded the project of African American Breast Cancer Outreach (AABCO) in 1998 (Adams *et al.* 2003). A community-based and culturally sensitive model was applied to increase the awareness of breast cancer and screening utilization among African American women from 1998-2003 (Israel *et al.* 1998; Adams 2007). The AABCO model can be introduced to implement into the three metropolitan areas with significant racial disparities in breast cancer mortality for African-Americans identified in this research. Health promotion activities can be planned on important holidays such as Mother's Day in churches, shopping malls, and beauty shops, partnering with local health organizations. Health education workshops can be organized by local

oncologists and general practitioners to improve awareness and attract higher rate of attendance from local communities. The barriers preventing women from screening can be ascertained through a well-designed survey, which can help address women's health concerns and improving the effectiveness intervention activities. A focus group with breast cancer patients and their family members steered by professional nurses can be a great way to raise the importance of mammography screening and breast health among women (Lee 2000).

6.3 Contributions

Although a lot of efforts have been devoted to research in health disparities, the mechanisms responsible for widening gaps of racial disparities in breast cancer are still not clear. This is partly due to the complexity in the interactions among risk factors affecting breast cancer outcomes (Lannin *et al.* 2002). This study exemplifies how GIS techniques could provide an innovative platform to investigate racial and geographic disparities of breast cancer. The insights from this dissertation could be applied to the targeted areas for more effective intervention programs to reduce these disparities. This study has made the following two overarching contributions towards untangling the complex relationships between risk factors affecting breast cancer and widening disparities.

First, the study focused on a larger study area at fine spatial aggregation levels. Most racial disparities research concentrated on smaller regions such as a few counties and SEER regions to combine all individual cases into different racial groups. The current study examined racial disparities in breast cancer for the whole state of Texas.

This research is novel as it investigated how racial disparities in breast cancer vary across geographic regions. Geographic disparities research can help investigate racial disparities in health rising from adverse environmental influence and different levels of access to health care services under different regions.

Second, this study addressed the fact that few studies have examined the direct connection of racial disparities between late-stage diagnosis and breast cancer mortality across geographic regions. This study applied regression models to the normalized statistics of racial disparities and investigated the association of both racial disparities using linear and logistic regressions. Most studies to date have utilized either single or a composite socioeconomic index in a categorical format to assess how socioeconomic status impact racial disparities of health, whereas this study produced a set of more comprehensive and numeric factors to represent socioeconomic status. Measurements of spatial accessibility were improved using the more accurate network-based distance and travel time instead of Euclidean straight-line distance. This study not only identified the census tracts in need of immediate attention to improve minority breast cancer health, but also pointed out some vital factors responsible for these significant racial disparities in breast cancer. This research discussed some intervention strategies to provide subsidized health insurance, financial support, and free mammogram utilization for underserved population among African-American and Hispanic females in needed communities.

6.4 Limitations and Future Work

The primary limitations in this study are elaborated as follows for the purpose of pointing out some future research directions. First, the modifiable areal unit problem

(MAUP) is a great concern for epidemiology studies. MAUP could lead to different statistical results based on the aggregation of the same individual-level data into different spatial units and zoning boundaries (Openshaw 1984). For example, the author explored racial disparities of breast cancer mortality at the county, zip code, and census tract levels before conducting this research and found the census tract level offered a more complete and better understanding of racial disparities in breast cancer mortality.

The second limitation is related to data quality and data processing. Incomplete geocoding of cancer data may result in the change of both racial disparities in this study. About 0.42% of breast cancer incidences were not mapped as a result of incompleteness/absence of an address, errors in an address, or a mailing address in the form of a post office box. About 12,480 cancer cases (8.34%) had missing information about the stage of cancer at diagnosis, which could impact the significance test of racial disparities in late-stage diagnosis across space. With respect to breast cancer mortality, the addresses of 13% of non-Hispanic whites, 10% of African-Americans, and Hispanic cases could not to be geocoded. The difference in geocoding rates among racial groups may have influenced the significance test of racial disparities in breast cancer mortality. Pertaining to data processing, age-adjusted incidence and mortality rates were calculated for each race using the 2000 US census population, but breast cancer cases were aggregated in the years from 1995 to 2005. Using the 2000 population to represent the exposed population in the period may cause an artificial inflation of incidence and mortality rates. Moreover, the increase in population varies among different racial groups. The small sample number issue may take place from calculating the cancer rates at such small geographical level of census tract. Statistical smoothing algorithms can be

applied to reduce the uncertainty within local small-area variations by considering neighborhood effects (Kafadar 1994). However, different algorithms have distinct underlying assumptions of spatial disease patterns and the difficulties in quantifying the uncertainties of smoothing rates and non-stationarity in datasets (Goovaerts 2005).

Third, lack of information on individual socioeconomic status and early screening detection place a significant barrier to investigate racial disparities in female breast cancer. This study assumed that population lived in the centroid of census tract to measure spatial accessibility to mammography facilities. In addition, the actual mammography utilization could be subject to a number of complicated factors including health insurance, financial obstacles, and physician recommendations.

Thus, future research on female cancer disparities could be extended to the study of mammography utilization difference across race, differences in health insurance, and health care provider variability across regions. Additional qualitative information could be gleaned through well-designed surveys that identify other psycho-social risk factors underlying cancer disparities such as the beliefs about the possibility of developing breast cancer, and the perceived importance of mammogram prevention and fear of finding breast cancer (Adams 2007; Russell and Shedd-Steele 2003).

Frequently combining cancer cases in broad categories such as Hispanic origin may obfuscate different risk factors underlying cancer disparities among subpopulations. Future research can be expanded to investigate racial disparities from a range of breast cancer outcomes among subpopulation groups. The study of carcinoma characteristics such as estrogen and progesterone receptors may shed an insight in respect to biological breast cancer treatment on different racial groups. The environmental inequality can be

placed as another emphasis to exploit the adverse effects harmful environmental exposure from their contaminated resident areas, especially for minority and disadvantaged groups (Abelsohn *et al.* 2002). The gene-environment interaction research can open a pioneering arena for cancer disparities by helping better understand the biological pathways of cancer disease caused by the physical and socio-culture environments (Wilson *et al.* 2002). This kind of research can aid in answering the questions such as how socioeconomic status determine people's health not only from affordability of health care utilization, but also the essential biological change at the molecule level through modifying life-style, human behaviors, and diary consumption in a long term (Eertmans *et al.* 2001).

In conclusion, this research helps identify the specific areas where immediate attention should be placed to reduce racial disparities in breast cancer outcomes in Texas. The socioeconomic status plays an important role in shaping individual health through neighborhood effects. New mammography facilities and health care clinics could be placed within impoverished and remote areas where minority and disadvantaged residents are more concentrated. Culturally sensitive community-based models could be implemented at targeted geographic areas to reduce the rates of late-stage diagnosis and improve the survival rates of breast cancer among disadvantaged groups including African-American and Hispanic women.

APPENDIX A. SUMMARY OF SELECTED PUBLICATIONS

The table below lists some selected publications and the findings presented in these publications

Study	Purpose	Study design	Findings	Future work
Bickell et al. (2009)	Evaluate a tracking and feedback registry effect on oncology consultations and use of adjuvant treatment.	Cohort study	Implementation of the tracking and feedback registry improved radiation oncology consultations and use of adjuvant treatment.	A tracking and feedback registry program could be an effective approach to reduce racial disparities.
Menashe et al. (2009)	Explore the racial disparities in breast cancer mortality among African-American women.	Population-based study	Disparities in mortality between blacks and whites have been widened, mainly driven by the higher death rates among blacks.	The etiology of the excess burden and therapeutic strategies should be placed on research emphasis.
Vona-Davis and Rose (2009)	Summarize the studies regarding the influence of socioeconomic on breast cancer tumor biology and prognosis.	Review	Poorer survival for Hispanic women was accounted for the socioeconomic disadvantage and later stage at diagnosis. African-American women had more aggressive tumors.	Future research should take an effort on the influence of social deprivation on breast cancer tumor characteristics.

Kim et al. (2008)	Investigate the reason that black women have poor survival.	Cohort: 1999-2006	Higher proportion of failure to complete adjuvant therapy among black women indicates that improving treatment follow-up will help increase survival for blacks.	Future research should aim at seeing if it is generalizable and what contribute blacks to fail complete adjuvant therapy.
Meliker et al. (2008)	Assess the survival disparities among African-American at different geographic levels.	Methodological study	A novel statistic method by accounting for population size was proposed to quantify the survival differences among African-American. The inconsistency occurred for the survival disparities at different geographical scales.	Calls more research to identify innate and modifiable factors responsible for the observed racial disparities.
Shack et al. (2008)	Analyze socioeconomic variations by region and age in England.	Cohort study: English cancer registries between 1998 and 2003.	The highest incidence of breast cancer occurred among the least deprived groups. Slight variation by region was observed in breast cancer incidence.	Data collection and timeliness may contribute the regional variation. Completeness of registration is less for deprived groups. Ecological measures of SES may not represent individuals.
Tarlov et al. (2008)	Investigate the association of breast cancer stage with characteristics of mammography facility locations.	Cohort study: 1996-1998.	Homicides neighbor effects were associated with stage of breast cancer. No association was observed for neighborhood characteristics besides crime in the neighborhood.	Future efforts may advance research related to factors of perceived threats effect on mammography use. Research needed to disentangle the effects of facility areas from neighborhood characteristics.
Vainshtein (2008)	Compare breast cancer incidence at	Review	Positive correlation was found between SES and breast cancer	The standard measurement of SES is absent from the

	different SES across races.		incidence. Racial disparities in breast cancer for Hispanics and Asian narrowed down in higher SES group.	current literature.
Brody et al. (2007)	Identify gaps of breast cancer research related to disparities.	Review	Disparities regarding to breast cancer were reported ranging from incidence, hormonal and lifestyle factors, genetic/biologic factors.	It is an overview of breast cancer research related to disparities and environmental issues.
Chu et al. (2007)	Examine the trend of disparity change in cancer mortality by race and socioeconomic.	Population-based study: 1990-2000.	Racial disparities increased the most for minorities at the lowest SES group.	Much work still need including early detection and treatment intervention due to the disparity increase.
Gerend and Pai (2007)	Comprehensively review social, economic, and cultural factors associated with mortality disparities.	Review: 1980-2006	The authors addressed social, economic and cultural factors associated with mortality disparities in breast cancer between African-American and whites.	Future research will benefit from study of the pathway of these determinants effects on health disparities. Various socioeconomic factors should be assessed on health disparities.
Gomez et al. (2007)	Identify the underuse of mammography screening among the Asian subpopulation.	Cohort study: 2001 California Health Interview Survey	Institutional, cultural and linguistic factors were important barriers for Asian women not following mammography use recommendation in addition to SES.	Effort should focus on culturally- and linguistically-appropriate strategies to implement intervention program.
Owusu et al. (2007)	Assess the association of survival with treatment receipt.	Cohort study: 1997-1999	Women older than 75 were less likely to receive guideline therapy which caused decrease in breast	Research need to emphasize the randomized clinical trial among the older population.

			cancer-specific survival.	
Wells and Roetzheim (2007)	Evaluate if Latinas received less mammography screening and what factors could explain it.	Review	Disparities in mammography screening were more pronounced in community samples and elderly Hispanic women. Lack of health insurance and lower SES were responsible for the lower screening-taken.	Predictor of mammography use including employment, nativity, and psychosocial variables should be assessed in future.
Carey et al. (2006)	The prevalence of subtype of breast cancer identified.	Cohort: Carolina Breast Cancer Study; 1993-1996.	Pre-menopausal African-American women are more likely to have Basal-like tumors and less likely to have luminal A tumor.	Future research needed to confirm that young African-American have higher prevalence of basal-like tumors.
Georgia et al. (2006)	Explore disparities in treatment and decision-making in minority women.	Review	Providing sufficient knowledge and appreciating the social and familial emotional context should be considered in helping patients making proper treatment options.	Research should explore how decision-making affect breast cancer diagnosis in the health-care delivery system.
Grann et al. (2006)	Evaluate the correlation of survival with geographic regions.	Cohort study: 1990-2001	Detroit has significantly higher mortality rates in the 11 SEER regions. Hawaii has the best survival. Overall, black had the worse survival.	Linking SEER database with individual socioeconomic and insurance data should become policy priority.
Sarfati et al. (2006)	Test hypothesis that socioeconomic and racial trends varies in New Zealand.	Population-based study: 1981-1999.	Due to mortality increase among Maori women, racial disparities widened. Higher income and education groups had significantly decrease of breast cancer mortality.	Studies regarding to stage differences are needed among Maori and pacific people related to non-Maori non-Pacific women.
Field et al.	Determine if the survival	Retrospective	African-American still experienced survival	More research should be conducted on how

(2005)	disparities were persistent among African-American.	cohort: 1993-1998	disparities despite access to health and having insurance.	socioeconomic factor and patient's interactions with health care system influence their receipt of health care.
Bobo et al. (2004)	Report the characteristics of on-schedule mammography rescreening	Prospective Cohort: interview	Race, move since index mammogram, reminders from physician, social support improved the reschedule screening rate on time.	Research needed to determine who belongs to the 20% not adheres to the rescreening schedule.
Glazier et al. (2004)	Reveal geographic patterns of mammography use in Toronto.	Cohort Study: 2000	Modifiable areal unit problem was investigated. Mammography intake had a strong positive association with income and immigration status.	The study may be limited by data coding error and accuracies. Area characteristics from 1996 Canadian Census were applied to women in 2000.
Gwyn et al. (2004)	Examine racial difference in the interval of medical consultation, diagnosis, and treatment among African-American.	Cohort study: 1990-1992	Access to care and lower socioeconomic may contribute to the observed delay in diagnosis and treatment among African-American women.	Additional research need to focus on other contributors to which population have the medical care delay.
Hsu et al. (2004)	Evaluate the geographical cluster among racial group in Texas	Population-based study: 1990-1996	The most likely cluster occurred along Gulf coast and Central Texas for non-Hispanic whites. Hispanics had a relative risk of 18% with excessive mortality burden in the West of Texas.	The detected cluster in Texas warrants the further research.
Sturgeon et al. (2004)	Illustrate patterns and trends of mortality by region, age	Population-based study: 1950-	Favorable historical pattern of mortality has been diminished in 1990s. African-American women	A continued need is required to monitor breast cancer continuum across region in order to

	and race group.	1999.	experienced the most disparities in South.	understand why the favorable trend erodes in South.
Chu et al. (2003)	Determine the treatment difference by examining survival rates.	Population-based study	Younger black women experienced racial differences in treatment for both ER+ and ER-. No survival disparities for older black suggest that Medicare may alleviate the situation.	The study is limited that there was no biological difference in the aggressiveness of breast cancer among racial groups.
Li et al. (2003)	Investigate breast cancer treatment, survival and mortality across races.	Retrospective cohort study: 1992-1998	Minority groups experienced more advanced and poor prognostic breast cancer tumors.	The study is limited by possible race misclassification, incomplete of treatment information, and lack of other risk factors.
Bradley et al. (2002)	Disentangle the influence of socio-demographic factors on the stage, treatment and survival.	Retrospective cohort: 1996-1997	After controlling for confounders, lower socioeconomic status was a predictor for unfavorable breast cancer outcome instead of race.	The study has limitations of underreporting women having radiation therapy, generability due to constrain to urban areas, ecologic poverty measurement, misclassifying data.
Coughlin et al. (2002)	Examine the disparities of rural and urban in mammography screening.	Cohort study: 1997-1998.	Hispanic women had relatively low screening rates. Women living in rural areas underutilized the preventive screening.	Classification of residence areas as rural and urban areas was problematic. The study may not be generalizable.
Legler et al. (2002)	Estimate mammography use across states based on county socio-demographic data.	Methodology study	Positive association was found for state variation of mammography use with number of intervention research.	Tailored research is encouraged to target the population in locations most needed.
Chu et al. (2001)	Four joints of estrogen and progesterone	Cohort study: SEER	Each racial group has the trend of ER+PR+>ER-PR-	The ER and PR data were reported from different laboratories.

	receptors were reported among multiple racial groups.	cancer registries between 1992 and 1997.	>ER+PR->ER-PR+. ER determines the subtypes of breast cancer tumors while PR modifies the subgroup.	The sample sizes were small for minority groups. Frequency of tumors was used instead of rates.
Richardson et al. (2001)	Examine the treatment difference between the women diagnosed by government program and their counterpart in the area.	Cohort study: 1992-1995	There were no statistical differences found among low-income women comparing with the women diagnosed with breast cancer in those areas.	Health education efforts should ensure that women receive the proper and complete treatment.
Caplan et al. (2000)	Determine the interval variation between abnormal mammography screening, breast cancer diagnostic and treatment.	Cohort Study: 1991-1995.	Women with abnormal mammograms, self-reported breast symptoms has shorter diagnostic interval than normal one. White has shorter intervals of diagnosis and treatment than other racial groups.	Further investigation needed to identify the factors associated with longer intervals between diagnosis and treatment.
Hall et al. (2000)	Investigate the association of body size and shape with risk of breast cancer.	Case-control study	Waist/hip ratio was positively associated with risk of developing breast cancer among all postmenopausal women. Premenopausal white women were observed to have an inverse association of body mass with risk of breast cancer.	Future research effort may be directed at better classification of fat distribution and inclusion of confounding factors of SES, physical activity, and diet.
May et al. (2000)	Address breast cancer rates using mammography screening data.	Cohort study: 1993-1996	American Indians/Alaska native had the highest detected breast cancer derived from mammography	The study may not be generalizable to the whole population due to using the data from the National Breast and Cervical Cancer

			screening. Mammography screening variation may account for some of the observed difference of breast cancer incidence.	Early Detection Program (NBCCEDP).
Moorman et al. (1999)	Investigate the association of obesity with stage of breast cancer among black and white.	Cohort: interview	Anthropometric factors, SES, and medical factors explain substantially differences of stage of breast cancer between black and white.	Information on mammography screening, education used as a measure of SES, no data regarding access to health care or cultural beliefs limited this study.
Lannin et al. (1998)	Evaluate SES and cultural factors on breast cancer stage.	Case-control study: interview; 1985-1996	SES and cultural belief together largely accounted for the observed stage difference among African-American women.	Generalization of the study could be applied.
Marchand (1991)	Examine the survival patterns among Japanese and Caucasians.	Review	The literature supports that weak negative association exist between obesity and survival.	Lower fat intake in Japanese which may explain the better survival for Japanese deserve the further study.

REFERENCES

- Abelsohn, A., B. L. Gibson, M. D. Sanborn, and E. Weir. 2002. Identifying and managing adverse environmental health effects: 5. persistent organic pollutants. *Canadian Medical Association Journal* 166 (12): 1549-1554.
- Adegoke, O. J. 2003. *Gene-gene and gene-environment interactions in the etiology of breast cancer*. Reported by U.S. Army Medical Research and Material Command Fort Detrick, Maryland (21702-5012).
- Adler, N. E., and J. M. Ostrove. 1999. Socioeconomic status and health: What we know and what we don't. In *Socioeconomic Status and Health in Industrial nations: social, psychological, and biological pathways*. *Annals of the New York Academy of Sciences* 896(1): 3-15.
- Adler, N. E., and D. H. Rehkopf. 2008. US disparities in health: Descriptions, causes, and mechanisms. *Annual Review of Public Health* 29: 235-252.
- Admas, M. L., H. Becker, J. J. Guidry, T. R. Robertson, C. A. Carrington, M. Erard, and K. A. Monsho. 2003. *A guide to enhance African American Women's participation in breast cancer screening outreach and case management services*. Austin, TX: the University of Texas at Austin School of Nursing.
- Adams, M. L. 2007. The African American breast cancer outreach project: Partnering with communities. *Family & Community Health* 30: S85.
- Adams, S. A., J. R. Hebert, S. Bolick-Aldrich, V. G. Daguise, C. M. Mosley and M. V. Modayil. 2006. Breast cancer disparities in South Carolina: early detection, special programs, and descriptive epidemiology. *Journal of South Carolina Medical Association* 102 (7): 231-239.
- Ahuja, N., D. Chang, and S. L. Gearhart. 2007. Disparities in colon cancer presentation and in-hospital mortality in Maryland: A ten-year review. *Annals of Surgical Oncology* 14(2): 411-416.

- Aldrich, J. 1997. R.A. Fisher and the making of maximum likelihood 1912–1922. *Statistical Science* 12(3): 162–176.
- Althuis, M. D., J. M. Dozier, W. F. Anderson, S. S. Devesa, and L. A. Brinton. 2005. Global trends in breast cancer incidence and mortality 1973-1997. *International Journal of Epidemiology* 34(2): 405-412.
- Alexander, D. D., J. Waterbor, T. Hughes, E. Funkhouser, W. Grizzle, and U. Manne. 2007. African-American and Caucasian disparities in colorectal cancer mortality and survival by data source: An epidemiologic review. *Cancer Biomarkers* 3(6): 301-313.
- American Cancer Society. 2008a. *Cancer Facts and Figures 2008*. Atlanta, GA: American Cancer Society.
- American Cancer Society. 2008b. *Texas Facts & Figures 2008*. Austin, TX: American Cancer Society, High Plains Division.
- American Cancer Society. 2004. *Cancer facts & figures 2004*. Atlanta: American Cancer Society.
- Andersen, R., S. Z. Lewis, A. L. Giachello, L. A. Aday, and G. Chiu. 1981. Access to medical care among the Hispanic population of the southwestern United States. *Journal of Health and Social Behavior* 22(1): 78-89.
- Anselin, L. 1995. Local indicator of spatial association-LISA. *Geographical Analysis* 27: 93-115.
- Arcury, T. A., W. M. Gesler, J. S. Preisser, J. Sherman, J. Spencer, and J. Perin. 2005. The effects of geography and spatial behavior on health care utilization among the residents of a rural region. *Health Services Research* 40(1): 135.
- Athas, W. F., M. Adams-Cameron, W. C. Hunt, A. Amir-Fazli, and C. R. Key. 2000. Travel distance to radiation therapy and receipt of radiotherapy following breast-conserving surgery. *Journal of the National Cancer Institute* 92(3): 269-271.
- Australian Institute of Health and Welfare and Australasian Association of Cancer Registries (AACR). 2003. *Cancer survival in Australia 1992-1997: geographic categories and socioeconomic status*. AIHW Cancer Series 22.

- Avruskin, G. A., G. M., Jacquez, J. R., Meliker, M. J., Slotnick, A. M., Kaufmann, and J. O., Nriagu. (2004). Visualization and exploratory analysis of epidemiologic data using a novel space time information system. *International Journal of Health Geographics* 3: 26.
- Ayanian, J. Z., B. A. Kohler, T. Abe, and A. M. Epstein. 1993. The relation between health insurance coverage and clinical outcomes among women with breast cancer. *New England Journal of Medicine* 329(5): 326-331.
- Baade, P. D., L. Fritschi, and J. F. Aitken. 2005. Geographical differentials in cancer incidence and survival in Queensland: 1996 to 2002. Brisbane, Queensland: Viertel Centre for Research in Cancer Control, Queensland Cancer Fund.
- Bach, P. B., L. D. Cramer, J. L. Warren, and C. B. Begg. 1999. Racial differences in the treatment of early-stage lung cancer. *New England Journal of Medicine* 34(16): 1198-1205.
- Baicker, K., A. Chandra, J. S. Skinner, and J. E. Wennberg. 2004. Who you are and where you live: How race and geography affect the treatment of Medicare beneficiaries. *Health Affairs* (Millwood) 23:VAR33.
- Baldwin, L. M., S. H. Taplin, H. Friedman, and R. Moe. 2004. Access to multidisciplinary cancer care. *Cancer* 100(4): 701-709.
- Ballard-Barbash, R., A. L. Potosky, L. C. Harlan, S. G. Nayfield, and L. G. Kessler. 1996. Factors associated with surgical and radiation therapy for early stage breast cancer in older women. *Journal of National Cancer Institute* 88(11): 716-726.
- Ballard-Barbash, R., S. H. Taplin, B. C. Yankaskas, V. L. Ernster, R. D. Rosenberg, P. A. Carney, W. E. Barlow, B. M. Geller, K. Kerlikowske, and B. K. Edwards. 1996. Breast cancer surveillance consortium: A national mammography screening and outcomes database. *American Journal of Roentgenology* 169 (4): 1001-1008.
- Baquet, C. R., S. I. Mishra, P. Commiskey, G. L. Ellison, and M. DeShields. 2008. Breast cancer epidemiology in blacks and whites: disparities in incidence, mortality, survival rates and histology. *Journal of National Medical Association* 100 (5): 480-488.
- Barnett, E., V. E. Braham, M. L. Casper, G. A. Elmes, and J. A. Halverson. 2001. *Men and heart disease: An atlas of racial and ethnic disparities in mortality*. 1st Ed. Atlanta, GA: Centers for Disease Control and Prevention.

- Bartlett, M.S. 1947. The use of transformations. *Biometrics* 3 (1): 39-52.
- Benjamini Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* 57(1): 289-300.
- Bentley, J. R., R. J. Delfino, T. H. Taylor, S. Howe, and H. Anton-Culver. 1998. Differences in breast cancer stage at diagnosis between non-Hispanic white and Hispanic populations, San Diego county 1988–1993. *Breast Cancer Research and Treatment* 50(1): 1-9.
- Bhaskara, A., M. Altamirano, V. Trisal, I. B. Paz, and L. L. Lai. 2008. Effectiveness of decentralized community-based screening, detection, and treatment of breast cancer in low-income, uninsured women. *The American Surgeon* 74(10): 1017-1021.
- Bickell, N. A., K. Shastri, K. Fei, S. Oluwole, H. Godfrey, K. Hiotis, A. Srinivasan, and A. A. Guth. 2008. A tracking and feedback registry to reduce racial disparities in breast cancer care. *Journal of the National Cancer Institute* 100(23): 1717-1723.
- Blot W. J., and J. R. Fraumeni. 1996. Cancers of the lung and pleura. pp. 637-665 in Schottenfeld D, Faumeni J.F.Jr (eds), *Cancer Epidemiology and Prevention*. Oxford University Press, New York.
- Black, D., J. Morris, C. Smith, and P. Townsend. 1980. *Inequalities in health: Report of a research working group*. London: Department of Health and Social Security.
- Blackman, D. J., and C. M., Masi. 2006. Racial and ethnic disparities in breast cancer mortality: are we doing enough to address the root causes? *Journal of Clinical Oncology* 24 (14): 2170-2178.
- Bobo, J. K., J. A. Shapiro, J. Schulman, and C. L. Wolters. 2004. On-schedule mammography rescreening in the national breast and cervical cancer early detection program. *Cancer Epidemiology Biomarkers & Prevention* 13(4): 620-630.
- Boyer-Chammard, A., T.H.Taylor, and H. Anton-Culver. 1999. Survival differences in breast cancer among racial/ethnic groups: a population-based study. *Cancer Detect Preview* 23: 463-473.

- Brach, C., and I. Fraser. 2000. Can cultural competency reduce racial and ethnic health disparities? A review and conceptual model. *Medical Care Research and Review* 57: 181-217.
- Bradley, C. J., C. W. Given, and C. Roberts. 2002. Race, socioeconomic status, and breast cancer treatment and survival. *Journal of National Cancer Institute* 94(7): 490-496.
- Brawley, O. W., and H. P. Freeman. 1999. Race and outcomes: Is this the end of the beginning for minority health research? *Journal of National Cancer Institute* 91(22): 1908-1909.
- Breen, N., K. A. Cronin, H. I. Meissner, S. H. Taplin, F. K. Tangka, J. A. Tiro, and T. S. McNeel. 2007. Reported drop in mammography: Is this cause for concern? *Cancer* 109 (12): 2405-2409.
- Brody, J. G., M. H. E. Kavanaugh-Lynch, O. I. Olopade, S. M. Shinagawa, S. Steingraber, D. V. Scholar, and D. R. Williams. *Identifying gaps in breast cancer research: Addressing disparities and the roles of the physical and social environment*. California Breast Cancer Research Program Special Research Initiatives.
- Brondolo, E, L. C. Gallo, and H. F. Myers. 2009. Race, racism and health: Disparities, mechanisms, and interventions. *Journal of Behavioral Medicine* 32(1): 1-8.
- Brown P. 1995. Race, class and environmental health: a review and systemization of the literature. *Environment Research* 69: 15–30.
- Brunsdon, C., A. S. Fotheringham, and M.E. Charlton. 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* 28: 281–98.
- Byers, T., J. Mouchawar, J. Marks, B. Cady, N. Lins, G. M. Swanson, D. G. Bal, and H. Eyre. 1999. The American cancer society challenge goals: How far can cancer rates decline in the U. S. by the year 2015? *Cancer* 86 (4): 715-727.
- Cambridge Dictionaries Online. 2001. Cambridge international dictionary of English. Online. Available from Internet: <http://www.dictionary.cambridge.org/>, accessed, November 3, 2009.

- Campbell, N. C., E. Murray, J. Darbyshire, J. Emery, A. Farmer, F. Griffiths, B. Guthrie, H. Lester, P. Wilson, and A. L. Kinmonth. 2007. Designing and evaluating complex interventions to improve health care. *British Medical Journal* 334(7591): 455.
- Caplan, L. S., D. S. May, and L. C. Richardson. 2000. Time to diagnosis and treatment of breast cancer: Results from the national breast and cervical cancer early detection program, 1991-1995. *American Journal of Public Health* 90(1): 130-134.
- Carey, L. A., C. M. Perou, C. A. Livasy, L. G. Dressler, D. Cowan, K. Conway, G. Karaca, M. A. Troester, C. K. Tse, and S. Edmiston. 2006. Race, breast cancer subtypes, and survival in the Carolina breast cancer study. *The Journal of American Medical Association* 295(21): 2492.
- Carter-Pokras, O., and C. Baquet. 2002. What is a "health disparity"? *Public Health Reports* 117(5): 426-434.
- Casper, M. L., E. Barnett, J. A. Halverson, G. A. Elmes, V.E. Braham, Z. A. Majeed, A. S. Bloom, and S. Stanley. 1999. *Women and heart disease: An atlas of racial and ethnic disparities in mortality*. Office for Social Environment and Health Research, West Virginia University, Morgantown, WV.
- Cassel, J. 1976. The contribution of the social environment to host resistance. *American Journal of Epidemiology* 104: 107-123.
- Cella, D. F., E. J. Orav, A. B. Kornblith, J. C. Holland, P. M Silberfarb, K. W Lee, R. L. Comis, M. Perry, R. Cooper, and L.H. Maurer. 1991. Socioeconomic status and cancer survival. *Journal of Clinical Oncology* 9(8): 1500-1509.
- Census Bureau. 2009. Online. Available from Internet: <http://www.census.gov/hhes/www/poverty/povdef.html>, accessed, March 3, 2010.
- Centers for Disease Control and Prevention. 2007. Use of mammography among women aged \geq 40 years-united states. *MMWR Morb Mortal Wkly Rep* 56: 49-51.
- Center to Reduce Cancer Health Disparities, National Cancer Institute. 2004. Online. Available from Internet: <http://crchd.nci.nih.gov/>, accessed, October 23, 2009.
- Chagpar, A. B., and K. M. McMasters. 2007. Trends in mammography and clinical breast examination: A population-based study. *Journal of Surgical Research* 140(2): 214-219.

- Chapa, J., and B. De La Rosa. 2004. Latino population growth, socioeconomic and demographic characteristics, and implications for educational attainment. *Education and Urban Society* 36(2): 130.
- Chapa, J., and R. R. Valencia. 1993. Latino population growth, demographic characteristics, and educational stagnation: An examination of recent trends. *Hispanic Journal of Behavioral Sciences* 15(2): 165.
- Chapman, K. 1979. *People, pattern and process: An introduction to human geography*. John Wiley & Sons.
- Chandra, A., and J. S. Skinner. 2003. *Geography and racial health disparities*. National Bureau of Economic Research, Cambridge, Mass.
- Chevarley F., and E. White. 1997. Recent trends in breast cancer mortality among white and black US women. *American Journal Public Health* 87: 775-781.
- Chin, M. H., A. E. Walters, S. C. Cook, and E. S. Huang. 2007. Interventions to reduce racial and ethnic disparities in health care. *Medical Care Research and Review* 64(5): 7S.
- Chlebowski, R. T., S. L. Hendrix, R. D. Langer, M. L. Stefanick, M. Gass, D. Lane, R. J. Rodabough, M. A. Gilligan, M. G. Cyr, and C. A. Thomson. 2003. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: The women's health initiative randomized trial. *Obstetrical & Gynecological Survey* 58(11): 735.
- Chu, K. C., B. A. Miller, and S. A. Springfield. 2007. Measures of racial/ethnic health disparities in cancer mortality rates and the influence of socioeconomic status. *Journal of the National Medical Association* 99(10): 1092-1010.
- Chu, K. C., C. A. Lamar, and H. P. Freeman. 2003. Racial disparities in breast carcinoma survival rates. *Cancer* 97(11): 2853-2860.
- Chu, K. C., R. E. Tarone, L. G. Kessler, L. A. Ries, B. F. Hankey, B. A. Miller, and B. K. Edwards. 1996. Recent trends in U.S. breast cancer incidence, survival, and mortality rates. *Journal of the National Cancer Institute* 88(21): 1571-1579.
- Chu, K. C., R. E. Tarone, and O. W. Brawley. 1999. Breast cancer trends of black women compared with white women. *Archives of Family Medicine* 8(6): 521.

- Chu, K. C., W. F. Anderson, A. Fritz, L. A. Ries, and O. W. Brawley. 2001. Frequency distributions of breast cancer characteristics classified by estrogen receptor and progesterone receptor status for eight racial/ethnic groups. *Cancer* 92(1): 37-45.
- Clarke, K. C., S. L. McLafferty, and B. J. Tempalski. 1996. On epidemiology and geographic information systems: A review and discussion of future directions. *Emerging Infectious Diseases* 2(2): 85-92.
- Clarke, M., R. Collins, C. Davies, J. Godwin, R. Gray, R. Peto, and S. Ebctcg. 1998. Tamoxifen for early breast cancer: An overview of the randomized trials. *Lancet* 351 (9114): 1451-1467.
- Claus, E. B., J. M. Schildkraut, W. D. Thompson, and N. J. Risch. 1996. The genetic attributable risk of breast and ovarian cancer. *Cancer* 77(11): 2318-2324.
- Coleman, W. 1982. *Death is a social disease: Public health and political economy in early industrial*. France University of Wisconsin Press Madison.
- Cook, W. K. 2008. Integrating research and action: A systematic review of community-based participatory research to address health disparities in environmental and occupational health in the USA. *Journal of Epidemiology and Community Health* 62(8): 668-676.
- Coughlin, S. S., T. D. Thompson, H. I. Hall, P. Logan, and R. J. Uhler. 2002. Breast and cervical carcinoma screening practices among women in rural and nonrural areas of the United States, 1998-1999. *Cancer* 94(11): 2801-2812.
- Davis, A. M., L. M. Vinci, T. M. Okwuosa, A. R. Chase, and E. S. Huang. 2007. Cardiovascular health disparities - A systematic review of health care interventions. *Medical Care Research and Review* 64(5): 29S-100S.
- Deaton, A. 2001. Health, inequality, and economic development. *Journal of Economic Literature* 41(1): 112-158.
- DeChello, L. M., and T. J. Sheehan. 2007. Spatial analysis of colorectal cancer incidence and proportion of late-stage in Massachusetts residents: 1995-1998. *International Journal of Health Geographics* 6: 20.
- Devesa, S., D. Grauman, W. Blot, G. Pennello, R. Hoover, and J. Fraumeni. 1999. Atlas of cancer mortality in the United States, 1950-94. *American Journal of Epidemiology* 149(11): S68.

- Devore, J. L. 2003. *Probability and Statistics for Engineering and the Sciences*. Duxbury Press, Pacific Grove, CA.
- Diez-Roux, A. V. 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health* 88(2): 216-222.
- Du, X. L., T. E. Meyer, and L. Franzini. 2007. Meta-analysis of racial disparities in survival in association with socioeconomic status among men and women with colon cancer. *Cancer* 109(11): 2161-2170.
- Duncan, G. J., and J. Brooks-Gunn. 1997. *Consequences of growing up poor*. Russell Sage Foundation.
- Eastern Research Group. 2001. Assessment of the availability of mammography services. Prepared for the Office of Policy, Planning and Legislation, Office of the Commissioner, Food and Drug Administration. Washington, DC: U.S. Food and Drug Administration, 2001. Online. Available from Internet: <http://www.fda.gov/downloads/RadiationEmittingProducts/MammographyQualityStandardsActandProgram/Reports/ucm126269.pdf>, accessed, September 15, 2009.
- Eertmans, A., F. Baeyens, and O. Van den Bergh. 2001. Food likes and their relative importance in human eating behavior: Review and preliminary suggestions for health promotion. *Health Education Research* 16 (4): 443-456.
- Eggleston, K. S., A. L. Coker, M. Williams, G. Tortolero-Luna, J. B. Martin, and S. R. Tortolero. 2006. Cervical cancer survival by socioeconomic status, race/ethnicity, and place of residence in Texas, 1995-2001. *Journal of Women's Health* 15(8): 941-951.
- Elledge, R. M., and S. A.W Fuqua. 2000. Estrogen and progesterone receptors. pp. 471-488 in J. Harris, M.E. Lippman, M. Morrow, C. Osborne (eds), *Diseases of the Breast*. Lippincott, Williams and Wilkins, Philadelphia.
- Elliott, P., and D. Wartenberg. 2004. Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives* 112(9): 998.
- Elster, A., J. Jarosik, J. VanGeest, and M. Fleming. 2003. Racial and ethnic disparities in health care for adolescents - A systematic review of the literature. *Archives of Pediatrics & Adolescent Medicine* 157(9): 867-874.

- Farley, T.A., and J.T. Flannery. 1989. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. *American Journal of Public Health* 79 (11): 1508-1512.
- Farrow, D. C., J. M. Samet, and W. C. Hunt. 1996. Regional variation in survival following the diagnosis of cancer. *Journal of Clinical Epidemiology* 49(8): 843-847.
- Federal, Provincial and Territorial Advisory Committee on Population Health. 1999. Toward a healthy future: second report on the health of Canadians. Ministry of Public Works and Government Services, Ontario, Canada.
- Feigelson, H. S., C. R. Jonas, L. R. Teras, M. J. Thun, and E. E. Calle. 2004. Weight gain, body mass index, hormone replacement therapy, and postmenopausal breast cancer in a large prospective study. *Cancer Epidemiology Biomarkers & Prevention* 13(2): 220-224.
- Feuer, E. J., L. M. Wun, C. C. Boring, W. D. Flanders, M. J. Timmel, and T. Tong. 1993. The lifetime risk of developing breast cancer. *Journal of National Cancer Institute* 85(11): 892-897.
- Field, T. S., D. S. M. Buist, C. Doubeni, S. Enger, H. Fouayzi, G. Hart, E. J. Korner, L. Lamerato, D. J. Bachman, and J. Ellis. 2005. Disparities and survival among breast cancer patients. *Journal of National Cancer Institute Monographs* 35: 88-95.
- Fisher, E. S., D. C. Goodman, and A. Chandra. 2008. *Regional and racial variation in health care among Medicare beneficiaries: A brief report of the Dartmouth atlas project: 1-42*. The Robert Wood Johnson Foundation's Aligning Forces for Quality Program.
- Flegal, K. M. 2005. Epidemiologic aspects of overweight and obesity in the United States. *Physiology & Behavior* 86(5): 599-602.
- Fleiss J. L. 1981. *Statistical methods for rates and proportions*. New-York: John Wiley and Sons.
- Fotheringham, A. S., C. Brundson, and M. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. West Sussex, England: Wiley & Sons.

- Fox, A. J., P. O. Goldblatt, and D. R. Jones. 1985. Social class mortality differentials: Artifact, selection or life circumstances? *Journal of Epidemiology and Community Health* 39(1): 919-933.
- Freeman, H. P. 2003. Commentary on the meaning of race in science and society. *Cancer Epidemiology Biomarkers & Prevention* 12(3): 232-236.
- Freeman, H. P., and K. C. Chu. 2005. Determinants of cancer disparities: Barriers to cancer screening, diagnosis, and treatment. *Surgical Oncology Clinics of North America* 14(4): 655-669.
- Freeman, H. P., and S. H. Reuben. 2002. *Voices of a broken system: Real people, real problems. President's Cancer Panel, Report of the Chairman 2000-2001.* Washington, DC: National Cancer Institute.
- Georgia, N. L., J. Polacek, M. C. Ramos, and R. L. Ferrer. 2007. Breast cancer disparities and decision-making among US women. *Patient Education and Counseling* 65(2): 158-165.
- Gerend, M. A., and M. Pai. 2008. Social determinants of black-white disparities in breast cancer mortality: A review. *Cancer Epidemiology Biomarkers & Prevention* 17(11): 2913-2923.
- Girgis, A., A. Boyes, R. W. Sanson-Fisher, and S. Burrows. 2000. Perceived needs of women diagnosed with breast cancer: Rural versus urban location. *Australian and New Zealand Journal of Public Health* 24(2): 166-173.
- Glazier, R. H., M. I. Creatore, P. Gozdyra, F. I. Matheson, L. S. Steele, E. Boyle, and R. Moineddin. 2004. Geographic methods for understanding and responding to disparities in mammography use in Toronto, Canada. *Journal of General Internal Medicine* 19(9): 952-961.
- Gomez, S. L., S. Tan, T. H. M. Keegan, and C. A. Clarke. 2007. Disparities in mammographic screening for Asian women in California: A cross-sectional analysis to identify meaningful groups for targeted intervention. *BMC Cancer* 7: 201.
- Goodman, E. 1999. The role of socioeconomic status gradients in explaining differences in US adolescents' health. *American Journal of Public Health* 89(10): 1522.

- Goovaerts, P. 2010. Visualizing and testing the impact on late-stage breast cancer incidence: a non-parametric geostatistical approach. *International Journal of Health and Place* 16: 321-330.
- Goovaerts, P. 2005. Geostatistical analysis of disease data: Estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *International Journal of Health Geographics* 4: 31.
- Goovaerts, P. and G.M. Jacquez. 2004. Accounting for regional background and population size in the detection of spatial clusters and outliers using geostatistical filtering and spatial neutral models: the case of lung cancer in Long Island, New York. *International Journal of Health Geographics* 3: 14.
- Goovaerts, P., J. R. Meliker, and G. M. Jacquez. 2007. A comparative analysis of aspatial statistics for detecting racial disparities in cancer mortality rates. *International Journal of Health Geographics* 6: 32.
- Gordon, D. 1995. Census based deprivation indices: their weighting and validation. *British Medical Journal* 49: S39-44.
- Grann, V., A. B. Troxel, N. Zojwalla, D. Hershman, S. A. Glied, and J. S. Jacobson. 2006. Regional and racial disparities in breast cancer-specific mortality. *Social Science & Medicine* 62(2): 337-347.
- Greenlee, R. T., M. B. Hill-Harmon, T. Murray, and M. Thun. 2001. Cancer statistics, 2001. *CA: A Cancer Journal for Clinicians* 51(1): 15-36.
- Greenwald, H. P., N. L., Polissar, and H. H., Dayal. 1996. Race, socioeconomic status and survival in three female cancers. *Ethnicity & Health* 1(1): 65-75.
- Gregorio D. I., M. Kulldorff, L. Barry, and H. Samociuk. 2002. Geographic differences in invasive and in situ breast cancer incidence according to precise geographic coordinates, Connecticut, 1991-95. *International Journal of Cancer* 100: 194-198.
- Gross, C. P., B. D. Smith, E. Wolf, and M. Andersen. 2008. Racial disparities in cancer therapy: did the gap narrow between 1992 and 2002. *Cancer* 112(4): 900-908.

- Gumpertz, M. L., L. W. Pickle, B. A. Miller, and B. S. Bell. 2006. Geographic patterns of advanced breast cancer in Los Angeles: Associations with biological and sociodemographic factors (United States). *Cancer Causes and Control* 17(3): 325-339.
- Guzman B. 2001. *The Hispanic Population: Census 2000 Brief*. Washington, DC: US Dept of Commerce, Economics and Statistics Administration.
- Gwyn, K., M. L. Bondy, D. S. Cohen, M. J. Lund, J. M. Liff, E. W. Flagg, L. A. Brinton, J. W. Eley, and R. J. Coates. 2004. Racial differences in diagnosis, treatment, and clinical delays in a population-based study of patients with newly diagnosed breast carcinoma. *Cancer* 100(8): 1595-1604.
- Haan, M. N., G. A. Kaplan, and S. L. Syme. 1989. Socioeconomic status and health: Old observations and new thoughts. pp 76-135 in J.P. Bunker, D.S. Gomby, and B.H. Kehrer (eds), *Pathways to Health: The Role of Social Factors*. The Henry J Kaiser Family, Menlo Park, CA.
- Hainer, R. 2009. Did Steve Job's money buy him a faster liver transplant? Online. Available from internet: <http://www.cnn.com/2009/HEALTH/06/24/liver.transplant.priority.lists/index.html>, accessed, October 8, 2009.
- Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate data analysis with readings*. Englewood Cliffs, NJ: Prentice-Hall.
- Hall, I. J., B. Newman, R. C. Millikan, and P. G. Moorman. 2000. Body size and breast cancer risk in black women and white women the Carolina breast cancer study. *American Journal of Epidemiology* 151(8): 754-764.
- Hall, S. A., J. S. Kaufman, R. C. Millikan, T. C. Ricketts, D. Herman, and D. A. Savitz. 2005. Urbanization and breast cancer incidence in North Carolina, 1995-1999. *Annals of Epidemiology* 15(10): 796-803.
- Harman, H. H. 1976. *Modern factor analysis*. University of Chicago Press.
- Harper, S., and J. Lynch. 2006. *Methods for measuring cancer disparities: Using data relevant to healthy people 2010 cancer-related objectives*. Bethesda, MD: National Cancer Institute.

- Harper S., and J. Lynch 2005. *Methods for Measuring Cancer Disparities: Using Data Relevant to Health People 2010 Cancer-Related Objectives*. NCI Cancer Surveillance Monograph Series. Bethesda, MD: National Cancer Institute.
- Harper, S., J. Lynch, S. C. Meersman, N. Breen, W. W. Davis, and M. E. Reichman. 2008. An overview of methods for monitoring social disparities in cancer with an example using trends in lung cancer incidence by area-socioeconomic position and race-ethnicity, 1992–2004. *American Journal of Epidemiology* 167(8): 889-899.
- Haynes, M. A., and B. D. Smedley. 1999. *The unequal burden of cancer: An assessment of NIH research and programs for ethnic minorities and the medically underserved*. National Academy Press.
- Haynes, R., G. Bentham, A. Lovett, and S. Gale 1999. Effects of distance to hospitals and GP surgery on hospital inpatient episodes, controlling for needs and provision. *Social Science and Medicine* 49 425-433.
- Healthy People 2010: Understanding and improving health. 2000. Washington, DC: US Dept of Health and Human Services, Public Health Service.
- Hebb, J. H., D. Fitzgerald, and W. H. Fan. 2003. Health care disparities in disadvantaged Medicare beneficiaries: A national project review. *Journal of Health Human Service Administration* 26(2): 153-173.
- Heck, K. E., D. K. Wagener, A. Schatzkin, S.S. Devesa, and N. Breen. 1997. Socioeconomic status and breast cancer mortality, 1989 through 1993: an analysis of education data from death certificates. *American Journal of Public Health* 87(7): 1218-1222.
- Henschke, U. K., L. D. Leffall, C. H. Mason, A. W. Reinhold, R. L. Schneider, and J. E. White. 1973. Alarming increase of the cancer mortality in the US black population (1950-1967). *Cancer* 31(4): 763-768.
- Hermon, C., and V. Beral. 1996. Breast cancer mortality rates are leveling off or beginning to decline in many western countries: Analysis of time trends, age-cohort and age-period models of breast cancer mortality in 20 countries. *British Journal of Cancer* 73(7): 955-960.
- Higgs, G. 2004. A literature review of the use of GIS-based measures of access to health care services. *Health Services and Outcomes Research Methodology* 5(2): 119-139.

- Hsu, C. W. E, H. Jacobson, and F. S. Mas. 2004. Evaluating the disparity of female breast cancer mortality among racial groups - a spatiotemporal analysis. *International Journal of Health Geographics* 3 (4).
- Hsu, J. L., S. L. Glaser, and D. W. West. 1997. Racial/Ethnic differences in breast cancer survival among San Francisco bay area women. *Journal of the National Cancer Institute* 89(17): 1311-1312.
- Hunt, K. K., G. L., Robb, E. A., Strom, and N. T., Ueno 2007. *Breast Cancer*. Springer.
- Ibrahim, S. A. 2007. Racial and ethnic disparities in hip and knee joint replacement: A review of research in the Veterans affairs health care system. *Journal of the American Academy of Orthopedic Surgeons* 15: S87-94.
- Ishibe, N., S. E. Hankinson, G. A. Colditz, D. Spiegelman, W. C. Willett, F. E. Speizer, K. T. Kelsey, and D. J. Hunter. 1998. Cigarette smoking, cytochrome P450 1A1 polymorphisms, and breast cancer risk in the nurses' health study. *Cancer Research* 58(4): 667-671.
- Israel, B. A., A. J. Schulz, E. A. Parker, and A. B. Becker. 1998. Review of community-based research: Assessing partnership approaches to improve public health. *Annual Review of Public Health* 19(1): 173-202.
- Jemal, A., R. Siegel, E. Ward, T. Murray, J. Xu, and M. J. Thun. 2007. Cancer statistics, 2007. *CA: A Cancer Journal for Clinicians* 57(1): 43-66.
- Jolliffe, I. T. 2002. *Principal component analysis*. Springer.
- Jones, B. A., S. V. Kasl, M. G. M. C. Curnen, P. H. Owens, and R. Dubrow. 1997. Severe obesity as an explanatory factor for the black/white difference in stage at diagnosis of breast cancer. *American Journal of Epidemiology* 146(5): 394-404.
- Jordan, H., P. Roderick, D. Martin, and S. Barnett. 2004. Distance, rurality and the need for care: Access to health services in south west England. *International Journal of Health Geographics* 3(1): 21.
- Joslyn S. A., and M. M. West. 2000. Racial differences in breast carcinoma survival. *Cancer* 88:114-123.

- Juarez, P. D., P. L. Robinson, and P. Matthews-Juarez. 2002. 100% access, zero health disparities, and GIS: An improved methodology for designating health professions shortage areas. *Journal Health Social Policy* 16(1-2): 155-167.
- Kafadar, K. 1994. Choosing among two-dimensional smoothers in practice. *Computational Statistics & Data Analysis* 18(4): 419-439.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20: 141-151.
- Kaplan, J. B., and T. Bennett. 2003. Use of race and ethnicity in biomedical publication. *Journal of American Medical Association* 289(20): 2709-2716.
- Kelsey, J. L., M. D. Gammon, and E. M. John. 1993. Reproductive factors and breast cancer. *Epidemiologic Reviews* 15(1), 36-47.
- Kempe, A., and P. H. Wise. 1993. The racial disparity in very-low-birth-weight - reply. *New England Journal of Medicine* 328(4): 286-286.
- Keppel, K. G. 2007. Ten largest racial and ethnic health disparities in the united states based on healthy people 2010 objectives. *American Journal of Epidemiology* 166(1): 97-103.
- Keppel, K. G., E. Pamuk, J. Lynch, O. Carter-Pokras, Insun Kim, V. Mays, J. Percy, V. Schoenbach, and J. S. Weissman. 2005. Methodological issues in measuring health disparities. *Vital and Health Statistic Series 2, Data Evaluation and Methods Research* 141: 1-16.
- Keppel, K. G., J. N. Percy, and R. J. Klein. 2004. *Measuring progress in Healthy People 2010*. Statistical Notes, No.25. Hyattsville, MD: National Center for Health Statistics.
- Khang, Y.H., J.W. Lynch, and G.A. Kaplan. 2004. Health inequalities in Korea: Age-and sex-specific educational differences in the 10 leading causes of death. *International Journal of Epidemiology* 33(2): 299-308.
- Kim, S. H., J. Ferrante, B. R. Won, and M. Hameed. 2008. Barriers to adequate follow-up during adjuvant therapy may be important factors in the worse outcome for black women after breast cancer treatment. *World Journal of Surgical Oncology* 6: 26.

- Klonoff, E. A. 2009. Disparities in the provision of medical care: An outcome in search of an explanation. *Journal of Behavioral Medicine* 32(1): 48-63.
- Krieger, N. 2002. Is breast cancer a disease of affluence, poverty, or both? The case of African American women. *American Journal of Public Health* 92(4): 611-613.
- . 2001. Theories for social epidemiology in the 21st century: An ecosocial perspective. *International Journal of Epidemiology* 30(4): 668-677.
- . 1992. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *American Journal of Public Health* 82(5): 703-710.
- Krieger, N., D., William, and N. Moss. 1997a. Can we monitor socioeconomic inequalities in health? A survey of U.S. health departments' data collection and reporting practices. *Public health Rep* 112: 481-491.
- . 1997b. Measuring social class in US public health research: Concepts, methodologies, and guidelines. *Annual Reviews in Public Health* 18(1): 341-378.
- Krieger, N., J. T. Chen, P. D. Waterman, D. H. Rehkopf, R. Yin, and B. A. Coull. 2006. Race/ethnicity and changing US socioeconomic gradients in breast cancer incidence: California and Massachusetts, 1978–2002 (United States). *Cancer Causes and Control* 17 (2): 217-226.
- . 2003. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: A comparison of area-based socioeconomic measures—the public health disparities geocoding project. *American Journal of Public Health* 93(10): 1655.
- . 2005. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: The public health disparities geocoding project. *American Journal of Public Health* 95 (2): 312-323.
- Krieger, N., J. T. Chen, P. D. Waterman, M. J. Soobader, S.V. Subramanian and R. Carson. 2002. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? The public health disparities geocoding project. *American Journal of Epidemiology* 156 (5): 471.

- Kutner, M. H., J. Neter, C. J. Nachtsheim, and W. Wasserman. 2004. *Applied Linear Statistical Models*. Homewood, IL: McGraw-Hill/Irwin.
- Lachin J. M. 2000. *Biostatistical methods: the assessment of relative risks*. New-York: John Wiley and Sons.
- Lannin, D. R., H. F. Mathews, J. Mitchell, and M. S. Swanson. 2002. Impacting cultural attitudes in African-American women to decrease breast cancer mortality. *The American Journal of Surgery* 184(5): 418-423.
- Lannin, D. R., H. F. Mathews, J. Mitchell, M. S. Swanson, F. H. Swanson, and M. S. Edwards. 1998. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. *Journal of American Medical Association* 279(22): 1801-1807.
- LaVeist, T. A. 2005. Disentangling race and socioeconomic status: A key to understanding health inequalities. *Journal of Urban Health* 82(2): 26-35.
- Lee, M. C. 2000. Knowledge, barriers, and motivators related to cervical cancer screening among Korean-American women: A focus group approach. *Cancer Nursing* 23(3): 168-175.
- Legler, J., N. Breen, H. Meissner, D. Malec, and C. Coyne. 2002. Predicting patterns of mammography use: A geographic perspective on national needs for intervention research. *Health Services Research* 2002(37): 929-947.
- Leon, D. A., and G. Walt. 2001. *Poverty, inequality, and health: An international perspective*. Oxford University Press Oxford.
- Li, C. I., K. E. Malone, and J. R. Daling. 2003. Differences in breast cancer stage, treatment, and survival by race and ethnicity. *Archives of Internal Medicine* 163(1): 49-56.
- Lillie-Blanton, M., R. M. Martinez, and A. Salganicoff. 2001. Site of medical care: do racial and ethnic differences persist. *Yale Journal Health Pol'y L. & Ethics* 1: 15.
- Lillie-Blanton M, and T. Laveist. 1996. Race/ethnicity, the social environment, and health. *Social Science Medicine* 43(1): 83-91.
- Lloyd G. E. R. 1978. *Hippocratic writings*. Harmondsworth: Pelican Classics.

- Long, J. A., M. I., Bamba, B. Ling, and J. A. Shea. 2006. Missing race/ethnicity data in Veteran's health administration based disparities research: A systematic review. *Journal of Health Care for the Poor and Underserved* 17(1): 128-140.
- Longley, P., and M. Batty. 1996. *Spatial Analysis: Modeling in a GIS Environment*. Cambridge, UK: Geoinformation International.
- Lynch, J. W., G. D. Smith, G. A. Kaplan, and J. S. House. 2000. Income inequality and mortality: Importance to health of individual income, psychosocial environment, or material conditions. *British Medical Journal* 320(7243): 1200.
- Macleod, U., S. Ross, C. Twelves, W. D. George, C. Gillis, and G. C. M. Watt. 2000. Primary and secondary care management of women with early breast cancer from affluent and deprived areas: Retrospective review of hospital and general practice records. *British Medical Journal* 320(7247): 1442-1445.
- MacNaughton, N. S. 2008. Health disparities and health-seeking behavior among Latino men: A review of the literature. *Journal of Transcultural Nursing* 19 (1): 83-91.
- Madigan, M. P., R. G. Ziegler, J. Benichou, C. Byrne, and R. N. Hoover. 1995. Proportion of breast cancer cases in the United States explained by well-established risk factors. *Journal of National Cancer Institute Cancer Spectrum* 87 (22): 1681.
- Magnusson, C. M. K., A. W. Roddam, M. C. Pike, C. Chilvers, B. Crossley, C. Hermon, K. McPherson, J. Peto, M. Vessey, and V. Beral. 2005. Body fatness and physical activity at young ages and the risk of breast cancer in premenopausal women. *British Journal of Cancer* 93: 817-824.
- Mandelblatt, J. S., C. B. Schechter, K. R. Yabroff, W. Lawrence, J. Dignam, P. Muennig, Y. Chavez, J. Cullen, and M. Fahs. 2004. Benefits and costs of interventions to improve breast cancer outcomes in African American women. *Journal of Clinical Oncology* 22(13): 2554.
- Marchand, L. L. 1991. Ethnic variation in breast cancer survival: A review. *Breast Cancer Research and Treatment* 18: 119-126.
- Marks, J., N. Lee, H. W. Lawson, R. Henson, J. K. Bobo, and M. K. Kaeser. 2000. Implementing recommendations for the early detection of breast and cervical cancer among low-income women. *Morbidity and Mortality Weekly Report* 49: 35-55.

- Marmot, M. G., M. J. Shipley, and G. Rose. 1984. Inequalities in death--specific explanations of a general pattern? *Lancet* 1(8384): 1003-1006.
- Marrone, S. 2007. Understanding barriers to health care: A review of disparities in health care services among indigenous populations. *International Journal of Circumpolar Health* 66 (3): 188-198.
- Maskarinec, G., and J. J. Noh. 2004. The effect of migration on cancer incidence among Japanese in Hawaii. *Ethnicity & Disease* 14(3): 431-439.
- May, D. S., N. C. Lee, L. C. Richardson, A. G. G. Giustozzi, and J. K. Bobo. 2000. Mammography and breast cancer detection by race and Hispanic ethnicity: Results from a national program (United States). *Cancer Causes and Control* 11(8): 697-705.
- McCarthy, E. P., R. B. Burns, S. S. Coughlin, K. M. Freund, J. Rice, S. L. Marwill, A. Ash, M. Shwartz, and M. A. Moskowitz. 1998. Mammography use helps to explain differences in breast cancer stage at diagnosis between older black and white women. *Annals of Internal Medicine* 128(9): 729.
- McGinnis, L. M. S., H. R. Menck, H. J. Eyre, K. I. Bland, C. E. H. Scott-Conner, M. Morrow, and D. P. Winchester. 2000. National cancer data base survey of breast cancer management for patients from low income zip codes. *Cancer* 88 (4): 933-945.
- McPherson, K., C. M. Steel, and J. M. Dixon. 2000. Breast cancer-epidemiology, risk factors, and genetics. *British Medical Journal* 321(7261): 624-628.
- Meade, M. S., and R. J. Earickson. 2005. *Medical geography*. The Guilford Press.
- Meersman, S. C., N. Breen, L. W. Pickle, H. I. Meissner, and P. Simon. 2009. Access to mammography screening in a large urban population: A multi-level analysis. *Cancer Causes and Control* 20 (8): 1469-1482.
- Meliker, J. R., Goovaerts, P., Jacquez, G. M., AvRuskin, G. A., Copeland, G. 2009. Breast and prostate cancer survival in Michigan: can geographic analyses assist in understanding racial disparities. *Cancer* 115(10): 2212-2221.
- Menashe, I., W. F. Anderson, I. Jatoi, and P. S. Rosenberg. 2009. Underlying causes of the black-white racial disparity in breast cancer mortality: A population-based analysis. *Journal of the National Cancer Institute* 101(14): 993-1000.

- Meng, L. 1997. Ethnic differences and factors related to breast cancer survival in Hawaii. *International Journal of Epidemiology* 26(6): 1151-1158.
- Merkin, S. S., L. Stevenson, and N. Powe. 2002. Geographic socioeconomic status, race, and advanced-stage breast cancer in New York City. *American Journal of Public Health* 92 (1): 64-70.
- Mezzetti, M., C. La Vecchia, A. Decarli, P. Boyle, R. Talamini, and S. Franceschi. 1998. Population attributable risk for breast cancer: diet, nutrition, and physical exercise. *Journal of National Cancer Institution* 90(5): 389-394.
- Moorman, P. G., B. A. Jones, R. C. Millikan, I. J. Hall, and B. Newman. 2001. Race, anthropometric factors, and stage at diagnosis of breast cancer. *American Journal of Epidemiology* 153(3): 284-291.
- Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37: 17-23.
- Morris, C. R. 2000. Increasing trends in the use of breast-conserving surgery in California. *American Journal of Public Health* 90(2): 281-284.
- Morrison, J., V. Garry, M. E. Harkins, M. Cohen, and J. Palermo. 2000. Regional differences in cardiovascular mortality in Minnesota. *Minnesota Medicine* 83(4): 41-46.
- Mullenix, P. S., T. A. Brown, M. O. Meyers, L. R. Giles, E. R. Sigurdson, M. C. Boraas, J. P. Hoffman, B. L. Eisenberg, and M. H. Torosian. 2005. The association of cytokeratin-only-positive sentinel lymph nodes and subsequent metastases in breast cancer. *The American Journal of Surgery* 189(5): 606-609.
- National Cancer Institute. 2001. *SEER Summary Staging Manual – 2000: Codes and Coding Instructions*. Bethesda, MD: National Cancer Institute.
- Nattinger, A. B., R. T. Kneusel, R. G. Hoffmann, and M. A. Gilligan. 2001. Relationship of distance from a radiotherapy facility and initial breast cancer treatment. *Journal of the National Cancer Institute* 93(17): 1344.
- Neal, R. D., and V. L. Allgar. 2005. Sociodemographic factors and delays in the diagnosis of six cancers: Analysis of data from the national survey of NHS patients: Cancer. *British Journal of Cancer* 92(11): 1971-1975.

- O'Brien, E. D., R. S. Bailie, and P. L. Jelfs. 2000. Cervical cancer mortality in Australia: Contrasting risk by aboriginality, age and rurality. *International Journal of Epidemiology* 29(5): 813-816.
- Openshaw S. 1984. *The Modifiable Areal Unit Problem*. Norwich, UK: GeoBooks.
- Owusu, C., T. L. Lash, and R. A. Silliman. 2007. Effect of undertreatment on the disparity in age-related breast cancer-specific survival among older women. *Breast Cancer Research and Treatment* 102(2): 227-236.
- Palmer, R. C., and E. C. Schneider. 2005. Social disparities across the continuum of colorectal cancer: A systematic review. *Cancer Causes & Control* 16(1): 55-61.
- Pattison, W. D. 1964. The four traditions of geography. *Journal of Geography* 63: 211-216.
- Pearcy, J. N., and K. G. Keppel. 2002. A summary measure of health disparity. *Public Health Reports* 117(3): 273-280.
- Pearson, K. 1901. LIII. on lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6* 2 (11): 559-572.
- Pellow, D. N. 2000. Environmental inequality formation: toward a theory of environmental injustice. *American Behavior Science* 43: 581-601
- Perou, C. M., T. Sørli, M. B. Eisen, M. Rijn, S. S. Jeffrey, C. A. Rees, J. R. Pollack, D. T. Ross, H. Johnsen, and L. A. Akslen. 2000. Molecular portraits of human breast tumors. *Nature* 406(6797): 747-752.
- Piette, J. D., and R.H. Moos, 1996. The influence of distance on ambulatory care use, death, and readmission following a myocardial infarction. *Health Services Research* 31(5): 573.
- Potosky, A. L., G. F. Riley, J. D. Lubitz, R. M. Mentnech, and L. G. Kessler. 1993. Potential for cancer related health services research using a linked Medicare-tumor registry database. *Medical Care* 31(8): 732-748.
- Prehn A., and D. West 1998. Evaluating local differences in breast cancer incidence rates: A census-based methodology (United States). *Cancer Causes Control* 9: 511-517.

- Ravallion, M. 1997. Can high-inequality developing countries escape absolute poverty? *Economics Letters* 56(1): 51-57.
- Redmond, L. J., B. A. Bowman, G. A. Mensah. 2005. *Health Disparities: A Selected Bibliography from the National Center for Chronic Disease Prevention and Health Promotion: January 2000–January 2005*. Atlanta (GA): Centers for Disease Control and Prevention.
- Reintgen, D., R. Giuliano, and C. E. Cox. 2000. Sentinel node biopsy in breast cancer: An overview. *Breast Journal* 6(5): 299-305.
- Remington, R. D., and M.A. Schork. 1985. *Statistics with applications to the biological and health sciences*. Englewood Cliffs, NJ: Prentice-Hall.
- Richardson, J. L., B. Langholz, L. Bernstein, C. Burciaga, K. Danley, and R.K. Ross. 1992. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. *British Journal of Cancer* 65(6): 922.
- Richardson, L. C., J. Schulman, L. E. Sever, N. C. Lee, and R. J. Coates. 2001. Early-stage breast cancer treatment among medically underserved women diagnosed in a national screening program, 1992–1995. *Breast Cancer Research and Treatment* 69 (2): 133-142.
- Ries, L., D. Melbert, M. Krapcho, D. G. Stinchcomb, N. Howlader, M. J. Horner, A. Mariotto, B. A. Miller, E. J. Feuer, and S. F. Altekruse. 2007. *SEER cancer statistics review, 1975-2005*. Bethesda, MD: National Cancer Institute.
- Riley, G. F., A. L. Potosky, C. N. Klabunde, J. L. Warren, and R. Ballard-Barbash. 1999. Stage at diagnosis and treatment patterns among older women with breast cancer an HMO and fee-for-service comparison. *Journal of American Medical Association* 281 (8): 720-726.
- Risser D. R., M. A. Williams, B. Mokry, P. Betts, and E. A. Miller. 2009. *Cancer Incidence & Mortality in Texas, 2001–2005*. Austin, Texas: Texas Department of State Health Services.
- Robert, S. A. 1998. Community-level socioeconomic status effects on adult health. *Journal of Health and Social Behavior* 39(1): 18-37.
- Robinson, J. L. 1976. A new look at the four traditions of geography. *Journal of Geography* 75(9): 520-530.

- Roetzheim, R. G., E. C. Gonzalez, J. M. Ferrante, N. Pal, D. J. Van Durme and J. P. Krischer. 2000. Effects of health insurance and race on breast carcinoma treatments and outcomes. *Cancer* 89, no. 11: 2202-2213.
- Rogerson, P. 2010. *Statistical Methods for Geography*. Los Angeles, CA: Sage Publication Ltd.
- Rogot, E., P. D. Sorlie, N. J. Johnson, and C. Schmitt. 1992. A mortality study of 1.3 million persons by demographic, social, and economic factors: 1979-1985 follow-up. US national longitudinal mortality study. Bethesda, MD: National Institutes of Health, National Heart, Lung, and Blood Institute.
- Ross, C. E., and J. Mirowsky. 2000. Does medical insurance contribute to socioeconomic differentials in health? *The Millbank Quarterly* 72 (2): 291-321.
- Ruffin, M. T., D. W. Gorenflo, and B. Woodman. 2000. Predictors of screening for breast, cervical, colorectal, and prostatic cancer among community-based primary care practices. *Journal of American board of Family Practice* 13(1): 1-10.
- Russell, K. M., and R. Shedd-Steele. 2003. The meaning of health in mammography screening for African American women. *Health Care for Women International* 24(1): 27-39.
- Sabesan, S., and P. Piliouras. 2009. Disparity in cancer survival between urban and rural patients—how can clinicians help reduce it. *Rural and Remote Health* 9: 1146.
- Saha, S., M. Freeman, J. Toure, K. M. Tippens, C. Weeks, and S. Ibrahim. 2008. Racial and ethnic disparities in the VA health care system: A systematic review. *Journal of General Internal Medicine* 23(5): 654-671.
- Sakamoto, Y., M. Ishiguro, and G. Kitagawa. 1986. Akaike information criterion statistics. Dordrecht, Netherlands: Reidel Publishing Company.
- Sarfati, D., T. Blakely, C. Shaw, D. Cormack, and J. Atkinson. 2006. Patterns of disparity: Ethnic and socio-economic trends in breast cancer mortality in New Zealand. *Cancer Causes and Control* 17(5): 671.
- Sequist, T. D., and E. C. Schneider. 2006. Addressing racial and ethnic disparities in health care: Using federal data to support local programs to eliminate disparities. *Health Services Research* 41(4): 1451-1468.

- Shack, L., C. Jordan, C. S. Thomson, V. Mak, and H. Moller. 2008. Variation in incidence of breast, lung and cervical cancer and malignant melanoma of skin by socioeconomic group in England. *BMC Cancer* 8: 271.
- Shambaugh, E. M., and C. E. Platz. 1977. *Summary staging guide for the cancer surveillance, epidemiology and end results reporting (SEER) programs*. Bethesda, MD: National Institute of Health.
- Shlens, J. 2005. A Tutorial on Principal Component Analysis. UCSD: Institute for Nonlinear Science.
- Shavers, V. L., and M. L. Brown. 2002. Racial and ethnic disparities in the receipt of cancer treatment. *Journal of the National Cancer Institute* 94(5): 334-357.
- Singh, G. K., Miller, B. A., Hankey, B. F., and Edwards B. K. 2003. *Area socioeconomic variations in US cancer incidence, mortality, stage, treatment, and survival, 1975-1999*. US Dept. of Health and Human Services, National Institutes of Health, National Cancer Institute.
- Smedley, B. D., A. Y. Stith, and A. R. Nelson. 2002. *For the Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Institute of Medicine Washington, DC: National Academy of Sciences.
- Smigal, C., A. Jemal, E. Ward, V. Cokkinides, R. Smith, H. L. Howe, and M. Thun. 2006. Trends in breast cancer by race and ethnicity: Update 2006. *CA: A Cancer Journal for Clinicians* 56(3): 168.
- Smith, G. D. 2003. *Health Inequalities: Lifecourse Approaches*. Policy Press.
- Stein, J. A., S. A. Fox, and P. Murata. 1991. The influence of ethnicity, socioeconomic status, and psychological barriers on use of mammography. *Journal of Health and Social Behavior* 32(2): 101-113.
- Steve, K., F. Jeanne, W. Bok, and H. Meera. 2008. Barriers to adequate follow-up during adjuvant therapy may be important factors in the worse outcome for black women after breast cancer treatment. *World Journal of Surgical Oncology* 6(1): 26.

- Sturgeon, S. R., C. Schairer, D. Grauman, L. E. Ghormli, and S. Devesa. 2004. Trends in breast cancer mortality rates by region of the United States, 1950–1999. *Cancer Causes and Control* 15(10): 987-995.
- Szasz A., and M. Meuser. 1997. Environmental inequalities: literature review and proposals for new directions in research and theory. *Current Sociology* 45(3): 99–120.
- Talley, L. I., W. E. Grizzle, J. W. Waterbor, D. Brown, H. Weiss, and A. R. Frost. 2002. Hormone receptors and proliferation in breast carcinomas of equivalent histologic grades in pre- and postmenopausal women. *International Journal of Cancer* 98(1): 118-127.
- Tan, A., D. H. Freeman, J. L. Freeman, D. D. Zhang, H. Dayal, and B. U. Philips. 2009. *The Cost of Cancer in Texas, 2007*. The Texas Cancer Registry, Texas Department of State Health Services.
- Tarlov, E., S. N. Zenk, R. T. Campbell, R. B. Warnecke, and R. Block. 2009. Characteristics of mammography facility locations and stage of breast cancer at diagnosis in Chicago. *Journal of Urban Health* 86(2): 196-213.
- Texas Health and Human Services Commission. 2007. *Demographic Profile: Texas by poverty income status in 2006, March 2007 CPS report using 2006 data*. Data compiled by Research and Evaluation Department, Center for Strategic Support.
- Thurstone, L. L. 1931. Multiple factor analysis. *Psychological Review* 38(5): 406-427.
- Tian, N., J. Wilson, and F.B. Zhan. 2010. Female breast cancer mortality clusters within racial groups in the United States. *International Journal of Health & Place* 16 (2): 209-218.
- Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46(2): 234-240
- Tropman, S. E., T. C. Ricketts, E. Paskett, T. A. Hatzell, M. R. Cooper, and T. Aldrich. 1999. Rural breast cancer treatment: Evidence from the reaching communities for cancer care (REACH) project. *Breast Cancer Research and Treatment* 56(1): 59-66.

- U.S. Census Bureau. 2009. 2006-2008 American Community Survey. Online. Available from Internet:
http://factfinder.census.gov/servlet/ACSSAFFFacts?_event=&geo_id=04000US48&geoContext=01000US%7C04000US48&street=&county=&cityTown=&state=04000US48&zip=&lang=en&sse=on&ActiveGeoDiv=&useEV=&pctxt=fph&pgsl=040&submenuId=factsheet_1&ds_name=DEC_2000_SAFF&ci_nbr=null&q_r_n_ame=null®=null%3Anull&keyword=&industry, accessed, September 27, 2009.
- U.S. Department of Health and Human Services. 2000. *Healthy People 2010: Understanding and Improving Health*. Washington, DC: Public Health Service.
- Vainio, H., and F. Bianchini. 2002. *Weight Control and Physical Activity*. France: IARC Press Lyon.
- Vainshtein, J. 2008. Disparities in breast cancer incidence across racial/ethnic strata and socioeconomic status: A systematic review. *Journal of the National Medical Association* 100(7): 833-839.
- Veith, I. 2002. *The Yellow Emperor's Classic of Internal Medicine*. University of California Press.
- Vona-Davis, L., and D. P. Rose. 2009. The influence of socioeconomic disparities on breast cancer tumor biology and prognosis: A review. *Journal of Women's Health* 18(6): 883-893.
- Wang, M., K. D. Burau, S. Fang, H. Wang, and X. L. Du. 2008. Ethnic variations in diagnosis, treatment, socioeconomic status, and survival in a large population-based cohort of elderly patients with non-Hodgkin Lymphoma. *Cancer* 113 (11): 3231-3241.
- Ward, E., A. Jemal, V. Cokkinides, G. K. Singh, C. Cardinez, A. Ghafoor and M. Thun. 2004. Cancer disparities by Race/Ethnicity and socioeconomic status. *CA: A Cancer Journal for Clinicians* 54(2): 78-93.
- Weiss H., S. Soong, E. Partridge, J. Carpenter, B. Bryant, and J. Waterbor. 1997. Cancer incidence among predominantly black, rural-poor populations in southern states. *South Medical Journal* 90: 986-992.
- Wells, B. L., and J. W. Horm. 1992. Stage at diagnosis in breast cancer: Race and socioeconomic factors. *American Journal of Public Health* 82 (10): 1383-1385.

- Wells, K. J., and R. G. Roetzheim. 2007. Health disparities in receipt of screening mammography in Latinas: A critical review of recent literature. *Cancer Control* 14(4): 369-79.
- Wennberg, J. E., and M. M. Cooper. 1994. *The Dartmouth atlas of health care 1996*. Hanover, NH: The Dartmouth Atlas Project.
- Whitehead, M. 1991. The concepts and principles of equity and health. *Health Promotion International* 6(3): 217-228.
- Williams, D. R., and S. A. Mohammed. 2009. Discrimination and racial disparities in health: Evidence and needed research. *Journal of Behavioral Medicine* 32(1): 20-47.
- Williams, D. R. 1996. Race/ethnicity and socioeconomic status: Measurement and methodological issues. *International Journal of Health Services: Planning, Administration, Evaluation* 26(3): 483-505.
- Wilson, S., L. Jones, C. Couseens, K. Hana. 2002. *Cancer and the Environment: Gene-Environment Interactions*. National Academies Press.
- Woods, L. M., B. Rachet, and M. P. Coleman. 2005. Choice of geographic unit influences socioeconomic inequalities in breast cancer survival. *British Journal of Cancer* 92: 1279-1282.
- Young, J. L. 2001. *SEER summary staging manual 2000: Codes and coding instructions*. National Cancer Institute, National Institutes of Health.
- Zar, J. H. 1999. *Biostatistical Analysis*. Upper Saddle River, NJ: Prentice Hall.

VITA

Nancy Tian was born in Houma, ShanXi Province, China on July 26, 1978 to Qibao Tian and Xingming Gao. Nancy grew up in a small village in northern China. In 1999, she graduated with a Bachelor degree in Geo-technical Engineering from Central South University, Changsha. After working as a construction engineer for one year, she returned to graduate school at Hunan University and obtained a master's degree in Engineering Mechanics in 2003. She then worked as a bridge designer in Beijing in 2004 and later supervised a hotel remodeling project in Xiamen in 2005.

Nancy immigrated to Austin, TX with her husband, Les Stewart for the purpose of fulfilling her dream of pursuing a doctoral degree. In 2006, she began the doctoral program in Geographic Information Science at the Department of Geography, Texas State University-San Marcos. During graduate school, she was very active in teaching undergraduate and graduate GIS labs. To enrich her academic experience, she also attended a variety of professional conferences related to GIS and Public Health. She will begin her postdoctoral work with the National Exposure Research Laboratory at Environmental Protection Agency in August 2010.

Email: nancytiangeo@gmail.com

This dissertation was typed by Nancy Tian.

