SPATIAL CLUSTER ANALYSIS OF FEMALE BREAST CANCER DIAGNOSIS IN MISSOURI: USING GIS AND SPATIAL ANALYST FUNCTIONS

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

FAUSTINE WILLIAMS, MS, MPH

Dr. David O'Brien, Dissertation Supervisor

Dr. Stephen Jeanetta, Co-Advisor

JULY 2012

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

SPATIAL CLUSTER ANALYSIS OF FEMALE BREAST CANCER DIAGNOSIS IN MISSOURI: USING GIS AND SPATIAL ANALYST FUNCTIONS

Presented by Faustine Williams, MS, MPH,

A candidate for the degree of doctor of philosophy,

And hereby certify that, in their opinion, it is worthy of acceptance.

Professor David O'Brien

Professor Stephen Jeanetta

Professor Mary Grigsby

Professor Chris Fulcher

Professor John Fresen

Professor Jeannette Jackson-Thompson

DEDICATION

This dissertation is dedicated to the following:

To my family, for their believe and confidence in me; and

To PEO and IPS women, for their wonderful support, love and care during my studies

ACKNOWLEDGEMENTS

I joined the Rural Sociology program at the University of Missouri, Columbia in the fall of 2009 to pursue my doctoral studies. Along the way, there are so many people who have made almost my three year (35 months) stay a memorable one, and also the timely completion of my dissertation.

First of all, thanks unto the Almighty God, who has promised never to leave nor forsake me. Your everlasting love and grace indeed is more than sufficient for me. I would not have come so far, without His protection, and direction. Your presence alone will I continue to seek after.

I will like to express my profound graduate to my dissertation committee members: Drs. David O'Brien, Stephen Jeanetta, Mary Grigsby, Chris Fulcher, John Fresen and Jeannette Jackson-Thompson for working tirelessly with me. To Dr. O'Brien, my mentor, advisor and dissertation supervisor, thank you very much for your insightful guidance, comments and taking off time your busy schedule to meet with me regularly. My co-advisor Dr. Jeanetta, you were not only my advisor rather you treated me like you own daughter and wanted me not only to succeed academically, but also in all areas of my life. I will forever be grateful for your amazing mentoring, support and care. Most of the exposures and skills I have acquired throughout the program were all because of your guidance. There is absolutely no word to describe my sincere appreciation. God richly bless you! Dr. Grigsby, I thank you for your comments that have helped to adequately shape my research. I will always remember the very first advice you gave me when we met..."it is always important in life to identify your strength, then build and develop it. It is the only way that will lead you to the top. Never tried to be anyone else…always be you"! Dr. Chris Fulcher, thank you for your great comments and recommendations that have enabled me to come out with a good research. Thank you also, for believing in me, and also for offering a graduate research assistantship at the Center for Applied Research and Environmental System (CARES). The opportunity provided me with lots of exposures and skills for my future career. Dr. John Fresen, wow! I don't know where to begin from. I really appreciate your selfless attitude and kindness. I am especially intrigued about how you set aside your precious time and worked with me hours on the entire research. Please accept my heartfelt gratitude for all you have and continue to do for me. God richly bless you! Dr. Jackson-Thompson thank you for allowing me use the Missouri Cancer Registry and Research Center (MRC-ARC) data for this study. I also want to express my gratitude for your great support, advising and comments provided to me. I will forever be grateful and very much appreciative you were all part of this research.

…to all faculty and staff in the Department of Rural Sociology, I very much appreciate your guidance and support throughout my studies. You were all always available to provide me with great feedback and mentorship. I was never turned away even though I asked the same question over and over again. Dr. Jere Gilles and the doctoral admission committee, I just want to say, thank you for making my dream of getting a Ph.D. come to pass by giving me admission into the program. So many good things have happened to me since I joined the program. Sincerely, I could not have gotten such an amazing exposure if I was not admitted into the program. I am so confident that the training and skills I have acquired over the years in the program has fully prepared me to be an advocate for change in my community.

…to all the staff at CARES for their help, and encouragement: Chris Barnett, Yan Barnett, Erin, Barbaro, Michael Barbaro, Richard Charrier, Dave Connett, James, H. Cutts III, Michael Hardly, Jonathan Lippmann, Anthony Strand, Angela Johnson, Angie Stewart, Eloi Espanol, and many others.

...the staff of American Cancer Society (ACS) map place, Dr. Bill Elder and the entire Office of Social and Economic Data Analysis (OSEDA) family, MCR-ARC especially, Chester Lee Schmaltz, and Mary Jane King, University of Missouri Master of Public Health Program, Department of Geography and Health Sciences Library for their valuable support.

…PEO/IPS women especially, Wanda Eason, Chapter GN in Kansas, and all the other Chapters for the parental support they extended to me, as well as their valuable friendship and companionship.

…Leah Graham for introducing me to Wanda Eason to be nominated for the PEO/IPS scholarship. Love you and God bless!

…Jeff Partridge – Department of Health and Senior Services for promptly answering all the data questions. Dr. K.C. Morrison for his advice, encouragement and also for offering me with graduate teaching assistantship for his African Politics class, and Dr. Sheryl A. Tucker for supporting and helping me through my transfer of program.

…my friends and colleagues in the Rural Sociology Department especially Michelle Kaiser, Charles Duvon and Iris Zachary, I will always appreciate your friendship.

…my own family: father (Evans Williams), mother (Rose Agbakpewu), step-mother (Christine Williams), sisters (Dela Akatse, Priscilla Williams, Francisca Williams, Vincentia Doh), uncle (Dickson Opare) and other relatives, for supporting me in many ways, and for their constant prayers, and uplifting.

…my dearest "uncle" Dr. Eric Bempah and his family, the Ghanaian community in Columbia, Missouri and beyond, for their motivation, and encouragement… let the Azonto begin!

…everyone else whose name is not mentioned here but has been a great part in my intellectual journey…my sincerest thanks to all of you. You have endowed me with good things and fruitful experiences. Thank you all and God bless!

TABLE OF CONTENTS

ACKNOWLEDGEMENTS II

TABLE OF CONTENTS VI

LIST OF ILLUSTRATIONS X

LIST OF TABLES XII

ABSTRACT XIII

CHAPTER ONE BACKGROUND AND INTRODUCTION 1

- Study Region 3
- The Burden of Cancer in Missouri over the Years 5
- Purpose of the Study 11
- Significance of the Study 12
- Anticipated Contribution and Potential Uses of Study Findings 12
- Definition of Terms 13
- Tumor Size and Stage at Diagnosis 13
- Types of Staging – Theoretical 13
- Clinical Staging 13
- Pathologic Staging 14
- Staging Systems 14
  - The TNM System 14
  - Stage Grouping 16
- Rural Health Clinic 17
- Critical Access Hospital 17
- Federally Qualified Health Center 18
- Operational Staging 18
- Potential Access Accessibility – Theoretical 19
- Accessibility Operational 19
- Theoretical Framework 19

Summary 22

CHAPTER TWO LITERATURE REVIEW 23

Introduction 23
Importance of Breast Cancer Preventive Services 24
Spatial Geographic Barriers to Access to Health Care Services 27
Social and Economic Factors in Relation to Access 31
Interaction Effects of Spatial Geographic and Social Factors 34
Defining and Measuring Rurality in America 37
Defining Spatial Geographic Isolation 38
Classification Scheme I: Urban Areas as Defined by the U.S. Census Bureau 40
Classification Scheme II: Core Based Statistical Area as Defined by OMB 41
Classification Scheme III: The Rural-Urban Continuum Code (RUCC) as Defined by USDA/ERS (Beale Codes) 42
Classification Scheme IV: The Rural-Urban Density Typology as Defined by Isserman (2005) 43
Classification Scheme V: The Index of Relative Rurality as Defined by Waldorf (2007) 44
Proposed Appropriate Rural Measurement 45
Objective of the Study 46
Summary 46

CHAPTER THREE DATA AND METHODS 49

Research Questions 49
Hypotheses 49
Study Design and Area 50
Data Sources and Description 50
Cancer Data 50
American Community Survey (ACS) 51
TIGER® Data 52
Environmental Systems Research Institute (ESRI) StreetMap 52

Health Care Facilities 53
Data Analysis Techniques 53
Independent Variables 53
Demography 53
Economy 54
Geography or Spatial Isolation 57
Data Preparation and Software 62
Unit of Analysis 65
Spatial Approach 67
Statistical Modeling and Hypothesis Testing 72
Model Assessment and Selection 73
Summary 74
CHAPTER FOUR ANALYSIS AND FINDINGS 75
Introduction 75
Characteristics and Relevance of the Sample 75
Marginal Distributions of the Predictors 77
Individual Predictors 77
County Level Predictors 80
Conditional Distribution by Derived County Size 88
Boxplot of Education by Race, Stage, Age and Adjacent 91
Distribution of Breast Cancer Incidence in Missouri 96
Distance Measure Description – Travel Time to the Nearest Medical Center in Missouri 105
Logistic Regression to Predict Stage at Diagnosis 112
Summary 117
CHAPTER FIVE DISCUSSION, CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS 119
Discussion 119
Conclusions 123

Implications, Future Research and Recommendations 125
Delimitations of the Study 126
Limitations of the Study 126

APPENDICES 128

A. COMMONLY USED RURAL DEFINITIONS 128
B. DEFINITION OF RUCC AND ALLOCATION OF RUCC TO MISSOURI COUNTIES 130
C. STEPWISE REGRESSION ANALYSIS 130
D. VIOLIN PLOT OF EDUCATION SCORE BY COUNTRY SIZE 144
E. SCATTER PLOT PERCENT BELOW POVERTY, EDUCATION AND ADJACENT 144
F. STEPWISE LOGISRIC REGRESSION FOR AIC SELECTION 145
G. PROPORTION DIAGNOSED LATE BY AGE 146

BIBLIOGRAPHY 147

VITA 160

LIST OF ILLUSTRATIONS

Figure 1.1. Ten Leading Types of Invasive Cancer, Missouri, 2007 6

Figure 1.2. Ten Leading Types of Invasive Cancers, by Gender, Missouri, 2007* 8

Figure 1.3. Ten Leading Causes of Cancer Deaths, Missouri, 2008 9

Figure 1.4. Ten Leading Types of Cancer Deaths, by Gender, Missouri, 2008* 10

Figure 2.1. Definition of Urban Areas 40

Figure 2.2. Definition of Core Based Statistical Areas 41

Figure 2.3. Categorization of U.S. Counties by the Rural-Urban Continuum Code 43

Figure 3.1. Late Stage Breast Cancer Distribution by Race, 2003-2005 61

Figure 3.2. Late Stage Breast Cancer Distribution by Race, 2006-2008 62

Figure 3.3. Summary of Methods Used in the Study 64

Figure 3.4. Rural-Urban Code by County 66

Figure 3.5. Total Distribution of Breast Cancer Diagnosed in Missouri, 2003-2008 68

Figure 3.6. Female Breast Cancer per 100,000 Population, 2003-2005 69

Figure 3.7. Female Breast Cancer per 100,000 Population, 2006-2008 70

Figure 4.1. Marginal Distribution of Individual Variables 77

Figure 4.2. Distribution of Important County Variables 81

Figure 4.3. Percentage Late Stage Distribution by Race and County Size, 2003-2005 ... 83

Figure 4.4. Percentage Late Distribution by Race and County Size, 2006-2008 84

Figure 4.5. Distribution of Female Breast Cancer Patients in Missouri
Metro/Nonmetro and Adjacent/Nonadjacent 85

Figure 4.6. Cancer Distribution by Metro and Rural Urban Adjacent, 2003-2003 86

Figure 4.7. Cancer Distribution by Metro and Rural Urban Adjacent, 2006-2008 86

Figure 4.8. Plots of Education and County Size 88

Figure 4.9. The Conditional Distributions of Education by County Size (smallest to largest, 1 = small; 6 = large) 89

Figure 4.10. Missouri County Education Score 90

Figure 4.11. Distribution of Education Score by Race, Age and Adjacent 92

Figure 4.12. Missouri County Poverty Score 94

Figure 4.13. Distribution of Blacks below FDL 95

Figure 4.14. Proportion of Breast Cancer Diagnosed Late per 100,000 Population 97

Figure 4.15. Proportion of Breast Cancer Diagnosed Late, 2003-2008 98

Figure 4.16. Late Stage Diagnosis for White Women, 2003-2005 100

Figure 4.17. Late Stage Diagnosis for Black Women, 2003-2005 101

Figure 4.18. Late Stage Diagnosis for White Women, 2006-2008 103

Figure 4.19. Late Stage Diagnosis for Black Women, 2006-2008 104

Figure 4.20. Time Travel in Minutes to Mammography Center 107

Figure 4.21. Time Travel to Hospitals in Missouri 108

Figure 4.22. Time Travel in Minutes to Rural Clinics in Missouri 109

Figure 4.23. Time Travel in Minutes to Critical Access Hospital in Missouri 110

Figure 4.24. Time Travel in Minutes to FQHC in Missouri 111

Figure 4.25. Probability of Late Diagnosis by Race, Age and County Education Score 116

LIST OF TABLES

Table 2.1 *The Rural-Urban Density Typology* 44

Table 3.1. *New Rural Type Defined* 59

Table 4.1. *Distribution of Study Population Diagnoses* 80

Table 4.2. *Description of the Derived Variable "County Population"* 82

Table 4.3. *Final Model from Stepwise Regression Selection* 113

Table 4.4. *Fit for Race and Education for Each Level of Age* 115

SPATIAL CLUSTER ANALYSIS OF FEMALE BREAST CANCER DIAGNOSIS IN MISSOURI: USING GIS AND SPATIAL ANALYST FUNCTIONS

Faustine Williams

Dr. David O'Brien, Dissertation Supervisor
Dr. Stephen Jeanetta, Co-Advisor

ABSTRACT

The stage at cancer diagnosis has a tremendous impact on type of treatment, recovery and survivor. In most cases the earlier the cancer is detected and treated the higher the survival rate for the patient. Various studies have indicated disparities in access to primary care especially access to screening services like mammography for early detection. The purpose of this research was to examine the role of spatial access to health care services on the probability of late detection of female breast cancer diagnosis in Missouri taking into account access and distance to clinics and hospitals. All cancer cases were categorized into two main groups: **early** defined as *in situ* and localized stages and **late** as regional and distant stages. Geographic information system (GIS), spatial analyst functions and logistic regression methods were used to analyze county-level incidence of female breast cancer in Missouri from 2004 to 2008. The GIS results showed that the majority of women in rural Missouri counties do not have access to screening and other health care services. Women had to travel over 60 minutes one way for medical care. This travel burden resulted in a higher probability of late detection. The logistic regression indicated that among younger white and black women, the effect of race and county-level educational score on late detection was similar. For the older group, the effect of race and in particular the lack of education on late detection was greater among blacks than whites. Over all, the age of a woman, race and county-level educational score of residence were the most statistically significant factors in predicting late stage cancer diagnosis among women in Missouri.

CHAPTER ONE
BACKGROUND AND INTRODUCTION

The leading cause of death globally is cancer, costing the world economy almost one trillion dollars per year (ACS & Livestrong, 2010). The total economic impact of premature death and disability from cancer worldwide was $895 billion in 2008, 20 percent higher than heart disease (ACS & Livestrong, 2010). In addition, GLOBOCAN estimates that by 2030, 21.4 million new cases of cancer will be diagnosed annually, an increase of almost 70 percent from 12.7 million in 2008. Of these cases approximately 13.2 million will die from the disease, up 72 percent in 2008 from 7.6 million (Ferlay et al., 2010; McCormack & Boffetta, 2011).

In the United States (U.S.) according to the National Institute of Health (NIH), the overall costs of cancer in 2007 were $226.8 billion: $103.8 billion for direct medical costs (total of all health expenditures) and $123.0 billion for indirect mortality costs (ACS, 2012). Most researches have indicated that the lack of insurance and other barriers are responsible for the increasing health care costs since many Americans are not able to receive the optimal health care until they are at terminal stage of their diseases. However, the number of uninsured Americans kept increasing. According to the U.S. Census Bureau, in 2009 over 50 million Americans were uninsured, an increase of 16.7 percent, from 46.3 million uninsured in 2008 (US Census Bureau, 2010). Of these almost one-third of Hispanics (32 percent) had no health insurance coverage. Uninsured patients and those from ethnic minorities are substantially more likely to be diagnosed with cancer at a later stage, when treatment can be more extensive and more costly (ACS, 2012). However, the critical element in reducing deaths from cancer is early diagnosis.

Even though there are different kinds of cancers, breast cancer has remained the most frequent malignancy affecting women across all racial and ethnic groups apart from skin cancer. In 2012 it has been estimated that 226,870 new cases of invasive breast cancer are expected to occur among women in the U.S.; about 2,190 new cases are expected in men. In addition to invasive breast cancer, 63,300 new cases of in situ breast cancer are expected to occur among women in 2012. Of these, approximately 85 percent will be ductal carcinoma in situ (DCIS) (ACS, 2012). Since 2004, in situ breast cancer incidence rates have been stable in white women, but increasing in African American women by 2.0 percent per year (ACS, 2012).

In spite of advances in medical technology leading to early diagnosis and treatment, breast cancer ranks as the second leading cause of death in women closely following lung cancer (ACS, 2012). An estimated 39,920 breast cancer deaths (39,510 women, 410 men) are expected in 2012. However, breast cancer death rates have steadily decreased in women since 1990, with larger decreases in younger women; from 2004 to 2008, rates decreased 3.1percent per year in women younger than 50 and 2.1percent per year in women 50 and older. The decrease in breast cancer death rates represents progress in earlier detection, improved treatment, and possibly decreased incidence (ACS, 2012).

Detection of cancer while it is still small and confined provides the best chance of effective treatment. Benefits of early detection include increased survival, increased treatment options and improved quality of life. Therefore, to improve access to screening services, the United States Congress passed the Breast and Cervical Cancer Mortality Prevention Act of 1990 (Public Law 101-354), which guided the Centers for Disease Control and Prevention (CDC) in establishing the National Breast and Cervical Cancer Early Detection Program (NBCCEDP) (CDC, 2011). The program operates in all 50 states, the District of Columbia, six US territories, and twelve American Indian/ Alaska Native organizations. In Missouri, the Breast and Cervical Cancer Control Program (BCCCP), now called the *Show Me Healthy Women* (SMHW) program, started in 1992. The goal of the program is to reduce breast and cervical cancer mortality and morbidity by increasing availability of cancer screening for early detection of breast or cervical cancer among women in high-risk populations. High-risk women include, women whose income is under 200 percent of the federal poverty level (FPL), are under 65 years of age, have little or no health insurance, and women with disabilities etc. (DHSS, 2011).

Study Region

Missouri is a state located in the Midwest part of the United States. Missouri comprises 114 counties and the independent city of St. Louis. The four largest urban areas are St. Louis, Kansas City, Springfield, and Columbia. According to the 2010 U.S. Census, the population of Missouri was 5,988,927, with a population density of 89.9 persons/km^2; over half of Missouri residents (3,294,936 people, or 55.0 percent) live within the state's two largest metropolitan areas - St. Louis and Kansas City. The racial and ethnic composition of the state's population is 81.0 percent White/Non-Hispanic, 11.6 percent African-American/Non-Hispanic, 3.5 percent Hispanic/Latino, 1.6 percent Asian/Non-Hispanic, 0.5 percent American Indian/Alaskan Native, 0.1 percent Native Hawaiian and other Pacific Islander, and 2.1 percent two or more races. Thirty-seven percent of Missouri's population is rural, equating to approximately 2.22 million people in rural areas. The fastest growing ethnic group in Missouri is the Hispanic population. Statewide, there was a 79.2 percent increase in Hispanics between the 2000 Census and the 2010 Census (U.S. Census Bureau, 2010 Census).

The Missouri Department of Health and Senior Services report indicated that the state mortality rate from all causes of death was 871.5 for 1999 to 2009. Of the 50 counties with an age-adjusted death rate from all causes that is statistically significantly higher than the state rate, 46 are rural. The majority of those counties are in the southern areas of the state (MO DHSS, 2011). However, according to United State Bureau of Economic Analysis Missouri's total personal income increased by 2.2 percent from 2009 to 2010. The state's growth rate lagged behind the U.S. increase of 3.0 percent. Missouri's per capita personal income grew by 3.7 percent from 2009 to 2010 (MO DHSS & Economic Research and Information Center, 2010). The State's urban areas had a higher median household income. Eighty-two of the 89 Missouri counties having a poverty rate greater than the overall state rate are rural. The average poverty rate for Missouri's rural counties was approximately 17.2 percent, while in urban counties the average poverty rate was approximately 13.1 percent (MO DHSS & Department of Economic Development, 2010).

Another characteristic closely tied to poverty as an indicator of the financial health of a community is the unemployment rate. In December 2010, 56 counties in Missouri had an annual average unemployment rate greater than the state (MO DHSS, 2010). The economic recovery continues, with jobs in all sectors impacted at varying levels and degrees. Nonetheless, in rural Missouri, the lack of educational attainment, as measured by the percentage of population without a high school education, is evident. Thirty-six rural counties have more than 20 percent of the population over 25 years of age without a high school education (MO DHSS, 2010).

Health insurance is an important determinant of health status, access and utilization of health care services. Health insurance is also highly correlated with income. Lack of insurance, along with reduced access to health care delivery services, is a dangerous combination that exists disproportionately in rural Missouri. According to the *2007 County-Level Study*, approximately 75 percent of all Missouri counties have a rate of individuals without insurance greater than the state rate (MO DHSS, 2007). Rural areas generally have higher rates of individuals without insurance than do urban counties (MO DHSS, 2007). Hence the aim of this research is to assess the impact of geographic location on stage at breast cancer diagnosis in the State.

The Burden of Cancer in Missouri over the Years

Over the last two decades, the breast cancer incidence rate has been decreasing after peaking at 142 per 100,000 women in 1999. The dramatic decline of almost 7 percent from 2002 to 2003 has been attributed to reduction in the use of menopausal hormone therapy (MHT), previously known as hormone replacement therapy (HRT), following the publication of results from the Women's Health Initiative in 2002; this study found that the use of combined estrogen plus progestin MHT was associated with an increased risk of breast cancer, as well as coronary heart disease (ACS, 2012). From 2004-2008, the most recent five years for which data are available, breast cancer incidence rates were stable (ACS, 2012).

Decreasing cancer-related morbidity and mortality requires continued focus on the cancer continuum. As part of the SMHW program, currently more than 1,180 Missouri women had been approved for cancer treatment through the breast and cervical cancer treatment (BCCT) program. However, in spite of this progress, breast cancer deaths in the state remain high. Based on data from the Missouri Cancer Registry and Research Center (MCR-ARC), 29,695 of Missouri's residents were diagnosed with invasive cancer in2007 (DHSS, 2010). This amounted to more than three new cases of cancer, diagnosed every hour of every day in Missouri. The five leading invasive cancers in 2007 were lung and bronchus; prostate; female breast; colon, rectum and rectosigmoid; and urinary bladder (Figure 1.1).

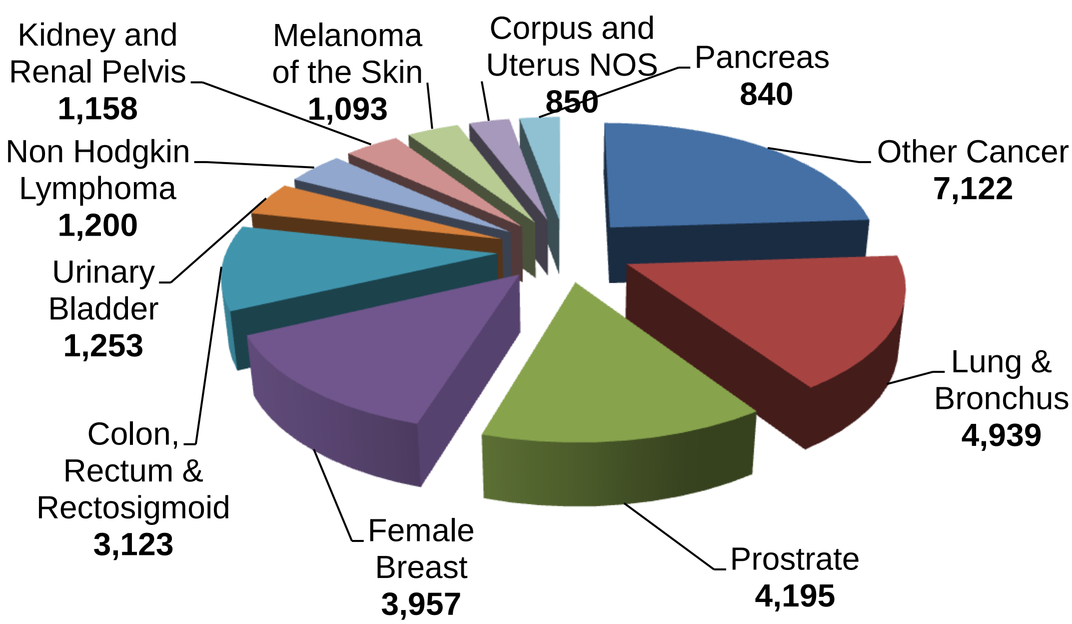


Figure 1.1. Ten Leading Types of Invasive Cancer, Missouri, 2007

Source: Missouri Department of Health and Senior Services, Cancer Registry, MICA www.dhss.mo.gov/data/mica/mica/cancer_19sites.php

Specifically, among females, the five leading cancers were breast; lung and bronchus; colon, rectum and rectosigmoid; corpus and uterus not otherwise specified (NOS); and non-Hodgkin lymphoma (Figure 1.2). These five sites accounted for 64.0 percent of all new cancer cases among women.

Female		Male	
Breast 3,957	(28%)	Prostate 4,195	(27%)
Lung & bronchus 2,189	(15%)	Lung & bronchus 2,749	(18%)
Colon, rectum & rectosigmoid 1,485	(10%)	Colon, rectum & rectosigmoid 1,640	(11%)
Corpus and Uterus, NOS 850	(6%)	Urinary bladder 946	(6%)
Non-Hodgkin lymphoma 592	(4%)	Kidney and Renal Pelvis 733	(5%)
Thyroid 494	(3%)	Melanoma of the skin 615	(4%)
Melanoma of the skin 478	(3%)	Non-Hodgkin lymphoma 608	(4%)
Kidney and Renal Pelvis 425	(3%)	Oral cavity and pharynx 555	(4%)
Pancreas 414	(3%)	Pancreas 390	(3%)
Ovary 403	(3%)	Leukemias 384	(2%)
Other cancers 2,899	(20%)	Other cancers 2,694	(17%)
All sites 14,186	(100%)	All sites 15,509	(100%)

*Excludes basal and squamous cell skin cancers and in situ cancer except urinary bladder. Percentages may not total 100 percent due to rounding.

Source: Missouri Department of Health and Senior Services. Cancer Registry MICA, www.dhss.mo.gov/data/mica/mica/cancer_19sites.php

Figure 1.2. Ten Leading Types of Invasive Cancers, by Gender, Missouri, 2007*

Similarly in 2008, 12,497 Missouri residents died from cancer, accounting for 22.2 percent of all deaths in Missouri (DHSS, 2010). Cancer is second only to heart disease, as a leading cause of death in Missouri. In 2008, the five leading causes of cancer deaths in Missouri were: lung, bronchus, and trachea; colon, rectum, and anus; female breast; pancreas; and prostate (Figure 1.3).

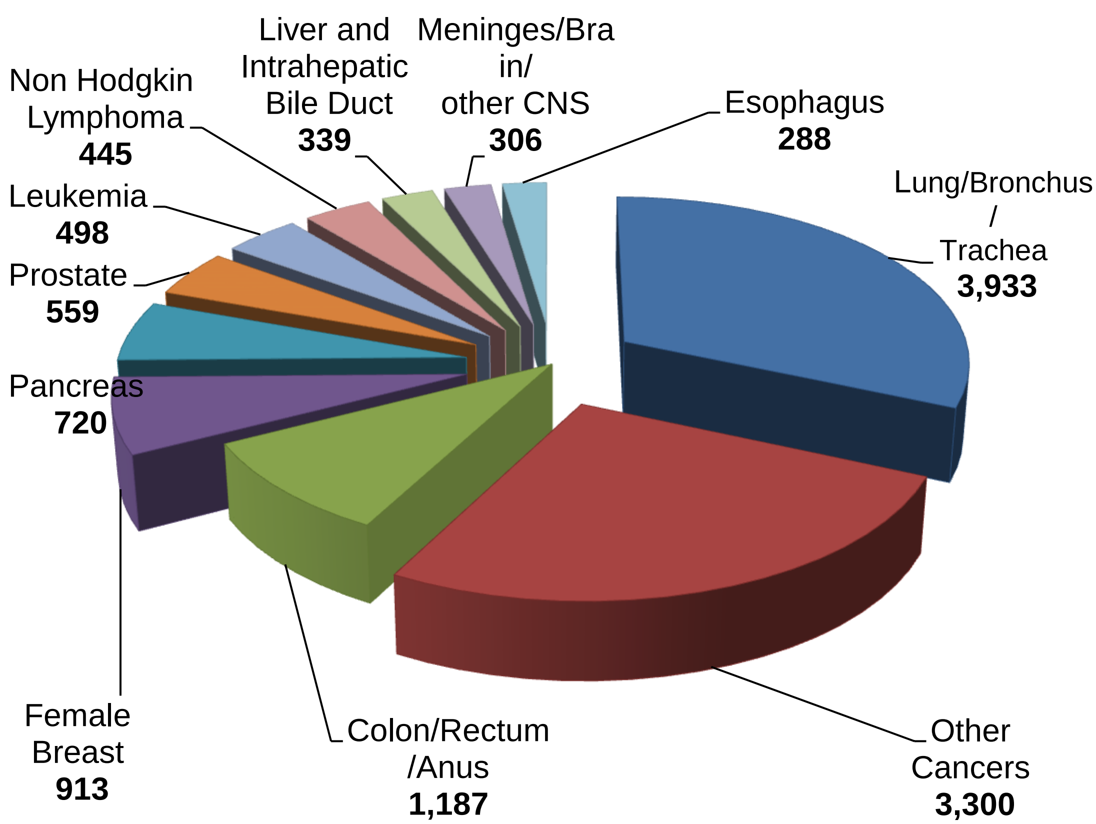


Figure 1.3. Ten Leading Causes of Cancer Deaths, Missouri, 2008

These five main leading causes of cancer deaths have not changed from the period of 1996-2000 to 2008. A comparison of female and main leading cause of deaths in Missouri is presented below (Figure 1.4).

Female		Male	
Lung/bronchus/trachea	1,680 (28%)	Lung/bronchus/trachea	2,253 (35%)
Breast	913 (15%)	Colon/rectum/anus	603 (9%)
Colon/rectum/anus	584 (10%)	Prostate	559 (9%)
Pancreas	359 (6%)	Pancreas	361 (6%)
Ovary	258 (4%)	Leukemia	290 (4%)
Corpus and Uterus, NOS	258 (4%)	Non-Hodgkin lymphoma	227 (3%)
Non-Hodgkin lymphoma	218 (4%)	Liver and Intra-hepatic Bile Duct	226 (3%)
Leukemia	208 (3%)	Esophagus	219 (3%)
Meninges/Brain and Other Parts of Central Nervous System	151 (3%)	Urinary Bladder	198 (3%)
Liver and Intra-hepatic Bile Duct	113(2%)	Kidney and Renal Pelvis	180 (3%)
Other cancers	1,261 (21%)	Other cancers	1,378 (21%)
All sites	6,003 (100%)	All sites	6,494 (100%)

* Excludes basal and squamous cell skin cancer and in situ cancer except urinary bladder. Percentages may not total 100 percent due to rounding.

Source: Missouri Department of Health and Senior Services, Death MICA
http://www.dhss.mo.gov/DeathMICA/index.html

Figure 1.4. Ten Leading Types of Cancer Deaths, by Gender, Missouri, 2008*

Further, the recent *Cancer Facts and Figures* (2012) by the American Cancer Society (ACS), estimates that 4,400 new breast cancers are expected in 2012 and almost 900 Missourian women will die from the disease (ACS, 2012). On the other hand, increasing access to preventive services to all women will reduce late stage cancer diagnosis, improve quality of life of survivors and also reduce the high mortality rates.

Purpose of the Study

In 2003 female breast cancer mortality in Missouri slipped from 26.1 deaths per 100,000 to 28 per 100,000 in 2005. The state is currently ranked 49th, just ahead of Louisiana and the District of Columbia (Health Management Associates Inc., 2011). Between 2000 and 2007 numerous researches have shown that after adjusting for individual risk factors, there are neighborhood differences in cancer screening, incidence, treatment and survival (DHSS, 2010). It is the relationship between place, race and poverty that can lead to the greatest disparities. Reducing such disparities requires action at several levels to maximize impact.

The purpose of this research was to examine the role of spatial access to health care services on the probability of late detection of female breast cancer diagnosis in Missouri taking into account access and distance to clinics and hospitals. This is necessary in order to ensure prompt and adequate access to health care services is available for all cancer patients regardless of place of residence and economic situation. At this time of the study, access is defined as an individual's ability to obtain medical services on a timely and financially affordable basis. Factors determining ease of access include availability of health care facilities and transportation to them as well as reasonable hours of operation (Jonas, Goldsteen & Goldsteen, 2007).

Significance of the Study

Detection of any breast and many cancers at an early stage is the key to improved survival and decreased mortality rates. However, review of the literature suggests that most of the past studies have focused solely on the effects of access and distance travel on early or late stage breast cancer diagnosis and treatment among females. This may be first study to understand the association between place of location, available health care services and the two broad groups of stage at diagnosis (early vs. late). Therefore, in this research the aim is to fill the gap in this area. This study will examine the impact of spatial and other demographic factors on diagnosis and treatment of breast cancer in Missouri taking into account available clinics and hospitals. Another important part of this study is to discover unidentified barriers to cancer diagnosis, screening and treatment among women in Missouri.

Anticipated Contribution and Potential Uses of Study Findings

Cancer costs billions of dollars in years of productive life lost. Above and beyond the financial costs, there are huge emotional costs related to losing loved ones prematurely. Reducing barriers to cancer care is critical in the fight to eliminate suffering and death due to the disease. Findings from this study will provide the basis for developing strategies aimed at improving access to breast cancer screening services for low-income, uninsured and underserved women in Missouri. Finally, based on the findings from this study recommendations will be made to the SMHW program and other state policy makers on actions that must be taken to improve on health care services to ensure early diagnosis and treatment.

Definition of Terms

The following terms are theoretically and operationally defined for the purposes of this study.

Tumor Size and Stage at Diagnosis

Cancer staging describes the extent or spread of the disease at the time of diagnosis. Proper staging is essential in determining the choice of therapy and in assessing prognosis. A cancer’s stage is based on the primary tumor’s size and whether it has spread to other areas of the body. According to Abeloff et al. (2004), the size of the tumor is inversely related to the survival rate of a patient with cancer. Larger tumors at the time of first diagnosis are associated with a higher risk of death from any kind of cancer, especially breast carcinoma.

Types of Staging – Theoretical

Cancer staging is done at the time of diagnosis, before any treatment is given. This staging is based on two major types of staging: (1) clinical staging, and (2) pathologic staging

Clinical Staging

This is an estimate of how much cancer there is based on the physical exam, imaging tests (x-rays, CT scans, etc.), and tumor biopsies. For some cancers, the results of other tests, such as blood tests, are also used in staging. The clinical stage is a key part of deciding the best treatment to use. It is also the baseline used for comparison when looking at the cancer's response to treatment.

Pathologic Staging

Pathological staging (also called surgical staging) relies on information obtained during surgery. Often this is surgery to remove the cancer and nearby lymph nodes, but sometimes surgery may be done to look at how much cancer is in the body and remove tissue samples. In some cases, the pathologic stage may be different from the clinical stage (for example, if the surgery shows the cancer has spread more than it was thought to have spread before surgery). The pathological stage gives the health care team more precise information that can be used to predict treatment response and outcomes (prognosis).

Staging Systems

A number of different staging systems are used to classify tumors:

The TNM System

The American Joint Committee on Cancer (AJCC) developed the *TNM classification system* as a tool for doctors to stage different types of cancer based on certain standards. It has replaced many of the older staging systems. In the TNM system, each cancer is assigned a T, N, and M category (AJCC, 2009).

T: Tumor

The T category describes the original (primary) tumor. The tumor size is usually measured in centimeters (2 and 1/2 centimeters is about 1 inch) or millimeters (10 millimeters = 1 centimeter).

- TX means the tumor can't be measured.
- T0 means there is no evidence of primary tumor (it cannot be found).

- Tis mean that the cancer cells are only growing in the most superficial layer of tissue, without growing into deeper tissues. This is also known as *in situ* cancer or *pre-cancer*.
- The numbers T1, T2, T3, and T4 describe the tumor size and/or level of invasion into nearby structures. The higher the T number, the larger the tumor and/or the more it has grown into nearby tissues.

N: Lymph Nodes

The N category describes whether or not the cancer has spread into nearby lymph nodes.

- NX means the nearby lymph nodes cannot be evaluated.
- N0 means nearby lymph nodes do not contain cancer.
- The numbers N1, N2, and N3 describe the size, location, and/or the number of lymph nodes involved. The higher the N number, the more the lymph nodes are involved.

M: Metastasis

The M category tells whether there are distant metastases (spread of cancer to other parts of body).

- MX means metastasis can't be evaluated.
- M0 means that no distant metastases were found.
- M1 means that distant metastases were found (the cancer has spread to distant organs or tissues).

Stage Grouping

Once the values for T, N, and M have been determined, they are combined, and an overall stage is assigned. For breast cancer for instance, five stages, ranging from zero to four, help explain the extent of disease in a patient at the time of diagnosis. In these stages, many sub-sections exist that help to more exactly diagnose a cancer.

Stage 0: This represents the finding that no evidence of a primary tumor, regional lymph node metastasis, or distant metastasis exists. Stage 1: Equates to the finding of a tumor which is 2cm or smaller at its greatest dimension. No regional lymph node metastasis or distant metastasis is noted.

Stage 2: Is used when the disease has spread to adjacent lymph nodes. This stage can be followed by either an A or B postscript. Stage 2A: Means that a tumor greater than 2cm but smaller than 5cm at its greatest dimension was found. Stage 2B: This represents a tumor greater than 5cm across at its greatest dimension. Metastasis to ipsilateral axiallary lymph nodes is also noted with no distant metastasis.

Stage 3: Describes a more advanced stage of disease and has A, B, and C post scripts. In stage 3 of the disease, tumor sizes can range anywhere from a quite small tumor to much larger sizes, but there is direct extension (spread) of the disease to the chest wall or skin. Metastasis to ipsilateral axillary lymph nodes fixed to one another or to other structures is possible.

Stage 4: Includes characteristics of all of the preceding stages along with “distant metastasis”, commonly known as spreading of the cancer to other parts of the body (Breast Cancer Organization, 2011; AJCC & ACS, 2009).

A different system of summary staging (in situ, local, regional, and distant) is used for descriptive and statistical analysis of tumor registry data by the Surveillance Epidemiology and End Results (SEER). (1) *in-situ,* pre-invasive malignancies, those that do not invade the basement membrane; (2) *localized,* invasive malignancies that are confined to the organ of origin; (3) *regional,* the cancer spread by direct extension to adjacent organs or tissues, and/or spread to lymph nodes considered regional to the organ of origin, but no further spread has occurred; (4) *distant,* the disease has spread beyond adjacent organs or tissues, and/or metastasis to distant lymph nodes or tissues; and (5) *unknown,* where the stage was either unknown or not recorded due to insufficient information available to determine stage of disease at diagnosis

Rural Health Clinic

A Rural Health Clinic (RHC) is a clinic located in a rural, medically under-served area. RHCs were established by the Rural Health Clinics Act (P.L. 95-210), (Section 1905 of the Social Security Act). The program was established to address an inadequate supply of physicians serving Medicare beneficiaries and Medicaid recipients in rural areas and to increase the utilization of non-physician practitioners. RHCs can be public, private or non-profit.

Critical Access Hospital

Critical Access Hospital (CAH) is a small, generally geographically remote facility that provides outpatient and inpatient hospital services to people in rural areas. The designation was established by law, for special payments under the Medicare program. To be designated as a CAH, a hospital must be located in a rural area; provide 24-hour emergency services; have an average length-of-stay for its patients of 96 hours or less; be located more than 35 miles (or more than 15 miles in areas with mountainous terrain) from the nearest hospital; or be designated by its State as a "necessary provider". CAHs may have no more than 25 beds.

Federally Qualified Health Center

Federally Qualified Health Center (FQHC) is defined as a clinic that is recognized and certified by the Centers for Medicare and Medicaid Services (CMS) that provides care to low income and medically underserved communities. FQHC was created by Congress, and the national network of community health centers to provide high quality affordable primary and preventive care for those whom other providers do not serve, regardless of an individual's ability to pay. In order to achieve the "federally qualified" status, clinics must adhere to the following key health center requirements: (1) Be located in or serve a high need community (designated Medically Underserved Area or Population); (2) governed by a community board composed of a majority (51 percent or more) of health center patients who represent the population served; (3) provide comprehensive primary health care services, as well as supportive services (education, translation and transportation, etc.) that promote access to health care; (4) provide services available to all with fees adjusted based on ability to pay; and (5) meet other performance and accountability requirements regarding administrative, clinical and financial operations.

Operational Staging

For public health research purposes the stages are frequently classified into two main groups – early and late or distant stages. Early stage - includes in situ and localized stages of disease, while regional and distant stages of disease are referred to us late stage

(Amornsiripanitch et al., 2010). Therefore, based on the aim of this study these two main staging categories will be used.

Potential Access Accessibility – Theoretical

Access: Is defined here as a concept representing the degree of "fit" between the clients and the system (Penchansky & Thomas, 1981).

Accessibility: The relationship between the location of supply and the location of clients, taking account of client transportation resources and travel time, distance and cost (Penchansky &Thomas, 1981). Thus, the geographical or locational relationship between services providers (e.g. hospitals, rural clinics, critical access hospitals, federal qualified health centers), and surrounding populations.

Accessibility Operational

Straight line, travel distance, and travel time are measures of accessibility for this study. Network travel time (e.g. 0-15, 15-30, 30-45, 45-60 and more than 60 minutes) and network mileage will be used for computation of potential access to health care facilities in Missouri. Straight line Euclidean distance from each county centroid to the nearest health facility will also be used to measure access in ArcGIS 10.

Theoretical Framework

Health care access and utilization behavior is complex and multifaceted. While issues of chronic diseases such as cancer, diabetes, etc., are the leading causes of death and disability around the world, it is difficult to mention a particular theory or model in either the social sciences or behavioral sciences that formed the basis of this research. I believe this research broadly touches on various theories such as structuralist and functionalist. However, the main basic assumptions underlying the solutions proposed in this research were Parsons' sick role, Mechanic's general theory of help seeking, and Andersen's health behavior model (Parsons, 1951; Mechanic, 1978; Mojtabai, Olfson & Mechanic, 2002; Aday & Andersen, 1974; Andersen, 1995).

According to Parsons' sick role theory, when an individual is sick, they adopt a role of being ill. This sick role has four main components: (1) the individual is not responsible for their state of illness and is not expected to be able to heal without assistance; (2) the individual is excused from performing normal roles and tasks; (3) there is general recognition that being sick is an undesirable state; and (4) to facilitate recovery, the individual is expected to seek medical assistance and to comply with medical treatment. Parsons' theory attempted to identify typically seen behavior in individuals who are ill (Parsons, 1951). However while a sick person may desire to get better, factors such as income, age, education, race and place of location could create some hindrances for that person and his desire to getting well. Unfortunately, Parson did not effectively address these issues.

The second fundamental theory is Mechanic's (1978) general theory of help seeking behavior. According to this theory, there are multiple levels of help seeking. Individuals experience symptoms; attempt to evaluate the significance of their symptoms and the likely consequences; determine whether they have a problem that requires intervention and could benefit from treatment; evaluate the benefits and costs of various treatments; and choose which health care providers to consult. On the other hand, while individuals may perceive a need and be willing to seek professional help, there are some objective factors such as financial problems, time or cost of care that influence this decision (Mechanic, 1978; Mojtabai, Olfson & Mechanic, 2002).

Access to health care has been demonstrated to act as an important determinant of the use of health services and resulting health outcomes (Campbell, Elliot, Sharp, Richie, Cassidy & Little, 2000). Various empirical studies have also demonstrated that individuals are more likely to report satisfaction with services (Young, Dobson & Byles, 2001) and utilize services when they are closer (Arcury et al, 2005; Pierce, Williamson & Kruse, 1998). Consequently, proximity to health care services can acts as a significant determinant of preventive health care use (Field & Briggs, 2001). Hence, the third framework underlying this study was based on Andersen's (1974) "Behavioral Model and Access to Medical Care" which was further modified in 1995 (Andersen, 1995). There are three characteristics within this framework that determines an individual's health services use.

1. **Predisposing Factors:** Variables that exist before the onset of the illness that describe the individual propensity to use services. Measures of this component include age, sex, race, religion, and values about health and illness.
2. **Enabling Characteristics:** Means or resources individual have available for the use of services. Individual or family resources include income and insurance coverage, while attributes of the community of residence include rural-urban character and region.
3. **Need Based Characteristics:** Level of illness that brings about health service use. Using this model the assumption of this research is that the lack of timely access to health care services may potentially cause adverse health outcome as

evidence by late stage cancer diagnoses and treatment as well as higher mortality rates when distance to health facilities and distance time travel are taking into account.

Summary

Differences in access to health care services and its resulting adverse health outcomes are major public health priorities. As a result, the U.S. Department of Health and Human Services (USDHHS, 2010) has continued to make it a national priority to improve the health and well-being of all Americans. At the same time, health care delivery is becoming more complex due to the growing divers population as well as the frequent changes in the provision of health care services. Improving health care access, reducing geographical differences in health outcomes, and eliminating disparities are essential social and political issues.

In addition, many inequalities exist within the U.S. health care systems. These disparities have been shown to restrict access to health care services especially to vulnerable populations, thereby leading to regional and local differences in health outcomes. Decreasing access and the growing number of at risk women breast cancer populations in Missouri will contribute to excessively higher breast cancer mortality in Missouri. This study therefore aimed at examining these county, rural and urban differences in health care access and making recommendation to policy makers in the state on how to bridge the disparity gap.

CHAPTER TWO
LITERATURE REVIEW

Introduction

Differences in access to health care services and the resulting adverse health outcomes are major public health priorities. As a result, the Institute of Medicine (IOM, 2002) and the U.S. Department of Health and Human Services (USDHHS, 2000), have identified the need for strategies to improve access to health care services and to support improvement of health outcomes (IOM, 2002). In addition, due to the varying degrees in access to health care services, Healthy People 2020 has designated several goals to improving the nation's health. Among these goals are: (1) attain high-quality, longer lives free of preventable disease, disability, injury, and premature death, and (2) eliminate disparities, and improve the health of all groups (USDHHS, 2010). Findings of this nature indicate that while most Americans have high quality health care available, gaps in health care access and health outcomes continue to exist. These differences are associated with age, education, race and ethnicity, gender, income and socioeconomic status (SES), place of residence and location of health care services.

Health care policy changes over the past decade have also drastically decreased access to health care services. The rural health environment in particular has been impacted by these changes in many ways (Bushy, 2000; Folland, et al., 2001). Significant decreases in health care services to an already vulnerable, at-risk rural population have compounded existing problems of resource disparities. Loss of community health services, health care professional shortages, rapidly rising cost, hospital closures, homecare cut backs, and tighter government payment schedules are just a few of the changes that have led to greater resource disparities for rural populations (USDHHS, 2010; Eberhardt, et al., 2001). Because of structural, financial and socio-cultural barriers in rural populations, they have fewer health care resources than urban populations. These rural resource disparities also lead to complex adverse health outcomes and rural health status disparities (Fryer, et al., 1999; Lovett, Haynes, Sunnenberg, & Gale, 2002; Lin, Allen, & Penning, 2002).

Importance of Breast Cancer Preventive Services

As noted by many researchers, one of the greatest successes in cancer control over the past two decades plus in the United States is the dramatic decline in the death rate (Evans, 2011 et al., 2011; Kopans, 2011). In 1989, the death rate for female breast cancer, corrected to the 2000 US standardized population, was 33.2 per 100,000. In 2007, it was 22.8 per 100,000.1 (Evans, 2011; Alterkruse et al., 2010). This was a 31.3 percent decrease, and American Cancer Society (ACS) epidemiologists estimate that this translates into more than 75,000 American women saved from a death from breast cancer. It is also predicted that if the decrease continues at this same rate, the mortality reduction will approach 50 percent by 2015 (Evans, 2011; Alterkruse et al., 2010).

The substantial decline in breast cancer deaths has predominantly been attributed to two main factors. The first is improvement mammography screening, and secondly technological advancements in medical and biomedical sciences leading to early detection and treatment (Peipins et al., 2011; Evans et al., 2011; Kopans, 2011). Other researchers have also reported that the most effective method of detecting early breast cancer and reducing cancer mortality is mammography screening. Coldman et al., (2007) reported that during the period from 1988 through 2003 in British Columbia, breast cancer deaths among women ages between 40 and 79 years who were screened annually decreased by 40 percent. Among women ages 40 to 49 years, there was a 39 percent mortality reduction at first screening. Similarly in Sweden, after 20 years of follow-up, Duffy et al., (2002) noted that women screened had a 44 percent lower risk of death from breast cancer across all age groups than those not screened. On the whole a 48 percent decrease in breast cancer deaths was found in women ages 40 to 49 years.

A population based mammography screening in the Netherlands which compared rates in 1986 to 1988 found that breast cancer mortality rates in women aged 55–74 years fell significantly in 1997 and subsequent years as predicted, reaching 19·9 percent in 2001. Prior to this, mortality rates had been increasing by an annual **0·3 percent** until screening was introduced. Thereafter mammography screening services were introduced, a decline of **1·7 percent** per year **(95% CI 2·39–0·96)** in women aged 55–74 years and of **1·2 percent** in those aged 45–54 **(2·40 to 0·07 percent)** were noted (Otto et al., 2003). The authors also noted that the turning point in mortality trends arose at around year 0. Adjuvant systemic therapy is unlikely to be the cause of this turning point, since the mortality rates continued to rise up to one year after implementation in municipalities where screening began after 1995. In spite of these achievements, in the 2009 U.S. Preventive Services Task Force (USPSTF) new report on breast cancer screening guidelines recommended that women should only undergo biannual mammography screening beginning at age 50. The decisions to start screening at an earlier age should be made solely on an individual basis such as medical or family history (Aragon, Morgan, Wong, Sharon, 2011; USPSTF, 2009; Nelson et al., 2009; Kopans, 2009; Evans, Poston, Poston, 2011; Peipins, 2011). Meanwhile, earlier USPSTF recommendations and the current recommendations of the American Cancer Society, and the American College of Radiology, and the Society of Breast Imaging recommend annual mammography screening every 1 to 2 years beginning at age 40 (Peipins et al., 2011; Lee et al., 2010; Smith, Cokkinides, Brawley, 2009; USPSTF, 2002).

The publication of the USPSTF guidelines generated a lot of national controversy and frustration among women advocate, research community, leading to many studies on benefits of mammography usage among women. Researchers such as Evans et al, (2011); Kopans, 2011; Hendrick & Helvie, 2011) reported that the meta-analysis of the randomized control trail (RCT) data made available to the USPSTF in formulating their screening recommendations showed a rather statistically significant benefit from invitation to screening in each of three subdivided age cohorts: 39–49, 50–59, and 60–69 years. Aragon et al. (2011) also noted that although the USPSTF recommendations advice against routine screening mammography for women aged 40–49 years, their results demonstrated that nearly one quarter of women in California with early breast cancer, which are likely to be screen detected, are in this younger age group and would be excluded from screening. The authors concluded that implementation of the USPSTF recommendations would disproportionately impact Hispanic, Asian/PI, and non-Hispanic black women. The sheer magnitude of early breast cancer cases among non-white women implies that the majority of young women could be significantly affected by the potential diagnostic delays resulting from these recommendations especially since the patient's quality of life depends on stage at diagnosis.

Spatial Geographic Barriers to Access to Health Care Services

Since the 19th century, distance to health care services has been recognized as a major barrier to health care access in the U.S. Access to health care services may also be fundamentally limited by proximity, which can be measured in travel time (Wang et al., 2008). A review of the literature revealed that distance and other geographical factors are often viewed as major intervening aspects for access to medical care and resultant health outcomes, specifically for the disadvantaged population from both developed and developing nations (Jordon, Roderick, Martin & Barnett 2004; Cromley & Cromley 2009; Peters et al., 2008). Studies in developing nations have shown that the absence of good roads and lack of proper communication particularly in the poor, remote and adverse geographic areas constrain access to health care resulting in poor health outcomes (Baker & Gesler 2000, Rahman & Smith 2000, Peters et al., 2008).

Owen, Obregon and Jacobsen (2010), analyzed the impact of geographic access to health services in rural Guatemala and indicated that the poorest communities in Alta Verapaz have the least geographic access to health center. This is consistent with other analyses that reported that the proportion of residents who sought care when ill, who were seen by a doctor when sick, and who visited a hospital for care all increased steadily from the poorest quintile through the richest quintile (Makinen et al., 2000; Khan et al., 2006; Onwujekew & Uzochukwu, 2004). A study from Kenya found that health facility use decreased significantly when access to health facilities required traveling more than 5 km, or approximately one hour of travel time (Noor et al., 2005). In Papua New Guinea it was found that people living more than 3.5km from a clinic were half as likely to seek care when ill as those living nearer to a health care facility (Muller, Smith, Mellor, Rate

& Genton, 1998), and a study from Pakistan showed differential use of health care facilities for those living more than 5 km from a town center than for those living closer (Noorali, Luby & Rahbar, 1999). Furthermore, a study by Oppong and Hodgson (2005) on spatial accessibility to health care facilities in the Suhum district of Ghana concluded that there is an urgent need for innovative measures to facilitate equal geographical accessibility as well as level of service utilization in order to ensure equity in health services throughout the country.

Similarly, in the United States, distance to health care services has been recognized as a major barrier to health care. According to Wang et al. (2008), access to health care services may be fundamental limited by proximity, which can be measured in travel time. Long distance travel time to health care services has been shown to influence both access and utilization. There is also an assumption that the greater the distance to be travelled, the higher the incidence of psychological morbidity and the poorer the compliance with treatment. For instance, evidence from general psychiatric clinics suggests that patients were more likely to miss appointments as the distance from the clinic increases (Campbell et al., 1991). In the case of cancer patients in particular, increased travel time to health care services has been associated with greater risk of presenting with advanced cancer and many complications.

Another study conducted by Campbell and colleagues (2000) in Scotland on rural factors and cancer survival revealed that increasing distance from a cancer center was associated with greater chance of the patient being recorded as a death certificate only *(DCO – patients for whom only the death certificate provides notification to the cancer registry)* case for stomach, breast and colorectal cancers. In addition, Campbell et al.

(2000) indicated that patients who reside far away from a cancer center are more likely to die on the first day of their diagnosis. Wang et al. (2008) argued that spatial access to primary care doctors and time travel are critically important in archiving high rates of early breast cancer detection in Illinois and surrounding environs. In contrast to these findings, Jones et al. (2008) reported that there was "no evidence of detrimental effects of long car journeys to hospital on cancer survival in Northern England" (Jones et al., 2008, pg. 274).

Geographic access to health care barriers can be classified into two main groups. Geographic (spatial) and socio-organizational (aspatial) access (Aday & Andersen, 1974; Penchansky & Thomas, 1981). Geographic access on the other hand refers to the presence of a staffed medical facility within reasonable travel time of a residence, while socio-organizational access encompasses a great variety of attributes that facilitate or hinder the use of health care services (Owen, Obregón, & Jacobsen, 2010). Geographic barriers are especially important for chronically and critically ill patients, like diabetes, any kind of cancer, asthma, HIV/AIDS etc., who live in rural areas. These patients may be unable to obtain regular treatment and needed care because they do not have access to health care facilities within a reasonable distance (Kerlikowske et. al. 1995).

Other studies on rural factors and survival from cancer in various countries have also noted that geographic location is strongly associated with survival that could also reflect stage at diagnosis and kind of treatment patients are likely to receive (Merkin, Stevenson & Powe, 2002; Brameld & Holman 2006; Jones et al., 2008, Onega et al., 2008; Wang, McLafferty, Escamilla & Liu, 2008; Meliker, Jacquez, Goovaerts, Copeland & Yassine, 2009). Similarly, a comparative study by Liu (2005) and colleagues on "the effects of a national breast and cervical cancer early detection program (NBCCEDP) and women's health network (WHN) on social disparities" in Massachusetts revealed that stage at diagnosis and type of treatment women received is strongly associated with social and demographic factors such as income, type of insurance or education as well as place of residence.

In addition to the above evidences on the impact of spatial factors on diagnosis and treatment of diseases, Baldwin (2008) and colleagues using SEER-Medicare databases found, that more than 25 percent of rural patients with colorectal cancer bypass their closest local small health providers. Patients in most remote area had to travel the longest distance to large rural or urban areas for surgical resections (Baldwin et al., 2008). Onega et al. (2008), who assessed geographic access to cancer care in the U.S. by analyzing traveling distance to nearest specialized cancer care, also revealed that rural dwellers had longer traveling distance to nearest specialized cancer centers than the overall U.S. population. Chan et al. (2006) evaluated how the traveling distance affects Medicare patients' access to health care using 1998 Medicare claims data; it was reported that residents in rural areas needed to travel 2 to 3 times farther to visit medical specialists than urban residents. Also, in Atlanta, the U.S. Census Bureau in 2000 indicated that more than 15 percent blacks do not have access to a private vehicle. Among whites, fewer than 4 percent do not have access to a private vehicle (U.S. Census Bureau, 2000).

A study on the spatial distribution of Chicago's low or no-cost mammography screening facilities, showed overall shorter travel time for low income residents. However, longer travel time and distances were shown for low income black neighborhoods than for other low income neighborhoods (Zenk, Tarlov & Sun, 2006). A study by Meersman, Breen, Pickle, Meissner & Simon (2009) in Los Angeles County showed that mammography use was higher in neighborhoods with a greater density of facilities. Distance to mammography facilities was also associated with late-stage breast cancer diagnosis among Latinas in Los Angeles County and among blacks in segregated areas in Detroit, Michigan as defined by zip codes (Meersman et al., 2009; Dai, 2010). McLafferty and Wang (2009) reported a J-shaped curve for late-stage breast cancer risk was described for women in Illinois with the most highly urbanized area (Chicago) and most isolated rural areas having the highest risk. All these studies have illustrated the significant role spatial factors have and continued to play in the diagnosis and treatment of cancer and other diseases over the years.

Social and Economic Factors in Relation to Access

Research has demonstrated a strong relationship between socioeconomic status (SES) and an increased risk of being affected by health disparities (Alder & Newman, 2002). It has been noted that the leading causes of death and disability have a disproportionate impact on African Americans, Alaska Natives, American Indians, Asian Americans, Hispanic Americans, and Pacific Islanders (Liburd, Giles & Mensah, 2006). Whether assessed by income, level of education, or occupation, SES clearly predicts the health status of an individual. A higher income level provides individuals with means to purchase health insurance and ensures access to health care on a consistent basis. Education has a direct impact on an individual's professional development and career opportunities, which influences access to health coverage. Occupational status has a significant impact on the health status of an individual especially since research has demonstrated that employed individuals have better health than unemployed individuals, otherwise known as the "healthy worker" effect (Alder et al., 2002). These three measurements of SES indirectly influence the impact of health disparities on minority populations but it is important to consider the three main determinants of health that are influenced by SES.

The three main determinants of health include: behavior and lifestyle, environmental exposure, and health care. It has been noted that behavior and lifestyle accounts for 80 percent of premature mortality, environmental exposure for 20 percent and health care for 10 percent (Lee & Paxman, 1997). Individuals of lower SES are more likely to live in poorer communities, which experience a higher degree of residential crowding, violence, and environmental pollution. Poorer housing quality further increases the risk of health conditions for individuals of lower SES. In addition, social environments have a significant impact on SES related health outcomes in regards to risk and prevalence of chronic and infectious diseases. As mentioned earlier, SES determines the ability to purchase health coverage, which has a direct effect on access to health care. Research has demonstrated that uninsured individuals are less likely to receive preventive and primary health care services than insured individuals (Alder et al., 2002). The most significant indirect pathway that influences SES is the impact of behavior and lifestyle. Lower SES is also associated with a sedentary lifestyle as well as poorer nutrition, both of which have an effect on the health status of an individual.

Economic and social factors such as poverty have been directly linked with low usage of mammography screenings (Campbell et al., 2009; MacKinnon, Duncan & Huang et al., 2007). Poverty and low income are associated with lack of health insurance and/or lack of access to primary care which in turn lead to low use of mammography screening (Wang et al., 2008). For example, a Florida study found that black women have lower breast cancer incidence but higher rates of mortality than white women (MacKinnon, Duncan & Huang et al., 2007). This paradox is due to black women not being able to receive regular breast cancer screenings, citing insurance problems and low socioeconomic status as the prime reasons. In contrast, this study found that the white population was wealthier on average and they could afford to obtain regular mammograms (MacKinnon al., 2007). Consistent with the literature on economic barriers, socio-economic deprivation was found to be associated with lower rates of treatment and survival in a study explaining "inequalities in access to treatment of lung cancer" patients in the U.K. (Jack, Gulliford, Ferguson, & Møller, 2006). Schuler et al. (2008) also reported in their study that, women with a lack of health insurance typically have lower rates of mammography utilization than do women with health insurance. Overall, a larger proportion of minority women than white women do not receive regular breast cancer screenings. Some of this is due to lack of health insurance. African American women and Hispanic women have higher rates of not being medically insured which partly accounts for their low rates of mammography screening (Schuler et al., 2008). Among people who do not have health insurance, Chinese and White, non-Hispanic women are less likely to receive a mammogram (Schuler et al., 2008).

Regardless of race or ethnicity, Campbell et al. (2009) noted that poverty has a strong effect on the probability of being diagnosed at the later stages of cancer. As poverty increases by 10 percentage points, the odds of being diagnosed at a regional or distant stage increase by a factor of approximately 1.07, an effect that does not differ by race or ethnicity. Analyzing the geographic differences in late-stage breast cancer in Illinois and the role of socioeconomic and spatial factors, Wang et al. (2008) found that people living in areas of high socioeconomic disadvantage were more likely to be diagnosed with late-stage breast cancer. The risk of late diagnosis was also higher for women living in areas with poor geographical access to primary care physicians, indicating a combination of spatial and socioeconomic barriers. Similarly, MacKinnon et al. (2007) found that minorities and socioeconomically disadvantaged people have lower incidence rates of breast cancer but higher mortality rates because they are unable to seek or obtain screening services. Even if disadvantaged people live near a screening center, they sometimes do not seek help because of economic, cultural and social barriers.

In sum, SES has a significant impact on the health status of individuals, especially minority populations. Reducing the burden of health disparities for minority populations can be achieved by addressing the main determinants of health as well as indirect assessments of SES (income, education and occupation) through appropriate public policy measures that include: reducing gaps in health coverage, improving economic conditions for minority populations, increasing educational opportunities for these populations, and introducing culturally sensitive health promotion efforts that will help reduce the burden of chronic and infectious diseases.

Interaction Effects of Spatial Geographic and Social Factors

Poverty rate is an important social determinate of well-being. However, defining rural poverty in America is as complex as the word rural, because rural America is not a homogeneous entity. While metro and non-metro areas in America have all experienced upward and downward trends in poverty rates over the years, the non-metropolitan rate has always exceeded the metropolitan-rate every year since poverty was first officially measured in the 1960s (Joliffe, 2005). For instance, in 2007, 15.4 percent of the non-metro populace (about 7.4 million people) lived in poverty, while the poverty rate in metro areas was 11.9 percent (U.S. Census Bureau, 2008). Of the 500 poorest U.S. counties, 459 are rural (Housing Assistance Council, 2002). Of 386 persistently poor counties, those with poverty rates greater than 20 percent in each decennial census since 1960, 340 are non-metro (Jolliffe, 2004).

Another study by the United Stated Department of Agriculture (USDA) and Economic Research Services (ERS) (2005) also noted that in non-metropolitan areas, only 16.6 percent of the people living in male-headed, single-adult families were poor; the poverty rate for female-headed families was as high as 37.1 percent. The high rate of poverty among female-headed families in these areas was attributed to lower labor force participation rates, shorter average workweek and lower earnings (USDA & ERS, 2005). The poverty rate was also highest in the completely rural counties (not adjacent to metro counties), with 16.8 percent of the population poor. The poverty rate in the largest metro areas was the lowest, with 11.5 percent of the population poor. Persistent poverty and degree of rurality are also linked. Nearly 28 percent of the people living in completely rural counties live in persistent poverty counties. In contrast, 7.5 percent of the people living in the most urban non-metro areas live in persistent poverty counties. A study by Snyder et al. (2006) on household composition and poverty among female headed households noted that the highest poverty rates among female-headed households occur among African American, Hispanic, and Native American, and among those living in central cities and nonmetropolitan areas. The study therefore concluded that these differentials highlight not only the importance role of race or ethnicity and residence for economic but also general well-being outcomes (Synder, McLaughlin & Findeis, 2006).

In addition, the substantial differences between metropolitan and non-metropolitan areas are not only socioeconomically linked but also area deprivation and low socioeconomic status have been shown to be powerful determinants of cancer mortality, incidence, and patient survival. For example examination of the rural-urban trends and pattern in cervical cancer mortality between 1950 to 2007 revealed that in 2007, the age-adjusted cervical cancer mortality rate for women in non-metropolitan areas was 2.9 deaths per 100,000 population, 22 percent higher than the rate of 2.3 deaths for those in metropolitan areas (Singh, 2011). Similarly, within counties with a poverty rate greater than 20 percent, the age-adjusted cervical cancer mortality rate for white women in non-metropolitan areas during 1999–2007 was 3.3 per 100,000 population, nine percent higher than the rate of 3.0 for white women in metropolitan areas. In high-poverty counties, non-metropolitan black women had 15 percent higher cervical cancer mortality than metropolitan black women. Additionally, within counties with a poverty rate less than 10 percent, the age-adjusted cervical cancer mortality rate for white women in non-metropolitan areas during 1999–2007 was 2.2 per 100,000 population, 16 percent higher than the rate of 1.9 for white women in metropolitan areas (Singh, 2011). Regarding breast cancer Greenlee and Howe (2009) reported that the largest jump in the proportion of distant stage diagnosis occurred often when going from counties with 20-29 percent below poverty to the highest level, 30-45 percent below poverty.

Investigating access to health care and colorectal cancer in Kentucky, Katirai (2012) reported that geographic access was a factor that was found to be significant for men but not women. Men who lived greater than 10 miles away from a health care facility had odds approximately 21 percent larger of being diagnosed at a late stage for colorectal cancer than otherwise similar men living closer to a health care facility. Travel time has also been associated with lesser quality treatment for depression (Fortney, Rost, Zhang & Warren, 1999). Distance also affects preventive care; due to the inconveniences of travel, rural residents may choose not to seek preventive treatment (Slifkin, 2002). The long distance travel inconvenience may also compound the financial barrier (Blazer, Landerman & Fillenbaum, 1995). While there is great concern regarding access to primary care services in rural areas, considering the higher incidence of chronic disease, access to specialty physician services is an equally pressing issue. Rural residents report fewer annual visits to health care providers than those in urban communities, even though they may report that they have a health care provider (Larson & Fleishman, 2003). In Healthy People 2010 by the U.S. Department of Health and Human Services (USDHHS, 2000), it is observed that heart disease, cancer, and diabetes rates for rural areas exceed those in urban areas. These findings have reinforced the recognition that geographical location and socioeconomic deprivation play an important role in health status especially since cancer stage is known to have a strong determinant on patient's survivability.

Defining and Measuring Rurality in America

According to the 2010 U.S. Census Bureau report, only 16 percent of Americans live in rural, passing the previous low of 20 percent in 2000 (Census Bureau, 2010). However, the issue of defining what constitutes rural or urban America is complex due to the numerous and conflicting definitions of rural. As noted by Brown and Schafft (2011), the word rural is ambiguous - there is no consensus among researchers and policy makers about how to define and classify “rural” and “urban”. Even among social scientists there is a great disagreement on the meaning and exact definition of the two. Research shows that there are over two dozen definitions that are currently in use by various federal agencies, let alone those employed by researchers, organizations, and local governments. The use of various definitions reflects the multidimensionality of these concepts – the defining criteria can be population size, population density, administrative boundaries, proximity to urban settings, and economic activities. In addition, researchers and policy makers face several challenges when defining or classifying rural and urban, such as defining thresholds and building blocks (geographic unit), and data availability (Flora & Flora, 2008; Waldorf, 2007; Isserman, 2005).

The most commonly used federal definitions are those by the Census Bureau, the Office of Management and Budget (OMB), and the Economic Research Service (ERS) of the United States Department of Agriculture (USDA). Appendix A provides summary on the various definitions.

Defining Spatial Geographic Isolation

Given that geographic access is an essential determining factor of a patient’s treatment seeking behavior, it is important to study and develop measures of spatial availability and accessibility of health care facilities for rural areas. Nonetheless, conceptualizing spatial geographic isolation is a very important but also a complex matter due to the many definitions of rurality. Depending on how rural regions are designated, research may produce varied results (Hewitt, 1989). The classifications of rurality apply different criteria, geographic units of analysis, and methodologies to designate rural areas. The classification of rural and urban has for years been characterized by debates on how to define rurality. Some places are rural or non-metropolitan under one definition, but not under others. Rural has often been considered as being “not urban” or “not metropolitan”. For example the rurality definitions of the Bureau of Census and the Office of Management and Budget (OMB), which are the most commonly used ones, are derived by exclusion, i.e., whatever areas not classified as urban or metropolitan are considered to be rural.

As noted earlier, the perception of rurality is multidimensional and its characterization is attached to particular objectives and views. Rural areas have been defined as particular types of regions and communities according to some objective measures, such as population density, commuting patterns, poverty or unemployment rates, or extent of wild areas and farmland (Beedasy et al., 2008). There is no one standard definition of rural that can satisfy all stakeholders or their goals. It is difficult to arrive at a single definition, as the classification has to suit different purposes. Nevertheless, a need exists to arrive at adequate definitions of rural that capture the diverse characteristics of rurality. Even though the concept of rurality is diverse, funding agencies and organizations have to make rural and urban delineations to administer policies and programs, to target resources to rural areas, to adjust Medicare and Medicaid health care reimbursement levels, or to establish eligibility for rural grant programs. There are several different types of spatial classification schemes which are described below.

Classification Scheme I: Urban Areas as Defined by the U.S. Census Bureau

The Census Bureau defines an urban area as a continuously built up territory with a total population of 2,500 or more, that is comprised of census block groups and blocks with a population density of at least 1,000 persons per square mile and surrounding blocks with an overall density of at least 500 people per square mile. All territory outside urban areas is defined as rural. Two types of urban areas are distinguished: urbanized areas and urban clusters (Figure 2.1).

- An urbanized area has at least 50,000 residents.
- An urban cluster has at least 2,500 residents but fewer than 50,000 residents.

All territory outside of urban areas is defined as rural. All persons residing in an urban area are referred to as urban residents. All persons residing outside an urban area are referred to as rural (Isserman, 2005; Waldorf, 2007).

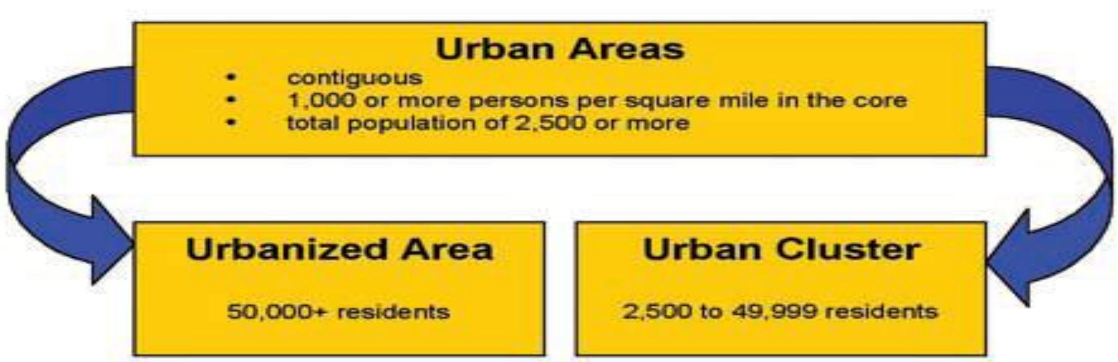


Figure 2.1. Definition of Urban Areas

Classification Scheme II: Core Based Statistical Area as Defined by OMB

The OMB group counties into metropolitan and non-metropolitan (Figure 2.2) (a new micropolitian system was added in 2003) based on population size in an urbanized area and outlying counties, and commuting patterns between them. The purpose of this classification is to "to provide nationally consistent definitions for collecting, tabulating, and publishing Federal statistics" – known as Core Based Statistical Areas (CBSA) (OMB 2000, 82228), hence does not equate to a rural-urban (Waldorf, 2007; Isserman, 2005).

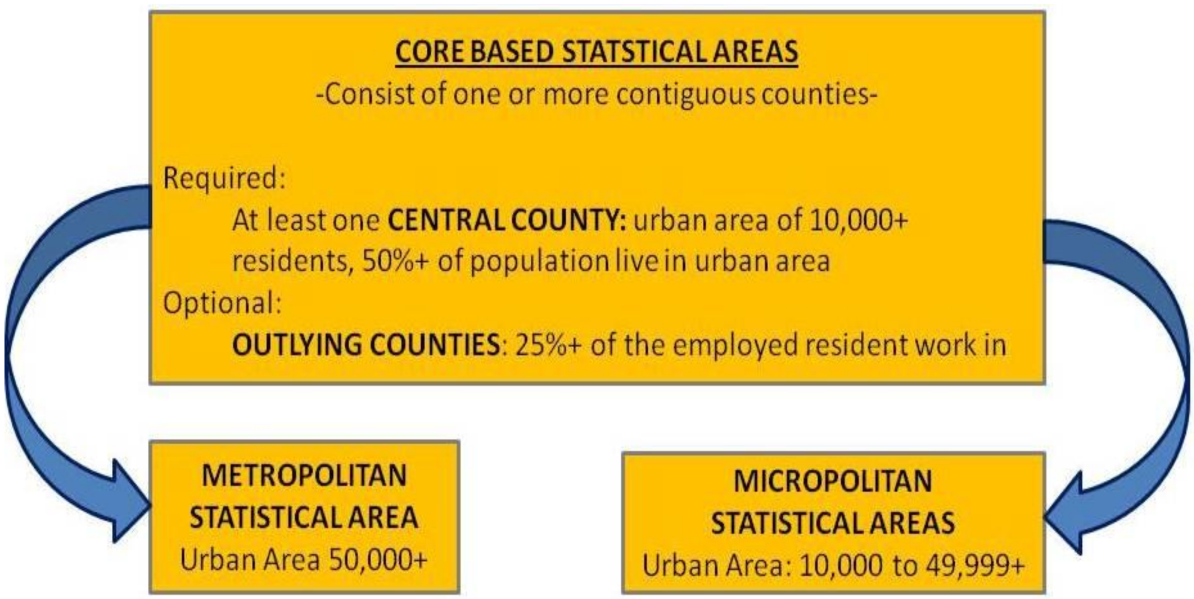


Figure 2.2. Definition of Core Based Statistical Areas

Classification Scheme III: The Rural-Urban Continuum Code (RUCC) as Defined by USDA/ERS (Beale Codes)

The ERS of the USDA probably has the most extensive definitions of rural. Some of the popular classification schemes are the Rural-Urban Continuum Code (RUCC), the Urban Influence Code (UI), and the Rural Urban Commuting Areas (RUCA). The RUCC and UI define rural and urban along county lines, while the RUCA uses the census tract as the building block for more precise information at a finer geographical scale. These classifications define counties or census tracts by size and their degree of urbanization or proximity to metro areas (Appendix B). The RUCC allocates counties to nine categories. It does so in three steps (Figure 2.3) (Waldorf, 2007).

- First step: Counties are distinguished by whether or not they belong to a metropolitan statistical area (MSA).

Second step:

- Metropolitan counties are further differentiated into three groups using the size of the MSA to which they belong as the distinguishing criterion;
- Non-metropolitan counties are further differentiated into six groups using the size of their urban population and adjacency to a metropolitan area as the distinguishing criteria.
- Third step: Numerical values (from 1 to 9) are assigned to the nine categories, with categories 1 to 3 representing metropolitan counties, and categories 4 to 9 representing non-metropolitan counties

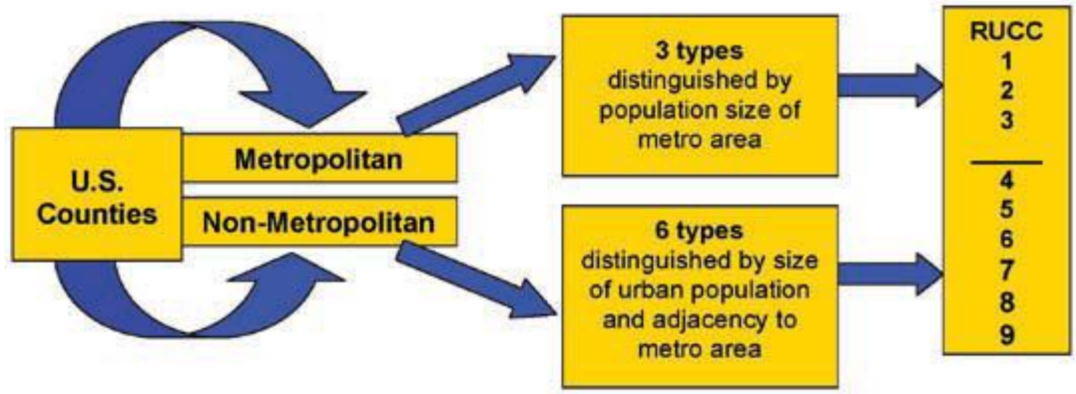


Figure 2.3. Categorization of U.S. Counties by the Rural-Urban Continuum Code

Depending on the definition, the shares of U.S. rural population and its socioeconomic characteristics vary substantially. The need for a clear definition to produce accurate research conclusions and efficient and well-targeted government programs has encouraged researchers to create more detailed and precise definitions that go beyond the metro/non-metro dichotomy and overcome the “county trap.” Isserman’s (2005) rural-urban density typology and Waldorf’s (2006) index of relative rurality are two illustrative examples.

Classification Scheme IV: The Rural-Urban Density Typology as Defined by Isserman (2005)

The rural-urban density typology was coined by Andrew Isserman (2005) as an alternative classification system. The goal of this classification is to help accurately distinguish between rural and urban within the constraint of countries that blend urban and rural. The “Rural-Urban Density Typology”, group counties into 4 areas: Rural, urban, mixed rural and mixed urban using these four criteria (Table 2.1).

- Percentage of urban residents

- Total number of urban residents
- Population density
- Population size of the county's largest urban area

Table 2.1

The Rural-Urban Density Typology

	Population Density (person per square mile)	% Urban	Population Size of Largest Area	Total Number of Urban Residents
Rural	<500	<10%	<10,000	
Urban	500+	90% +		50,000+
Counties meeting neither the rural nor the urban criteria are classified as mixed. A population density criterion is used to differentiate between 'mixed rural and 'mixed urban'.				
Mixed				
Mixed Rural	< 320			
Mixed Urban	320+			

Classification Scheme V: The Index of Relative Rurality as Defined by Waldorf (2007)

Waldorf (2007) believed that the "threshold trap" identified by Isserman creates artificial similarities and artificial separations. Therefore to address this problem, she proposed an alternative measure, called the "Index of Relative Rurality" (IRR). The index takes several dimensions of rurality into account and measures the degree of rurality on a scale from 0 to 1, with "0" indicating extremely low rurality and "1" indicating extremely high rurality. Specifically, the index simultaneously incorporates four dimensions of rurality:

- Population size: other things being equal, a county with a larger population size is considered less rural than a county with a smaller population size;
- Population density: other things being equal, a county with a higher population density is considered less rural than a county with a lower population density;
- Percentage of urban residents: other things being equal, a county with a higher percentage of urban residents (as defined by the U.S. Census Bureau) are considered less rural than a county with a lower percentage of urban residents;
- Distance to metropolitan areas: other things being equal, a county in close proximity to a metropolitan area is considered less rural than a remote county far away from a metropolitan area.

These four dimensions are expressed on compatible scales and subsequently linked so that a score of 0 is assigned to the least rural (most urban) county and a score of 1 is assigned to the most rural county.

Proposed Appropriate Rural Measurement

From the various categorizations on rural-urban, it is clear that rurality is much more complex than many people think. Throughout America, rural counties differ not only in terms of population, density and proximity to urban city but also culturally. Consequently, these factors will also have great impact on access to health care services as well as the wellbeing of individuals. Therefore taking all the definitions and classifications on rural-urban into account, the Beale codes or the rural-urban continuum codes that is an extension of the OMB classification was applied in this research. This classification scheme distinguishes metropolitan (metro) counties by the population size of their metro area, and nonmetropolitan (nonmetro) counties by degree of urbanization and adjacency to a metro area or areas. The metro and nonmetro categories have been subdivided into three metro and six nonmetro groupings, resulting in a nine-part county codification. Further, the codes allow researchers working with county data to break such data into finer residential groups beyond a simple metro-nonmetro dichotomy, particularly for the analysis of trends in nonmetro areas that may be related to degree of rurality and metro proximity. Lastly, because the Missouri Cancer Registry and Research Center uses this classification to identify rural areas in Missouri, in order to be able to assess effectively the problem regarding access and distance travel, this method appeared most appropriate. Appendix B was used as a guide in classifying all counties in the state.

Objective of the Study

The overall aims of this research were:

- Identify counties with high rates of breast cancer in Missouri;
- Assess the impact of access and distance travel to health care facilities on diagnosis and treatment of breast cancer in Missouri;
- Contrast the difference in cancer diagnosis in metropolitan; and nonmetropolitan in Missouri using the RUCC classifications
- Propose recommendation based on findings.

Summary

Access to cancer preventive services like mammography is currently the most effective method of detecting early breast cancer and reducing breast cancer mortality.

Yet the most recent guidelines from the U.S. Preventive Services Task Force (USPSTF) recommend that women undergo biennial mammography screening beginning at age 50. Decisions to start screening at an earlier age should be made on an individual basis Earlier USPSTF recommendations and the current recommendations of the American Cancer Society, the American College of Radiology, and the Society of Breast Imaging recommend annual mammography screening every 1 to 2 years beginning at age 40. Studies have also shown that despite increases in mammography use over the past two decades, population-based surveys have consistently demonstrated that a substantial proportion of women were not up-to-date on screening (Peipins et al., 2011; Lee et al., 2010; Smith, Cokkinides & Brawley, 2009; USPSTF, 2002).

Factors associated with mammography utilization have been explored in a large number of studies and reviews that have focused on characteristics related to socioeconomic status and health systems that may be barriers to or facilitators of screening. Among the often cited factors are income, insurance status, usual source of care, out-of-pocket expenses, client reminders, and recommendations for screening by health care providers (Campbell et al., 2009; Liu, 2005; Schuler et al., 2008). Access to care has also been described in terms of number of services available and transportation to those services. Mammography capacity, or the availability of machines, shows considerable geographic variability at the county level and has been shown to be an important factor in mammography usage and in late stage breast cancer diagnosis. Geographic accessibility is also commonly measured as distance to services. It is intuitively apparent that more sparsely populated locations may be at a spatial disadvantage with respect to access to medical care; and geographical distance as a barrier to breast cancer screening and treatment has been described for several rural areas. In contrast with rural areas, distances to facilities in urban areas are shorter and multiple means of transportation are often available for residents. Spatial accessibility in urban areas can nonetheless pose a challenge; especially for historically disadvantaged populations that are more likely to depend on public transportation.

Finally, economic research has demonstrated a spatial mismatch between dispersed urban employment opportunities and residential locations that is exacerbated by public transportation systems that fail to connect these areas (Campbell et al., 2000; Wang et al., 2008; Jordon, Roderick, Martin & Barnett 2004; Cromley & Cromley 2009; Peters et al., 2008). The purpose of this study was to assess the impact of spatial access to health care facilities on incidence of late stage female breast cancer diagnosis in Missouri.

CHAPTER THREE
DATA AND METHODS

The purpose of this research was to examine the role of spatial access to health care services on the probability of late detection of female breast cancer diagnosis in Missouri taking into account available clinics and hospitals. The primary interest was the relationship between spatial (geographic) isolation, distance to health care facilities and stage at breast cancer diagnosis. The stage or size of a breast tumor and how far it has spread are some of the most important factors in predicting the prognosis of a woman with this disease. Therefore, this study used geographic information system (GIS), spatial analyst functions and logistic regression methods to analyze county-level incidence of female breast cancer in Missouri from 2003 to 2008 taking into consideration place of residence and access to health care.

Research Questions

There are two central research questions in this study. The first was to what extent does spatial geographic access to diagnostic facilities have on the stage at which breast cancer is diagnosed? This question assumed that other factors that tend to inhibit access to early diagnosis, such as race and poverty (SES) etc., are confounded with spatial isolation, especially in the case of remote rural regions.

The second question was to what extent are the effects of other social factors such as race, age and poverty associated with later diagnosis of breast cancer?

Hypotheses

Two hypotheses were contemplated:

H_1 – Women with breast cancer in more remote non-metropolitan regions will, on average, be diagnosed at a more advanced stage than women in metropolitan areas over time.

H_2 – The negative effect of race, age, education and poverty on stage of breast cancer diagnosis will be increased by living in more remote non-metropolitan areas; i.e., a statistical interaction effect.

Study Design and Area

This was a retrospective observational study of female breast cancer incidence in the state of Missouri, using county as the unit of analysis. The study was approved by the University of Missouri Institutional Review Board. Descriptive design was used to describe the situation on the ground without any manipulation of variables. A GIS network analyst was used to calculate distance time travel to receive medical care.

Data Sources and Description

The following secondary datasets were used for the analysis: Missouri Cancer Registry and Research Center (MRC-ARC) cancer data, American Community Survey (ACS), TIGER® data, Environmental Systems Research Institute (ESRI) StreetMap, and Missouri health care facilities shape files.

Cancer Data

The study population, to whom we hope to extrapolate our findings, consists of all women in the state of Missouri, and even perhaps to women in all states who live with comparable education, access to health facilities, poverty and so forth. Because the study examined the extent of geographic access on breast cancer stage at diagnosis, the sample was restricted to Missouri females who have been diagnosed with breast cancer and whose case had been reported to MCR-ARC. The cancer incidence cases were provided by the MRC-ARC and covered the period of 2003 to 2008. For the purposes of this study all cancer cases were categorized into two main groups: **early** defined as *in situ* and localized stages and **late** as regional and distant stages. Cancer registry data included stage at diagnosis, age, race, county of patient's residence, and year of diagnosis. Overall, there were 29,410 cases of breast cancer diagnosed during the period under consideration. Eight hundred and seventy-four (874) cases, approximately three percent cases were excluded because either the patient was missing data on stage at diagnosis, race, place of residence or both. Two race classifications, white and black were used because these are the major racial groups in the state of Missouri. Finally, analysis was performed on 28,536 cases. The main limitation of the cancer data set is it does not contain individual patient's educational and poverty information. As a result, county-level education, and poverty characteristics were used to compute weighted average score for education, poverty and female head of households. Also, due to restrictions governing the cancer data usage, the four stages at diagnoses were combined into two. Lastly to ensure patient's rights and privacy are protected, county was used as the unit of analysis rather than block or tract groups.

American Community Survey (ACS)

This is a count-level survey which provides year to year information on all states. The most recent five year data from 2005-2009 was downloaded taking into account the following variables: Total female population by county, poverty which was calculated as the number of female below and above the federal poverty line (FPL), education, and female head of household. These county-level variables were later weighted to compute a composite score from the data as a measure of county-level educational attainment and poverty status because these data were not available in the cancer registry data. It was therefore assumed that the higher the weighted county education score the higher the educational status for that county and the higher the poverty score the poorer the county.

TIGER® Data

Topologically Integrated Geographic Encoding and Referencing (TIGER) is a county cartographic boundary files containing location in terms of latitude and longitude was downloaded from the States (U.S) Census Bureau (2010). These data were used in ArcGIS 10 to map and analyze distribution of cancer cancers in all counties in Missouri, and also distance and travel time from the centriod of each country to the nearest health care facility.

Environmental Systems Research Institute (ESRI) StreetMap

This is an enhanced street dataset that works with Esri's ArcGIS® software to provide geocoding, routing, and high-quality cartographic display for the entire United States, Canada, and Europe. StreetMap Premium works with ArcGIS Server and ArcGIS Desktop to help achieve the highest address geocoding match rates and generate the best routes and driving directions. ESRI StreeMap is specifically designed to support research, analysis and decision making for transportation issues at the national, regional, state, and local levels because it has data on all the roads and speed limits.

Health Care Facilities

These are shapefiles containing information on all the hospitals, rural health clinics, critical access hospitals and mammography centers in Missouri. These shape files was merged with patient and county data in ArcGIS 10.

Data Analysis Techniques

Independent Variables

Three main categories of independent variables were used: Demography, Economy refers to as county measure variable and Geography or Spatial Isolation. Table 3.1 shows the summary of spatial isolation definitions used in this study.

Demography

Race: Breast cancer is the most common cancer in women in the United States irrespective racial or ethnic groups. Nevertheless, the burden of cancer does not fall equally across all groups, and racial and/or ethnic disparities in diagnosis, survivorship and mortality particularly among African Americans (Warner & Gomez, 2009). Studies have shown that minority populations are more likely to live in poverty for a variety of reasons including racial discrimination, economic inequality etc. than whites (Rupasingha & Goetz, 2007; Voss et al., 2006; Crandall & Weber, 2004). For instance, African American men are 50 percent more likely than whites to be diagnosed with prostate cancer and 200 percent more likely to die of prostate cancer (ACS, 2009). White women are more likely to be diagnosed with breast cancer, though black women are more likely to die of breast cancer (ACS, 2009). According to United States Census Bureau (2010) only 11.6 percent of Missourians are blacks compare to 12.6 percent nationwide. It is therefore necessary to examine breast cancer pattern between white and black in Missouri to ascertain the differences in stage at diagnosis.

Age: Breast cancer is less common among "young" women usually anyone under 40 years of age. In the United States, about 5 percent of all breast cancer cases occur in women under age 40 (ACS, 2010; Ruddy & Partridge, 2012; Kheirelseid, Boggs & Curran et al., 2011). Breast cancer diagnosis in younger women is more difficult than in elderly women. The reason is because younger women generally have denser breast tissue than older women. As a result, by the time a lump in a younger woman's breast can be felt, the cancer often is advanced. In addition, studies have also shown that breast cancer in younger women tend to be more aggressive and less likely to respond to treatment (ACS, 2010; Ruddy & Partridge, 2012; Kheirelseid, Boggs & Curran et al., 2011). Women who are diagnosed with breast cancer at a younger age are more likely to have a mutated (altered) BRCA1 or BRCA2 gene (Komen for the Cure, 2012). Using the patient's age at diagnosis, four age groupings: 18-39 (1); 40-49 (2); 50-64 (3); and 65 and over (4) were used to assess the impact of age on incidence of female breast cancer diagnosis in Missouri.

Economy

Socioeconomic Status (SES): Socioeconomic status is known to be a powerful predictor of health and well-being (Feinstein 1993; Adler et al., 1994; Fein 1995). There are three distinct components of social determinants that have been widely reported in the literature. These include; socioeconomic determinants (e.g., age, sex, and education), psychosocial risk factors (e.g., social support, self-esteem, chronic stress, isolation) and community and societal characteristics (e.g., income inequality, social capital including civic involvement, level of trust) (Ansari, Carson, Ackland, Vaughan & Serraglio, 2003). For cancer patients, low SES is known to be associated with poor survival and increased incidence (Booth, Li, Zhang-Salomons & Mackillopo, 2010). Bradley, Given and Roberts (2002) indicated that,

> "the type of insurance a woman has also appears to play a role, in that privately insured women have, in general, a more favorable stage of disease at breast cancer diagnosis than do women who are insured through Medicare or Medicaid— who, in turn, have a more favorable stage of disease at breast cancer diagnosis than do uninsured women"

While many studies especially in the United States have found a strong association between SES and stage of cancer diagnosis (Byers, Wolf & Bauer et al., 2008; Woods, Racher & Coleman, 2006; Clegg, Reichman & Miller et al., 2009), other studies did not find any association (Wrigley, Roderick, Goerge, Smith, Mullee & Goddard, (2003); Thomson, Hole, Twelves, Brewster & Black, (2001); Brewster, Thomson, Hole, & Black, 2001). Similarly, Webster et al. (2002), Devesa and Diamond (1980), Gorey et al. (1998), Mackillop et al. (2000), Yostet al. (2001) examining community and individual level SES on breast cancer stage at diagnosis mentioned that SES, higher educational attainment and income as measured at the community level are also associated with higher incidence of breast cancer.

At the same time, the number of people considered living in poverty in America keeps rising. In 2009 the nation's official poverty rate was 14.4 percent, up from 13.2 percent in 2008 — the second statistically significant annual increase in the poverty rate since 2004 (U.S. Census Bureau, 2010). There were 43.6 million people in poverty in 2009, up from 39.8 million in 2008 — the third consecutive annual increase (U.S. Census Bureau, 2010). According to Kaiser Permanent (2010) state health facts, between 2009 and 2010, in Missouri 16 percent (770,100) white compare to 14 percent (27,512,700) white nationwide were considered as poor or with incomes less than 100 percent the FPL. On the other hand, 36 percent (247,200) black were considered to be poor in the State of Missouri, 38 percent (247,200) were living in poverty compare to 36 percent (13,378,600) nationwide. Using the ACS county-level data as a weighted measure for county educational attainment, education was coded into four groups: no high school diploma equals (1); high school graduate equals (2), some college education as (3); and bachelor and beyond equals (4). The *weighted average* formula was as follows:

$$County\ education\ score = \frac{1 * hsdp + 2 * hsdip + 3 * someco + 4 * bach}{hsdp + hsdip + someco + bach}$$

To obtain the percent of population living in poverty at each county, percent poverty for each county was also computed taking into account the number of people living below and above the FPL.

$$Percent\ poverty = \frac{popbelow}{popbelow + popabove}$$

Where: *popbelow* = number of people living below the FPL

popabove = number of people living above the FPL

Prior studies on female headed households and poverty have all concluded that there is a strong correlation between poverty and family structure type. More specifically women headed households are known to be poorer than men headed households (Snyder et al., 2006; Eggebeen, Snyder & Manning, 1996; Synder & McLaughlin, 2004). It is also well documented that women who are generally head of household tend to experience poor health throughout life as well as subsequent adverse health effects than do men (Cohen, 1994). Two formulae were used to compute percent of female headed households in Missouri.

$$pheadc = \frac{total\ female\ head\ of\ household}{county\ population}$$

Where *pheadc* represents the proportion of female headed household using the total county population

$$pheadf = \frac{total\ female\ head\ of\ household}{female\ population}$$

Where *pheadf* represents the proportion of female headed households using the total female population in each county.

Geography or Spatial Isolation

Place of Residence: Geographic location or place is important for breast cancer patients especially those who live in rural areas. These patients may be unable to obtain regular screening because they do not have access to health care within a reasonable distance (Kerlikowske et al., 1995). Urban populations generally have greater access to health services than rural populations, and this disparity in access is particularly acute when it comes to specialty health care, such as diagnostic and treatment services for cancer (Huang, Dignan, Han & Johnson, 2009; Jones, Haynes, Sauerzapf, Crawford, Zhao & Forman 2008; Chan, Hart, & Goodman, 2006; Arcury, Preisser, Gesler & Powers, 2005; Punglia, Weeks, Neville & Earle, 2006). Lack of access to health services is also likely to reduce opportunities for second opinions and personal health care choices. In addition to travel time and access, rural populations are also likely to be poor and have limited access to public transportation system than their urban counterparts (Arcury et al., 2005; Maheswaran, Pearson, Jordan & Black, 2006; Celaya et al., 20067). A study in Kentucky on distance to mammography facilities and stage at breast diagnosis, Huang, Dignan, Han and Johnson (2009) found that more rural women (15.7 percent), as opposed to urban women (4.4 percent), had to travel distances of over 15 miles to seek medical care. To effectively assess the impact of spatial isolation on breast cancer diagnosis in Missouri, the Beale Code or RUCC was used to derive a new type of rurality as depicted in Table 3.1. It was assumed that increase spatial isolation in areas that are not near metropolitan or urban town will inversely lead to higher rate of distant or late stage breast cancer.

Table 3.1.

New Rural Type Defined

Type of Rurality	County	RUCC
	FIRST CATEGORY	
Metro Large	Bates, Caldwell, Clay, Clinton, Franklin, Jackson, Jefferson, Lafayette, Lincoln, Platte, Ray, St. Charles, St. Louis, Warren, Washington, St. Louis city	1
Metro Medium	Christian, Dallas, Greene, McDonald, Polk, Webster	2
Metro Small	Andrew, Boone, Buchanan Callaway, Cole, DeKalb, Howard, Jasper, Moniteau, Newton, Osage	3
Urban Large & Adjacent	Johnson, Pettis, St. Francois	4
Urban Small & Adjacent	Cape Girardeau, Marion, Phelps, Pulaski, Scott	5
Urban Small & not Adjacent	Audrain, Barry, Barton, Carroll, Cedar, Cooper, Crawford, Douglas, Gasconade, Henry, Iron, Laclede, Lawrence, Livingston, Miller, Nodaway, Pike, Randolph, Ste. Genevieve, Saline, Taney, Wright	6
Urban very Small & not Adjacent	Adair, Butler, Camden, Dent, Dunklin, Grundy, Harrison, Howell, Linn, Macon, Madison, Mississippi, New Madrid, Pemiscot, Perry, Stoddard, Vernon	7
Rural & Adjacent	Dade, Daviess, Gentry, Hickory, Holt, Maries, Montgomery, Morgan, St. Clair, Stone	8
Rural & not Adjacent	Atchison Benton Bollinger Carter Chariton Clark Knox Lewis Mercer Monroe Oregon Ozark Putnam Ralls Reynolds Ripley Schuyler Scotland Shannon Shelby Sullivan Texas Wayne Worth	9
	SECOND CATEGORY	
Metro Large	all counties in RUCC 1	
Metro Medium	all counties in RUCC 2	
Metro Small	all counties in RUCC3	
Urban Large	all counties in RUCC 4 & 5	
Urban Small	all counties in RUCC 6 & 7	
Rural	all counties in RUCC 8 & 9	
	THIRD CATEGORY	
Metro	all counties in RUCC 1, 2 & 3	
Nonmetro Adjacent	all counties in RUCC 4, 6 & 8	
Monmetro & not Adjacent to Metro	all counties in RUCC 5, 7 & 9	
	FOURTH CATEGORY	
Metro	all counties in RUCC 1, 2 & 3	
Nonmetro	all counties in RUCC 3, 5, 6, 7, 8 & 9	

Figures 3.1 and 3.2 show the combined percentage of late stage female breast cancer diagnoses by racial and rural residential type using the 2003 RUCC. The racial breakdown of the data into black and white racial groups revealed that from 2003 to 2005 (Figure 3.1) the percentage of late stage breast cancer for blacks far exceeded that of whites in almost all the nine rural type counties. Overall, more than 50 percent of all late stage diagnoses occurred in metro medium (55.6 percent) and urban very small (57.1 percent) and adjacent counties, with completely rural and not adjacent counties accounting for almost 67 percent of all black late cases. For whites most of the diagnoses were recorded in urban small (37.1 percent), adjacent urban very small and not adjacent (35.7 percent), and completely rural and not adjacent (34 percent) counties. Between 2006 and 2008 (Figure 3.2), again, the proportion of late diagnoses among blacks were the highest. A total of 71.4 percent of all black late cases were in urban very small and adjacent to metropolitan counties while for whites, most of the diagnoses occurred in completely rural areas.

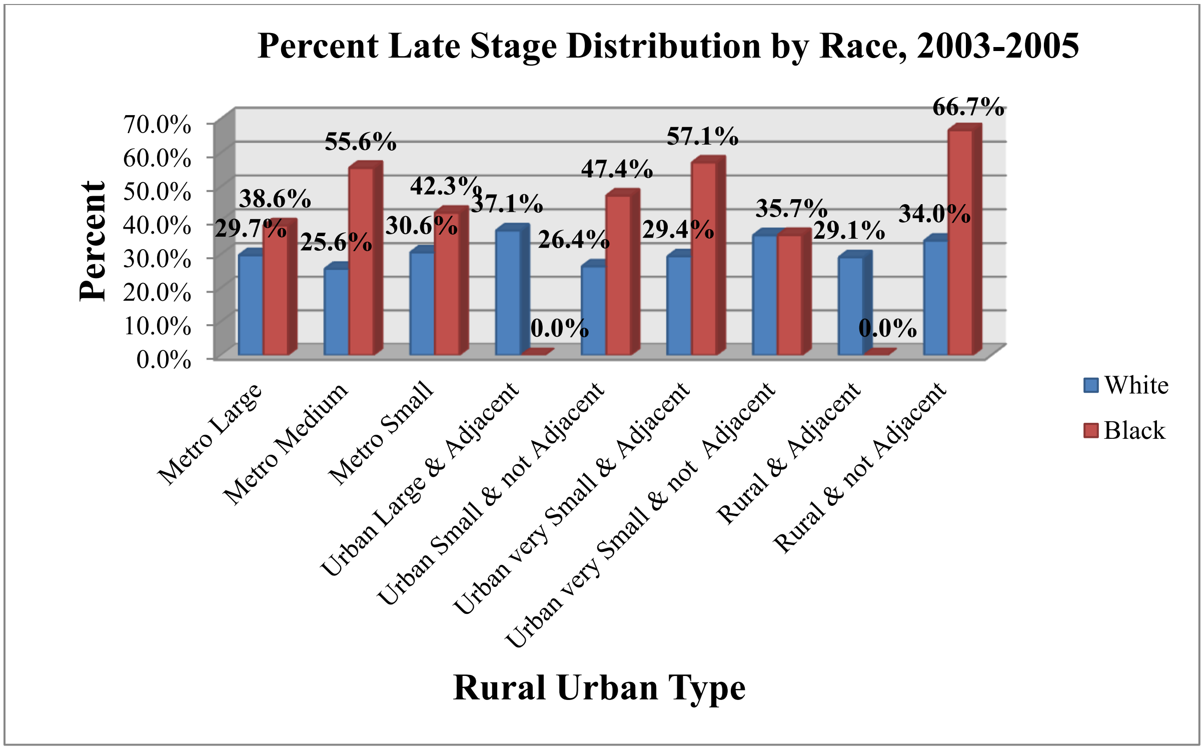


Figure 3.1. Late Stage Breast Cancer Distribution by Race, 2003-2005

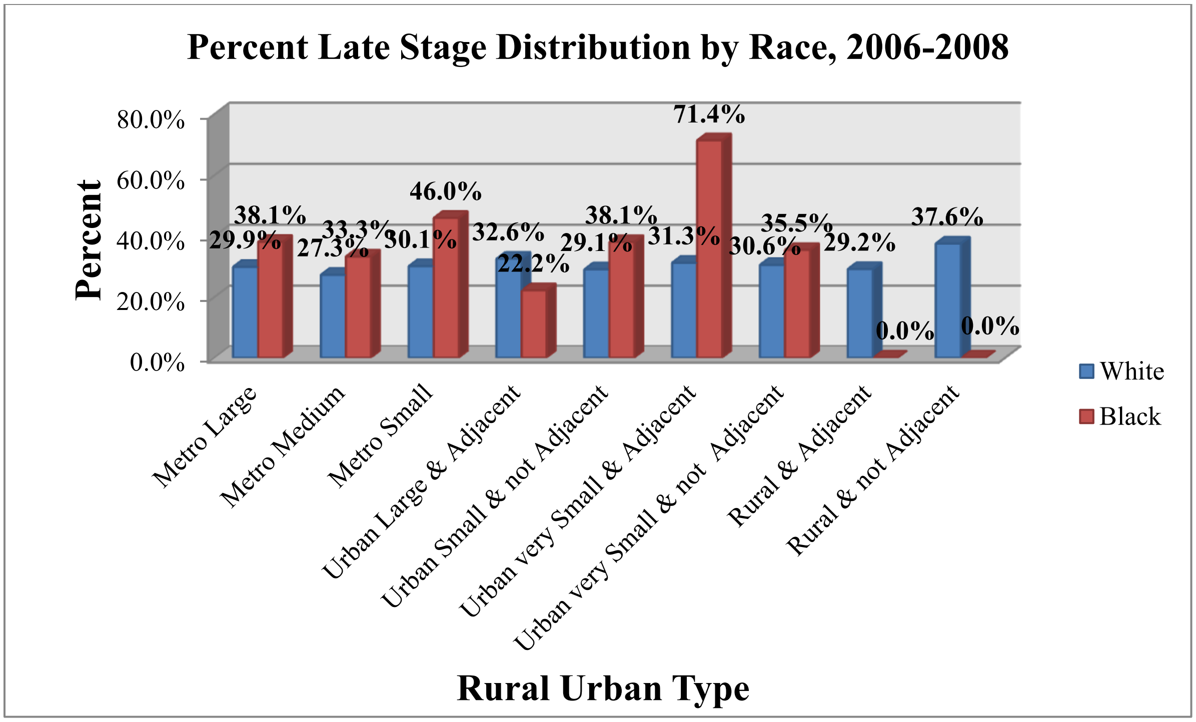


Figure 3.2. Late Stage Breast Cancer Distribution by Race, 2006-2008

Data Preparation and Software

After collecting all necessary data needed for this study, they were cleaned and then excel was used to merge the cancer data and ACS data (Figure 3.3). The merged data file was imported into ESRI ArcGIS 10. ArcGIS is an integrated geographic information system (GIS). ArcGIS helps to create and make maps. It can also be used for compiling geographic data, analyzing mapped information, sharing and discovering geographic information, and managing geographic information in a database. GIS uses shapefiles to depict shaped landmarks such as lakes or waterways. A TIGER shapefile (boundary file) of Missouri contiguous areas was downloaded from Missouri Spatial Information Service (MSDIS) website. The state shapefile was joined to the excel data file using a common field called Federal Information Processing Standards (FIPS) Codes. Missouri Plane Coordinate System was used to project all counties in the state. In GIS information is displayed as layers – a representation of different themes, such as roads, cities etc. The layer also helps to visually display information on a map.

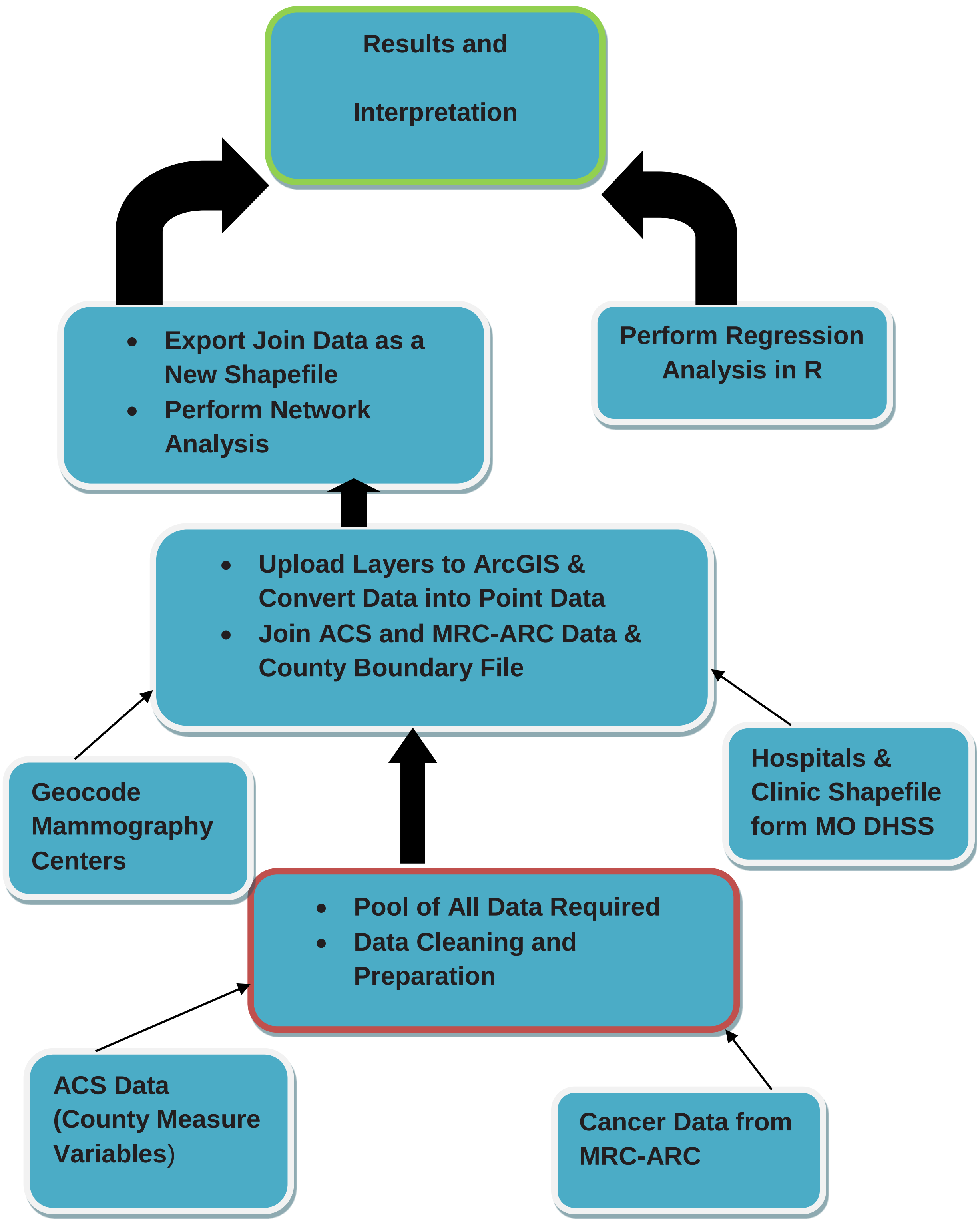


Figure 3.3. Summary of Methods Used in the Study

Unit of Analysis

The unit for analysis in this study is the individual counties in Missouri (114 plus St. Louis City). It is important to mention that county (s) as the unit of analysis poses some challenges. As noted by Lobao et al., (1999) all spatial units raise concern about containment of social processes or diffusion effects between units. Counties moreover are situated within other scales of government that influence internal relationships. However, due to limitations of the dataset, county was used taking into account the type of rurality based on the 2003 Beale Code or RUCC (Figure 3.4).

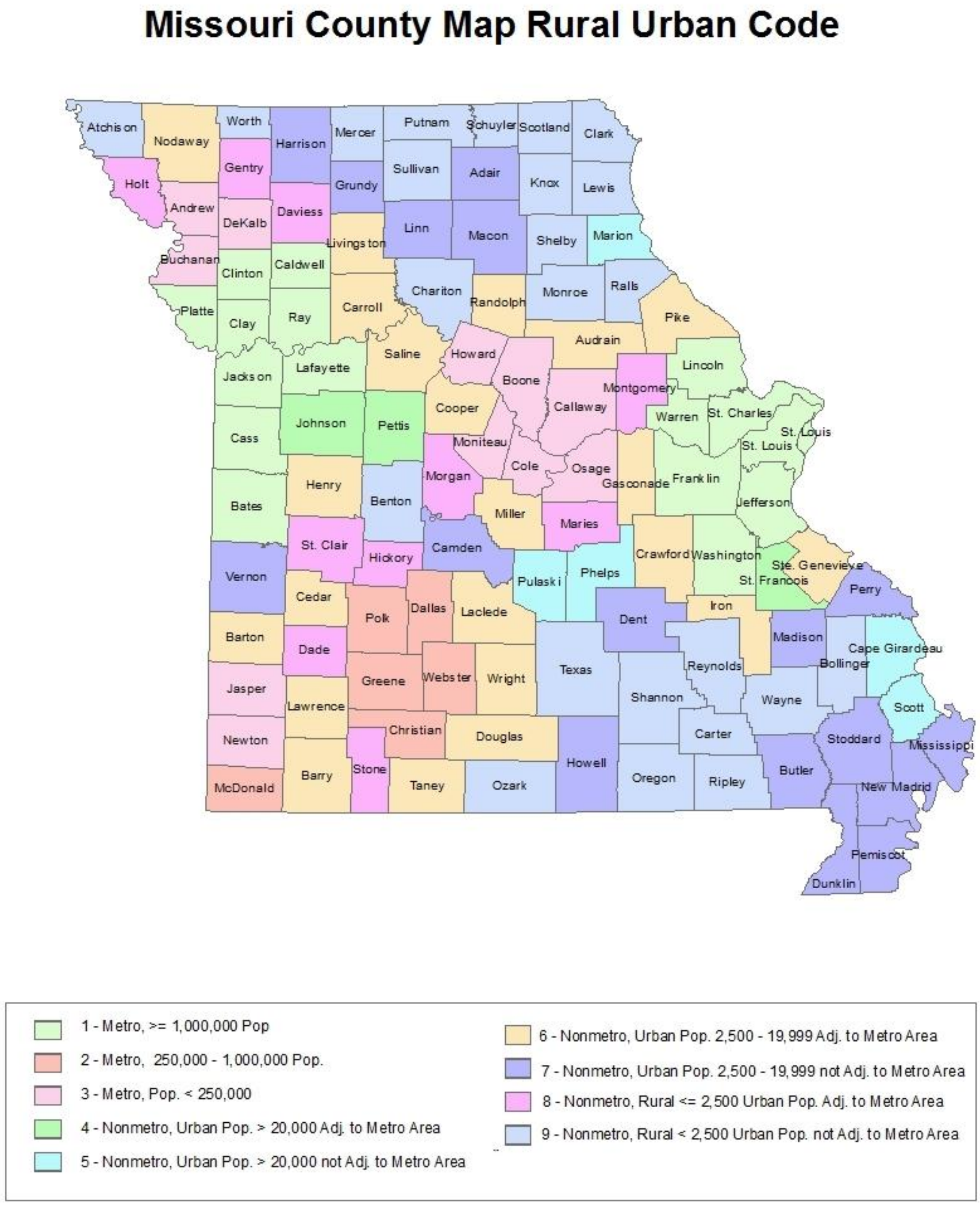


Figure 3.4. Rural-Urban Code by County

Spatial Approach

Geographical location of patients and health care services are important enabling factors for access to health care (Aday & Andersen, 1974). The analysis of this study used geoprocessing procedure to measure spatial effects by mapping various types of health care facilities such as mammography screening centers, hospitals, rural clinics, critical access hospitals and federally qualified health centers with patient's county of residence to calculate the distance and travel time from each county centroid to the nearest facility.

Geocoding is a process in GIS that matches each record in the database with the spatial database using for example TIGER. One of the benefits of geocoding is that it determines the provider or patient's Zip Code and then matches it to the relevant longitude and latitude in the database. For the purpose of this study, the next step was to geocode all health care facilities addresses in Missouri in order to be able to use it in calculating the time travel from the centroid of each county to each health care provider center. Even though Zip Codes were purposely designed for mail delivery only, and not for data analyses and mapping, it was considered useful for this study.

To assess patterns, trend and relationships as well as identify any specific geographic pattern of breast cancer diagnosis in Missouri, some spatial analysis techniques were applied to allow visualization of the distribution and also to generate chloropleth maps. The final step was to map the spatial results according to natural breaks. As noted by Mitchell (1999), natural breaks are classes based on natural groupings inherent in the data. ArcMap identifies break points by identifying class breaks that best fit each similar group values and maximize the differences between classes. In this way values within each class are likely to be more similar, and values between classes different.

Figure 3.5. Total Distribution of Breast Cancer Diagnosed in Missouri, 2003-2008

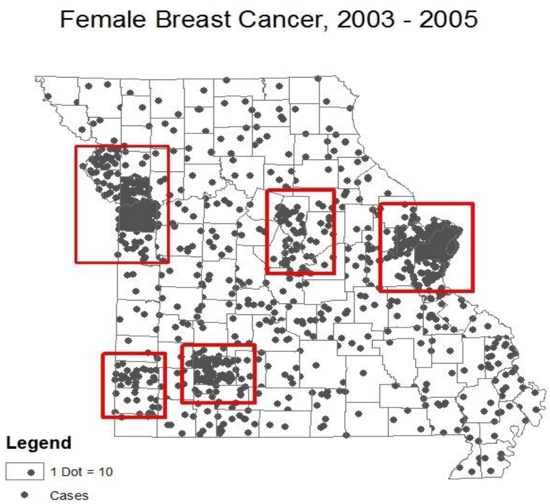


Figure 3.5A

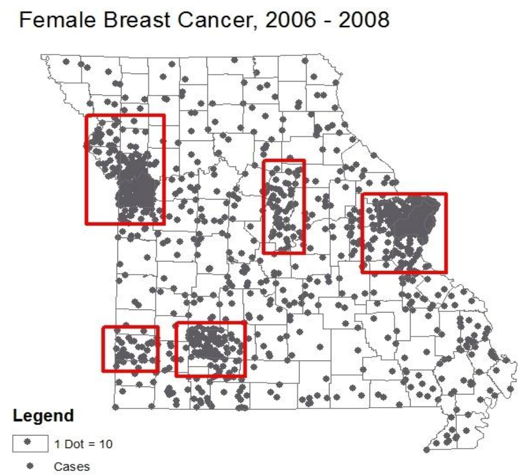


Figure 3.5B

The map labeled Figure 3.5A and 3.5B above show total number of breast cancer diagnosed among women from 2003-2005, and 2006-2008 in Missouri. The figures also exhibit considerable clustering pattern during the six year period. These were mainly around the five major metropolitan areas namely; Jaskson, Boone, St Louis County and City, Greene and Jasper Counties and their environs. However, this result in a way is misleading because we did not take into consideration the total population. Therefore to truly access the impact of breast cancer in each county, we performed geostatistical analysis based on each county population. Figures 3.6 and 3.7 depict the distribution of breast cancer rate per 100,000 female population from 2003 to 2008.

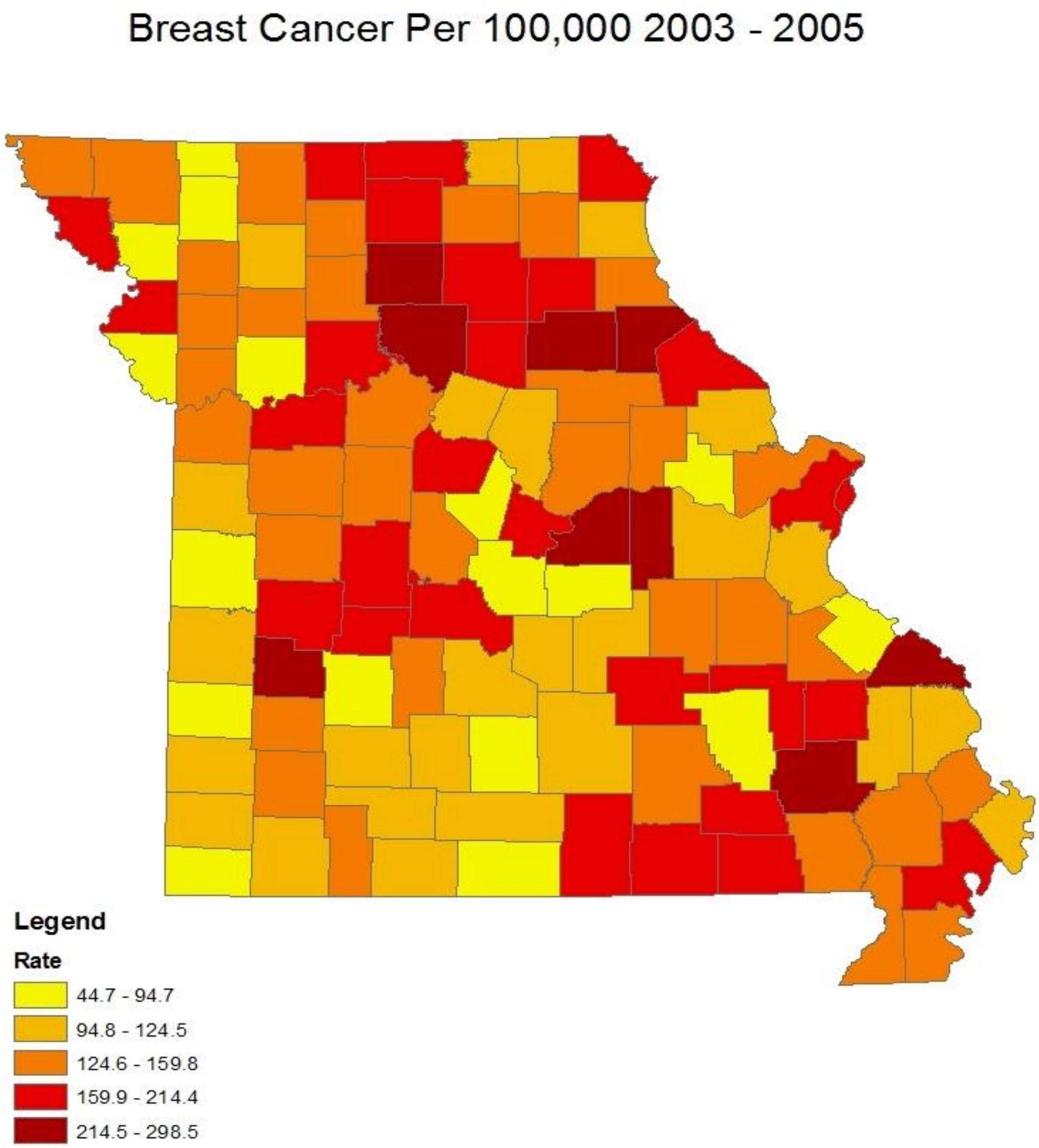


Figure 3.6. Female Breast Cancer per 100,000 Population, 2003-2005

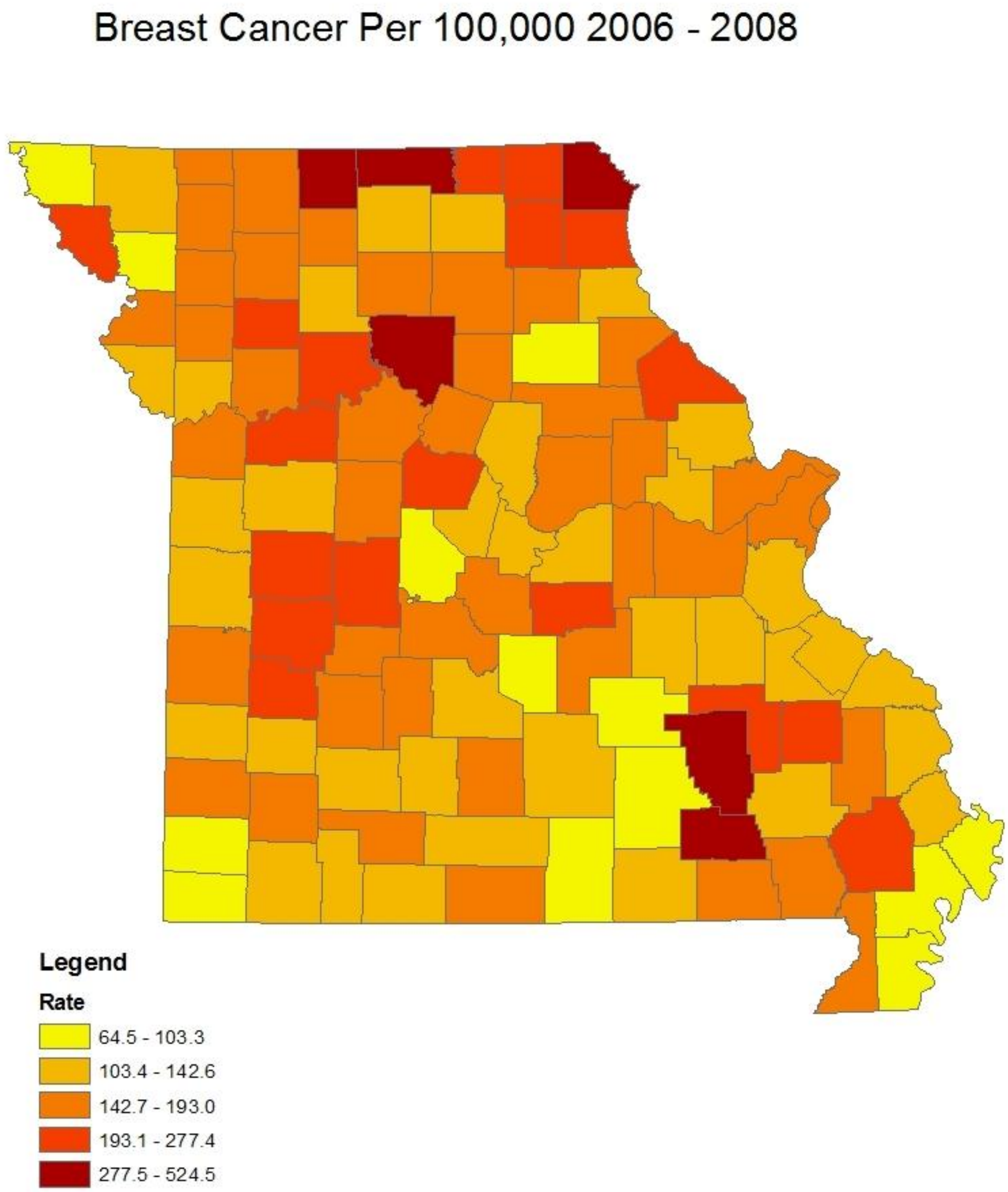


Figure 3.7. Female Breast Cancer per 100,000 Population, 2006-2008

The figures on female breast cancer per 100,000 population revealed large geographic differences over time in proportion of women diagnosed with breast cancer throughout Missouri. For instance, between 2003 and 2005, there were nine counties which recorded the highest (214.5 – 298.5 per 100,000 females) occurrence of the disease. Eight of these were all in nonmetropolitan counties; namely Linn, Chariton, Monroe, Ralls, Gasconade, Perry, Wayne and Dade Counties. The only metropolitan county was Osage in the South Central region. The same pattern of distribution was repeated between 2006 and 2008. The five counties with the highest incidence of 277.5 – 524.4 per 100,000 female population occurred in Mercer, Putnam, Clark, Chariton, Reynolds and Cater Counties. Another important observation is that, during the six year period, Chariton County which is considered a nonmetropolitan rural area was always among the counties with the highest female breast cancer rate in the state.

Network analyst travel time is considered the best measure of distance and therefore an excellent measure of access to health care services. Network analyst also provides benefit to determine the least cost network paths between a particular origin and destinations. Therefore, to determine the centroid of each county, each county feature polygon was converted to points. Second, using network analyst closest facility function distance travel time zones or service areas of less than 15 minutes, 15 to 30 minutes, 30 to 45 minutes, 40 to 60 minutes, 60 to 75 minutes and more than 75 minutes around each county centriod (origins) and hospitals (destinations) were created. This calculation took into account access to various types of health care services such as hospitals, mammography centers, rural health clinics, federally qualified health centers, and critical access hospitals in each county. Finally a thematic layer was generated and overlaid on the previous county map layer to estimate travel time travel and spatial isolation impact on stage at diagnosis.

Statistical Modeling and Hypothesis Testing

Logistic regression is a statistical model that examines the relationship between a binary outcome (dependent) variable, such as presence or absence of disease, usually denoted by $Y = 1$ if present, and $Y = 0$ if absent, and a set of predictor (explanatory or independent) variables such as patient characteristics, demographic and personal or pertinent information that may have a bearing on the response variable. These are usually denoted by $X_1, X_2, \cdots, X_k$, where k represents the number of predictors or factors. If we denote the probability of the presence of a disease by p, that is $P(Y = 1) = p$ and the probability of absence by $1 - p$, that is $P(Y = 0) = 1 - p$ then the odds are defined as the ratio of p to $1 - p$, i.e.

$$odds = \frac{probability\ of\ presence}{probability\ of\ absence} = \frac{p}{1 - p}$$

Each regression coefficient describes the size of the contribution of the corresponding predictor variable to the outcome. The effect of the predictor variables on the outcome variable is commonly measured by using the odds ratio of the predictor variable, which represents the factor by which the odds of an outcome change for a one unit change in the predictor variable. The odd ratio is estimated by taking the exponential of the coefficient e.g., (exp[β1]) (Agresti, 2002).

In this case, the response or dependent variable was the early ($Y = 0$) or late ($Y = 1$) diagnosis of breast cancer in the patient. This was called *stage* in the data. The demographic predictor variables considered were: *age, race, county of diagnosis.* The derived county variables considered were, *percent below poverty, county level education score, adjacent, county population size* and *female headed households.* Logistic regression analysis was performed using the statistical software R.

Model Assessment and Selection

An important question that arises when faced with a number of predictor variables is which variables are important and which are not? One also hopes to rank the importance of the predictor variables. This is called variable or model selection. Model selection is an essential process in quantitative data analysis. It is also necessary because, it informs us about which main effect and interaction terms to include in the analysis. Secondly, in logistic regression, the order in which variables are selected and fitted in a model is an important consideration in order to ensure that the model explains the effect or contribution of each of the predictors on the probability of early or late detection of breast cancer, and at the same time without over fitting the data. For this study, stepwise selection was used to determine the principal variables and their sequence of inclusion.

Akaike's information criterion (AIC) was derived from the concept of entropy, which is a measure of disorder of a system. AIC values provide a means for model selection. Models with smaller AIC are preferred. Using the smallest AIC as a criteria for model choice, stepwise regression was performed with *stage (DV) and age, race, poverty, education, adjacent and head of household (IVs).* The detail of stepwise logistic regression model selection is attached in appendix C.

Summary

The purpose of this chapter was to discuss the various data sources and analytical methods used. It also showed how the RUCC was used to generate a new meaningful rural type. The independent variables used in the analysis were also discussed. Distance from place of patient residence to health care facilities has been found to be highly correlated with treatment choice and survival and quality of life among breast cancer patients. Using network analyst functions in ArcGIS 10, it was indicated how time travel distance to mammography facility, hospitals, rural clinic, federally qualified health center, and critical access health center was measured using different time such as less than 15 minutes, 15 to 30 minutes, 30 to 45 minutes etc. so that the shortest and longest distance travel to the closest facility in each county can be determined and compared. This chapter also discussed how some essential county-level characteristics were used to computer average weighted scores as a proxy for individual variables. Finally, this chapter also showed how the stepwise model assessment selection was used to arrive at the “goodness of fit” primary variables and sequence for the logistic regression.

CHAPTER FOUR
ANALYSIS AND FINDINGS

Introduction

The purpose of this research was to examine the role of spatial access to health care services on the probability of late detection of female breast cancer diagnosis in Missouri taking into account the access and distance to clinics and hospitals. The derived variables of county-level education, county population size, percent below poverty and adjacency were thought to be indicators of access and distance to health care facilities. The primary interest was in the relationship between spatial (geographic) isolation, distance to health care facilities and stage at breast cancer diagnosis. This chapter focuses on the results of the analysis. The sections are presented chronologically in the following order: Characteristics of the Sample, Marginal Distributions of the Predictors, Distribution of Breast Cancer Incidence in Missouri, Distance Measure Description – Travel Time to the Nearest Medical Care Center, Logistic Regression to explain Stage at Diagnosis and Summary on the Chapter

Characteristics and Relevance of the Sample

The study sample consisted of all women in the state of Missouri, ages 18 years and older, who were diagnosed with breast cancer from 2003 to 2008. Breast cancer cases included non-invasive and invasive cancers. Although the study population has not been defined, as might commonly be expected, it is our hope that the findings and quantification will carry over to women in general, where ever they may be, who suffer disadvantages in education, poverty, and access to health care. It may seem obvious that the factors of age, race and education have a bearing on the probability of late detection. So an important aspect of this study is really the quantification and relative importance of the factors affecting the probability of late detection in the community of woman. Although there may well be variation in the magnitude of these predictors, in different situations, for example in different states, or even different countries, we do feel that there will be some uniformity of carry over effect, giving pertinence to the findings of this study to women in general and thus nullifying the spatial and temporal peculiarities of the sample data.

For the purpose of this study, the stage at diagnosis was divided into two groups. **Early** stage refers to *in situ* and localized stages and **Late-stage** include regional and distant stages. There were a total of 29,410 female breast cancers diagnosed in the MCR-ARC cancer registry file of which 874 were unknown stages or cases missing race, age, and place of diagnosis were eliminated. After the exclusions, 28,536 cases were included for analysis (Figure 4.1). These cases were further classified into four age categories, two main racial groups as white and black, and year at diagnosis (Figure 4.1).

It is generally assumed in the breast cancer literature that, the risk for developing breast cancer increases with age. In general, about 77 percent of women diagnosed with breast cancer each year are over age 50, and almost half are age 65 and older (Komen for Cure, 2012). In women 40 to 49 years of age, there is a one in 68 risk of developing breast cancer. In the 50 to 59 age group, that risk increases to one in 37 (Komen for Cure, 2012). Guided by this information, the age at diagnosis for patients was classified into four divisions: 18-39, 40-49; 50-64, and 65 and over. This classification was also used to assign an arbitrary number from 1 to 4 to represent each of the categories.

Overall, the younger age group reported the least number of cancer cases during the six year period. The distributions of diagnosed cases for each of the age group are as follows: Age 1 (1,253 – 4.4 percent), Age 2 (4,784 – 16.8 percent), Age 3 (10,247 – 35.9 percent) and Age 4 (12,252 – 42.9 percent). In effect, because age 4 – 65 years and older accounted for almost 43 percent of all total cancers in the state, the strongest conclusion for this study therefore falls on this category (Figure 4.1B).

Figure 4.1. Marginal Distribution of Individual Variables

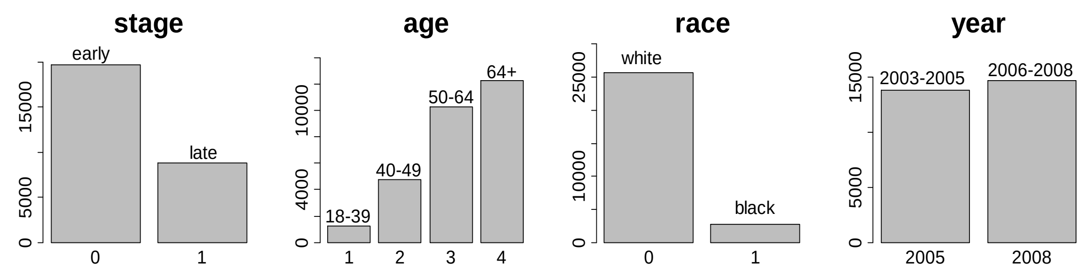


Figure 4.1A **Figure 4.1B** **Figure 4.1C** **Figure 4.1D**

Marginal Distributions of the Predictors

Individual Predictors

The bar plots above, in Figure 4.1, describe marginal distributions of the demographic variables for individuals. These marginal distributions inform us of where the bulk of the data lie and hence where the strength of the study conclusions will also lie. Specifically, the first graph on the left, entitled stage, Figure 4.1A, informs us that approximately two thirds (19,690 – 69 percent) of the individuals were diagnosed at early stage, and one third (8,846 – 31 percent) diagnosed at late-stage. This also means that between 2003 and 2008 the incidence rate for the state was 289.6 per 100,000 female population. This far exceeded Missouri Department of Health and Senior Services (DHSS) 2007/2008 Healthy People objective to reduce late stage female breast cancer to 43.9 per 100,000 population (DHSS, 2010). One of the implications of this analysis is that, the state of Missouri is still far away from achieving this goal. There is need for more education regarding the importance of breast cancer identification and preventive services in Missouri.

The graph labeled age, in Figure 4.1B, informs us of breast cancer diagnosis in the state. It is observable that majority of the cases were in the 50+ age group while only a small proportion of the cases were reported among the younger age group. The speculated cause of this is attributed to the fact that elderly have access to health insurance in the form of Medicare, which enabled them to have preventive services, while the younger groups generally do not have insurance to encourage them to go for mammography services. Secondly, the elderly group is usually known to be particular about their health care but since the younger groups are usually active and strong they may not consider it necessary to go for regular mammography examination. Thirdly, the elderly group is more aware of the possibility of breast cancer whereas the younger group tends to be oblivious of personal health care issues in general.

The graph titled race (Figure 4.1C) describes the racial distribution of cancer on white and black. Of those diagnosed with female breast cancer in Missouri, 25,743 (90.2 percent) were white while only 2,793 (9.8 percent) were black. This can be attributed to the small number of minority groups like African American in the state. Also, the model will be biased towards the white majority. Figure 4.1D, labeled year, shows an almost even distribution of breast cancer cases in the two periods from 2003-2005 (13,856) and

2006-2008 (14,680). Table 4.1 provides descriptive statistics summary on the study population. The information provided in Table 4.1 is based on each variable type hence does not total to sample size.

Table 4.1.

Distribution of Study Population Diagnoses

Factors	**Number (N=28536)**	**Percentage (%)**
	Metro Type	
Metro Area	21,106	74
Nonmetro Area	7,430	26
	Adjacent Type	
Adjacent	24,998	88
Nonadjacent	3,538	12
	Stage	
Early	19,690	69
Late	8,846	31
	Race	
White	25,743	90.2
Black	2,793	9.8
	Age	
18-39	1,263	4.4
40-49	4,784	16.8
50-64	10,247	35.9
65+	12,252	42.9

Note: Totals do not add up to 28,536 because percent is based on each factor type

County Level Predictors

Figure 4.2 provides bar plots to describe marginal distributions of the important county variables. Using the nine rural urban continuum code (RUCC) definition, Figure 4.2A inform us of the total distribution of cancer cases diagnosed between 2003 and 2008 by county population size based on the Beale Code. It is clear from the graph that overwhelming majority of the cases fell into RUCC 1 which is a large metropolitan area of more than one million population. The analysis therefore, will be biased towards code 1, and the model is good where the data are abundant and weak where the data are sparse.

Figure 4.2. Distribution of Important County Variables

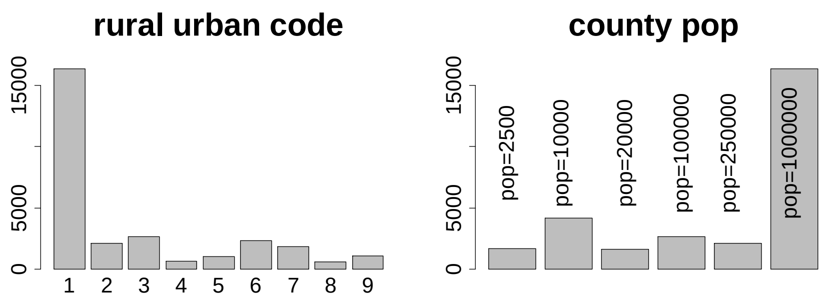


Figure 4.2A **Figure 4.2B**

To measure spatial isolation, the rural urban code was re-computed, taking into account six main groups based on the population size of each county. Figure 4.2B shows the recomputed county population distribution. This was derived using the RUCC definitions. The motivation for computing this was to disentangle size and adjacent and non-adjacent as a predictors which was not possible with the RUCC nine categories. The labels, printed sideways, of pop=2500, 10000, and so on do not represent actual counts but were extracted from the RUCC classification. These codes, defined in Table 4.2, are imprecise but are none the less ordinal.

Table 4.2.

Description of the Derived Variable "County Population"

Code	Size	Meaning
Pop=2500:	1	Included counties whose population was less than 2500
Pop=10000:	2	Includes counties from 2500 to 19999
Pop=20000:	3	Includes counties of 20000 or more
Pop=100000:	4	Includes counties fewer than 250000
Pop=250000:	5	Includes counties of 250000 to one million
Pop=1000000:	6	Includes counties exceeding one million

Considering that the overwhelming majority of diagnoses were reported in large metropolitan areas, it was decided to use the new derived county population sizes and calculate actual breast cancer percentages of late diagnosis by race. Although more than 90 percent of all cases diagnosed were whites, the distribution of late diagnosis by race demonstrates that for the black minority, the proportion of late diagnosis far exceeded that of the white majority in every location throughout the state (Figures 4.3 & 4.4).

This is both amazing and not amazing all at the same time: It is amazing that the proportion of late diagnoses in the black community so consistently exceeds that for the white community across all county sizes. Sadly, this is not so amazing because it is, intuitively, almost what one expects considering the sad history of inequality and discrimination in our community. Perhaps this also provides a research opportunity to explore this anomalous situation.

Referring to the last comparison in Figure 4.4, where the proportion of blacks having late diagnosis is zero, is consistent with earlier findings in Figure 3.2 where proportion of late stage diagnosis among black in rural adjacent and rural non-adjacent counties was also zero. We wonder if there is a problem with the data in this region or if proportion of blacks diagnosed during this period were un-staged or unknown.

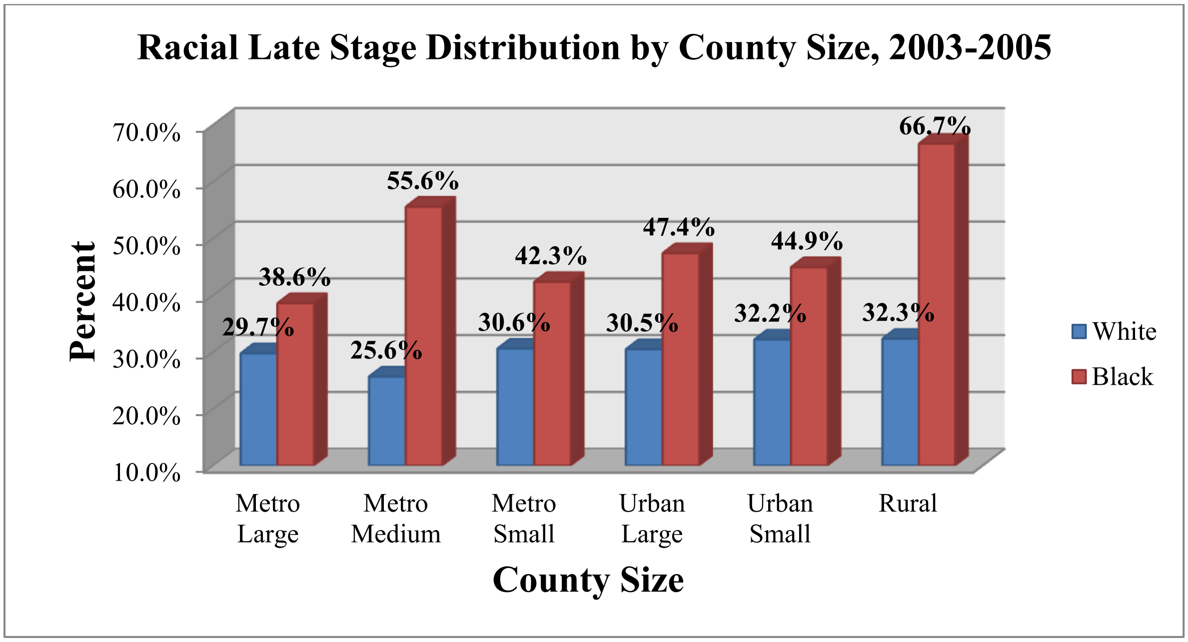


Figure 4.3. Percentage Late Stage Distribution by Race and County Size, 2003-2005

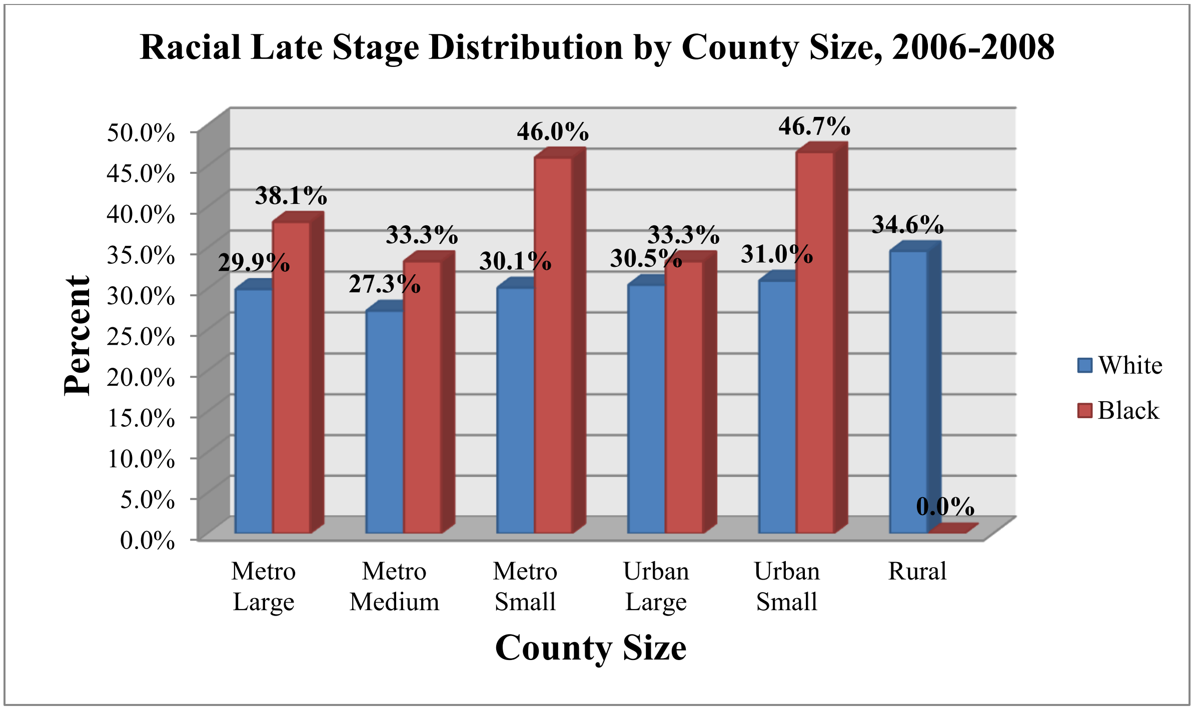


Figure 4.4. Percentage Late Distribution by Race and County Size, 2006-2008

Figure 4.5 displays the frequency distribution of female breast cancer patients diagnosed in Missouri from 2003 to 2008 by county population size – metro vs. nonmetro and counties adjacent to a metropolitan area and those that are not adjacent. Again, the distribution indicates that the majority of cancer diagnoses were found among women living in metropolitan and adjacent to metropolitan counties. Possibly, because people living in urban areas may have easy access to health care services, they are therefore being diagnosed early more frequently than those in rural areas who may not have easy access to health care services and may have to travel long distances for medical care.

Figure 4.5. Distribution of Female Breast Cancer Patients in Missouri Metro/Nonmetro and Adjacent/Nonadjacent

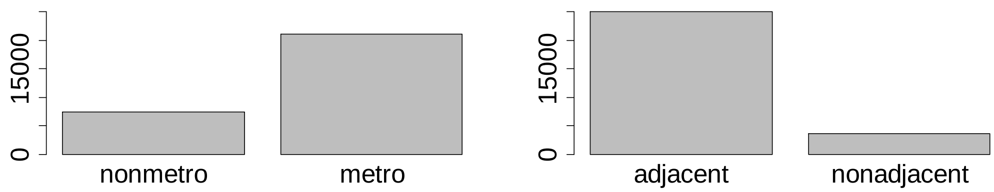


Figure 4.5A **Figure 4.5B**

In Figures 4.6 and 4.7, the percentage late-stage distribution for each race group was examined by taking into account the three broad rural-urban types. This was done to separate metropolitan counties from those that either rural or urban and adjacent to metropolitan area and those counties that are considered completely rural and not adjacent to metropolitan area. Here also, parentage late-stage was calculated by race. With the exception of 2006-2008 (Figure 4.7), where the distribution was almost evenly spread between white (32.2 percent) and black (35.2 percent) in completely rural counties, the proportion of blacks with late diagnosis significantly exceeded that for whites (Figures 4.6 & 4.7).

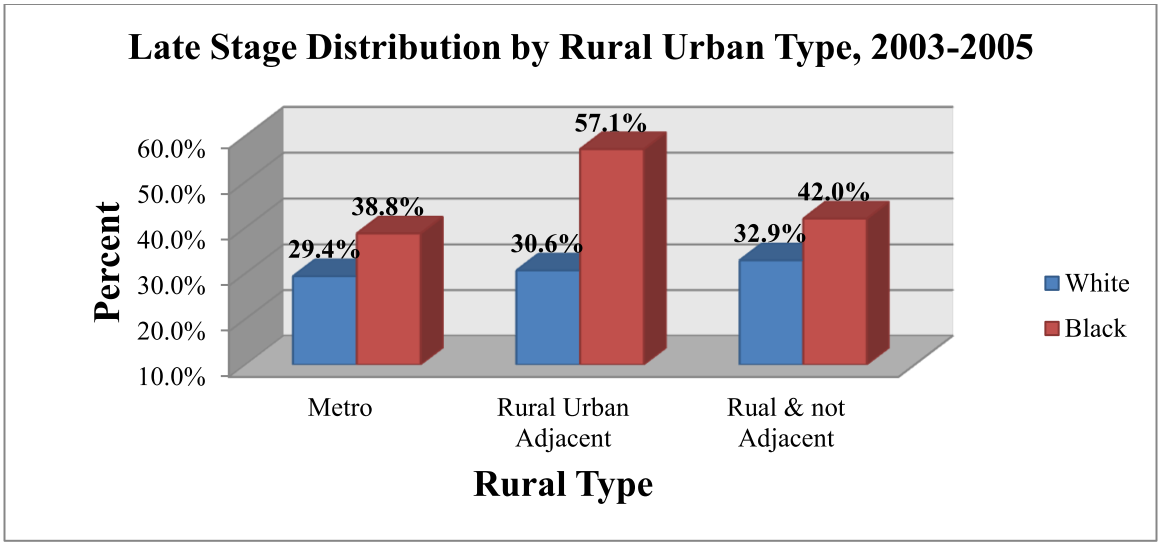


Figure 4.6. Cancer Distribution by Metro and Rural Urban Adjacent, 2003-2003

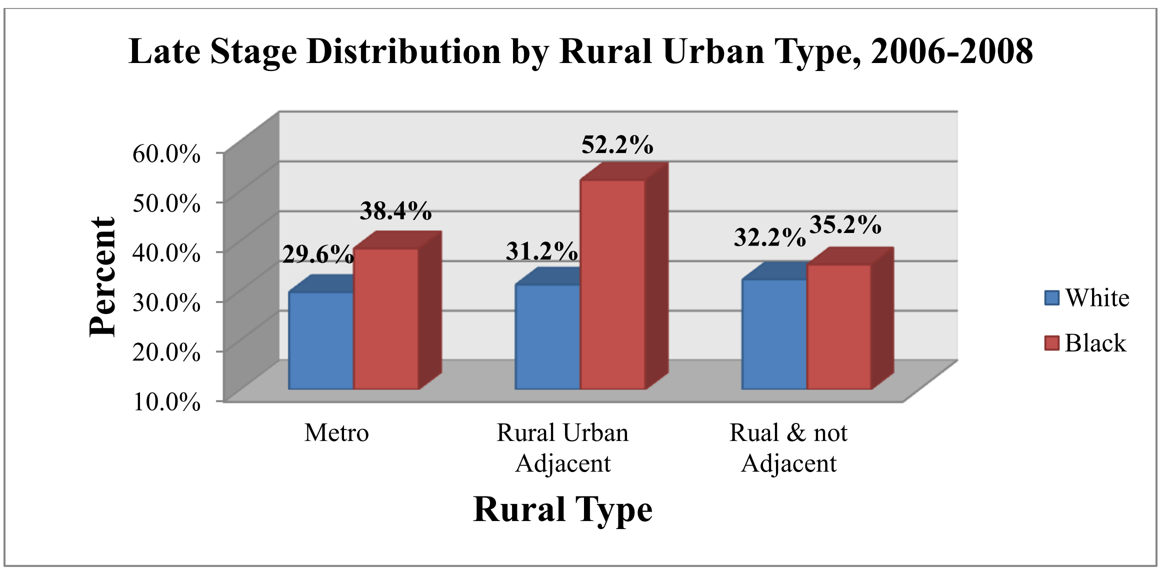


Figure 4.7. Cancer Distribution by Metro and Rural Urban Adjacent, 2006-2008

Education attainment is an important SES indicator because SES is usually measured in terms of level of education, income and occupation. Research in the social and medical fields has all indicated that low SES highly correlates to lower education, poverty and poor health (Campbell et al., 2009; Schuler et al., 2008; MacKinnon et al., 2007). The cancer data used in this study did not contain individual educational attainment or poverty information. However, in view of the importance of these variables, county education scores were computed as weighted average by assigning 1 to people with no high school diploma; 2 – high school diploma; 3 – some college degree; 4 – bachelor. This means that the higher the weighted score, the greater the proportion of educational attainment in that county.

The histogram in figure 4.8A displays a distribution, skewed to the left, of county level education. Counties with education score of two represent counties in which the majority of the population had no college education or bachelor degrees and may be considered "poorly-educated" counties. Those with a score of three represent counties in which the majority of the population had some college education or bachelor degrees, and may be considered as "well-educated" counties. The spike in the histogram, from 2.8 through 2.9 indicates that the majority of counties are well educated. By way of example, Boone County has highest educational score of 3.014 and Mississippi County has the lowest educational score of 2.016 (Figure 4.10). Figure 4.8B displays the regression of education score on size: the regression plots the mean of the conditional distributions of education for each value of size. The regression is evidence that county education attainment score increases with size of a county, but not linearly. This suggests that people in urban and metropolitan areas have higher education than those in rural areas.

Figure 4.8. Plots of Education and County Size

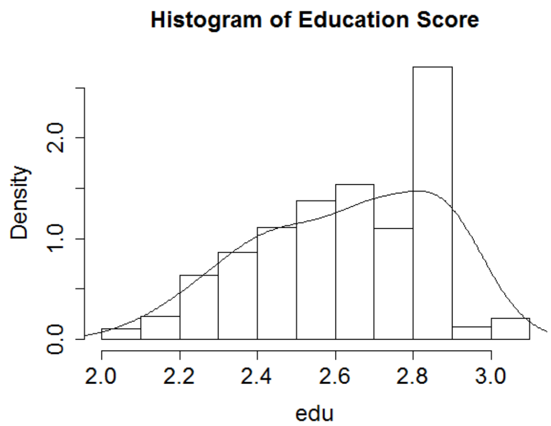


Figure 4.8A

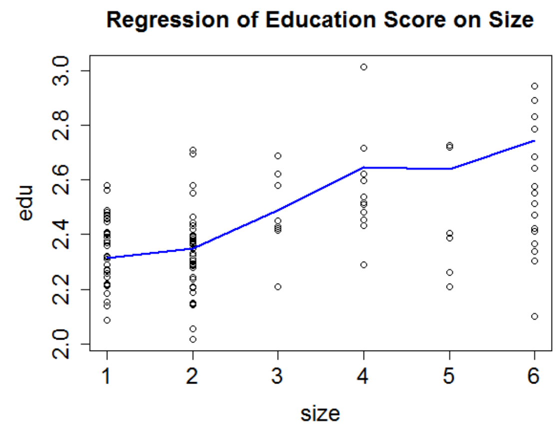
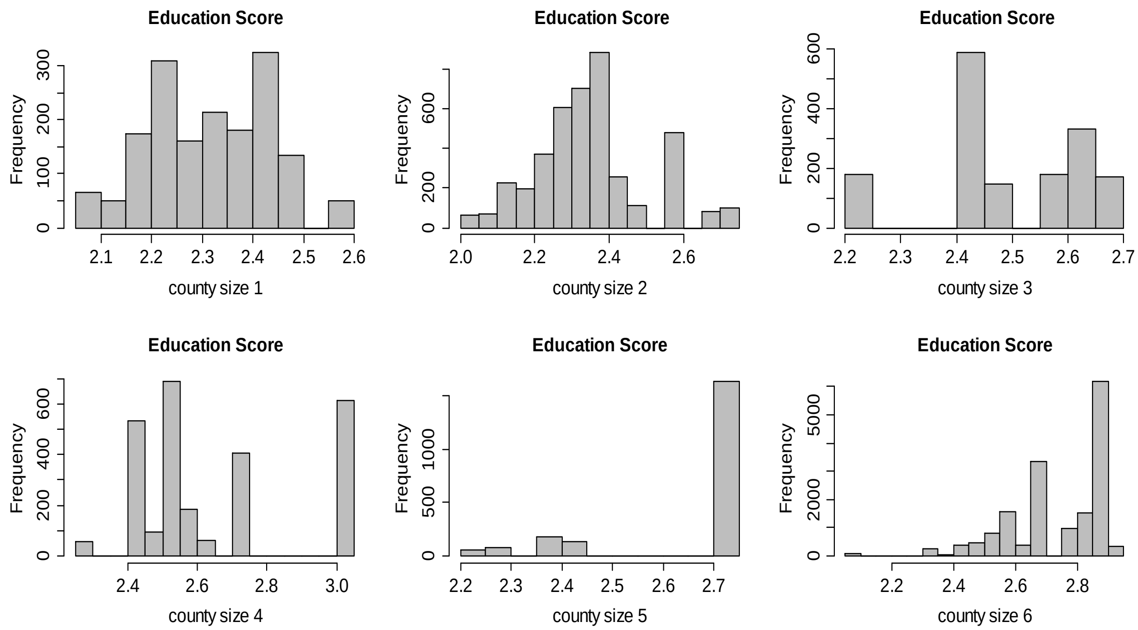


Figure 4.8B

Conditional Distribution by Derived County Size

Using the six county population sizes derived earlier and the weighted educational score for each county, the histograms in Figure 4.9 shows that the conditional distributions of education by county size are radically different in shape and location. The distribution of education in small counties is almost symmetrical and centered at about 2.3. In large counties, the distribution of education is skewed to the left and J-shaped. This is partial motivation why linear regression of education on county size is not feasible and which resulted in regression in figure 4.8B. Appendix D provides violin plot on education by county size.

Figure 4.9. The Conditional Distributions of Education by County Size (smallest to largest, 1 = small; 6 = large)



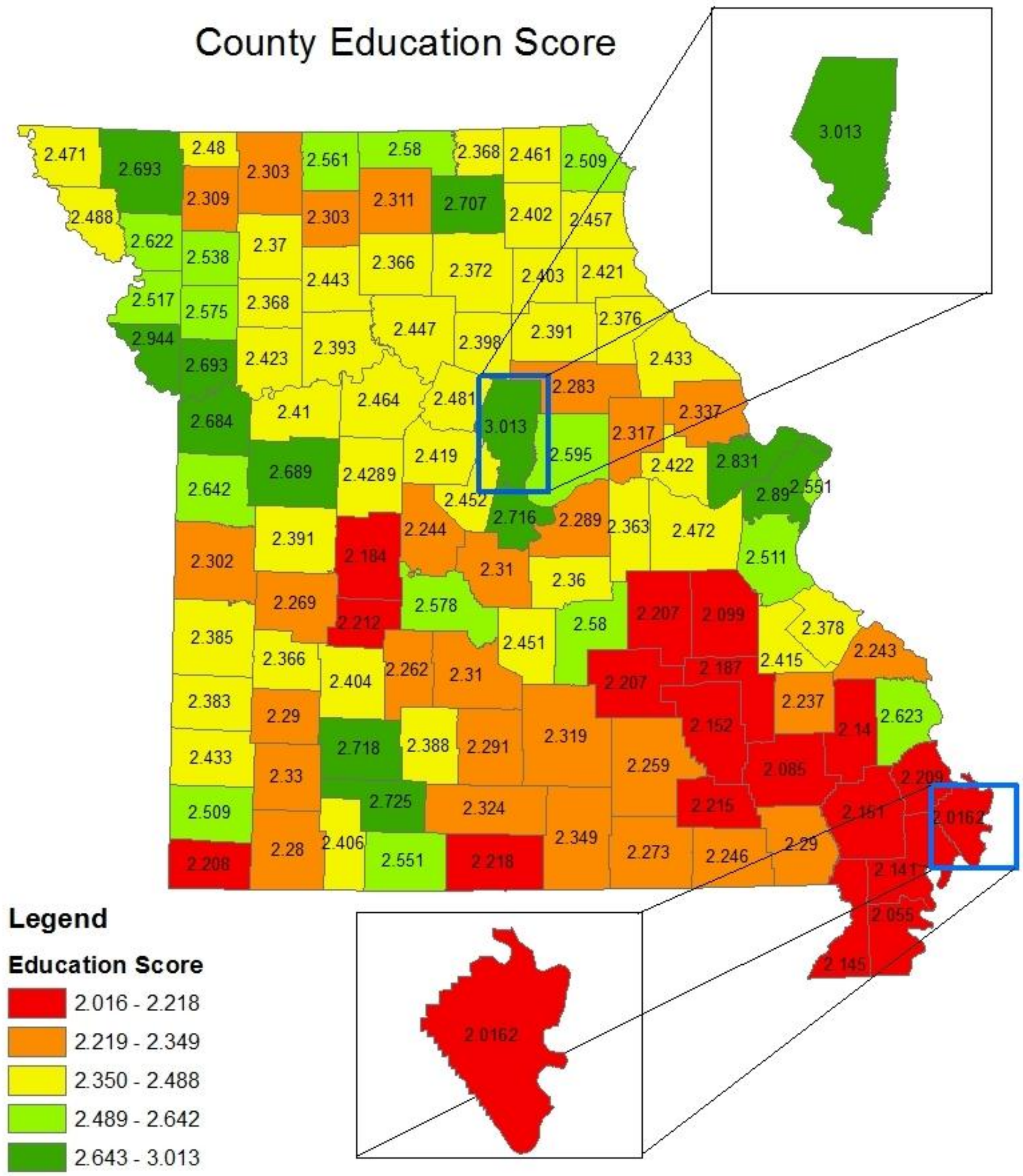


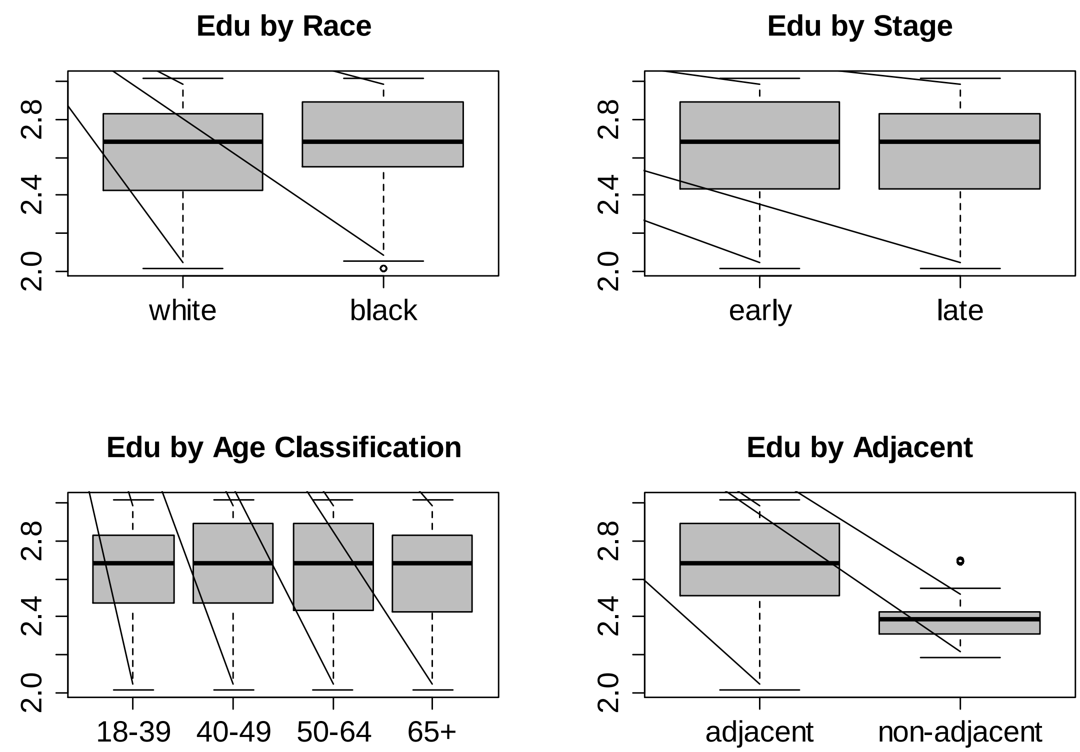
Figure 4.10. Missouri County Education Score

Boxplot of Education by Race, Stage, Age and Adjacent

Education, race, age and place of residence have been reported in previous studies to be highly correlated with cancer stage at diagnosis (Campbell et al., 2009; Schuler et al., 2008; MacKinnon et al., 2007). A boxplot is a five number summary that graphically depict the minimum, maximum, median, lower and upper quartiles. The degree of dispersion and skewness in the data is reflected in a boxplot: if a boxplot is symmetrical then the distribution is symmetrical. If the boxplot is asymmetrical then the distribution is also asymmetrical.

Boxplots were used to graphically compare level of education by race, stage age and type of metropolitan adjacency. Comparing the average education of white and black, the first graph in Figure 4.11 labeled education by race, shows an almost equal median county educational score for white (2.6) and black (2.7). On the other hand, the lower and upper quartiles differ substantially. Secondly, even though the total population of blacks in Missouri is less than12 percent, they appear to have slightly higher education scores than do whites. Similar distribution pattern were seen between education by stage and the four age classifications. While the means for all four age categories are similar, the elderly population especially 40-49 and 50-64 year old appear to have higher educational scores than do the young age group 18-39. In general, residents of metropolitan or adjacent to metro areas have higher levels of education scores than their nonmetropolitan counterparts. This could be attributed to the fact that, weighted measure rather than each individual actual educational attainment were used in the analysis.

Figure 4.11. Distribution of Education Score by Race, Age and Adjacent



The last graph in chart on Figure 4.11 displays education scores for adjacent and nonadjacent counties. Overwhelmingly, there is a significant relationship between educational attainment and geographic place of residence. Individuals living in counties closer to metropolitan areas, are approximately twice more likely to be educated than those in rural or nonadjacent metropolitan counties. In 2005, Ulubasoglu and Cardak examining comparison of rural-urban educational attainment for 1964-1999 concluded that rural educational quality is not only heavily dependent up the colonial history but also geographical characteristics such as landlockedness and a country's surface area. It is therefore not surprising that the boxplot of education by adjacent shows a strong association between higher educational attainment and type of rurality or rural status in Missouri.

The poverty score map in Figure 4.12 shows the proportion of people living below the FPL in each county. St. Charles County had the lowest poverty score (4.6 percent) and Pemiscot had the highest poverty score (31 percent) in the state. The corresponding colors on the two maps (Figures 4.10 and 4.12) suggest a negative correlation between the educational score and the poverty score: Where education scores are low, poverty is high, and where educational scores are high, poverty is low. This corroborates our intuitive understanding of the strong connection between education and poverty, and illustrates, the immerse value of education and its contribution to financial freedom, awareness and access to health care, improved self-esteem and higher expectations of quality of life. Although this may seem obvious, this is hard empirical evidence against mere speculation.

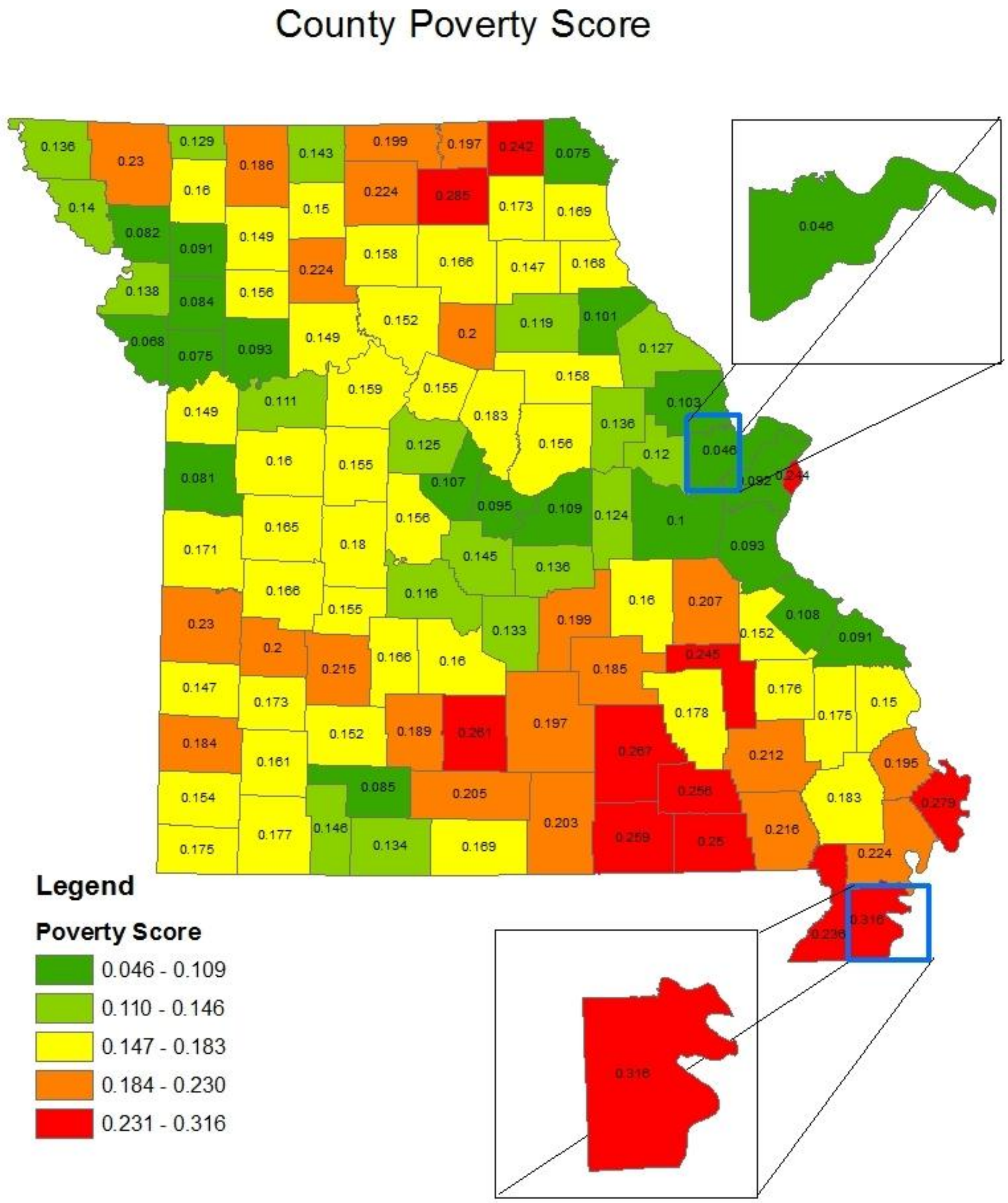


Figure 4.12. Missouri County Poverty Score

Figure 4.13 plots the proportion of blacks living below the FPL in Missouri as a function of proportion of people living below *(pbelow)* as estimated using a Kernel regression, commonly called a Kernel Smooth. The dotted line represents the average proportion of blacks living below the FPL throughout Missouri. The solid line indicates that, as *pbelow* increases above 0.20, the proportion of blacks living below poverty line increases shapely from the overall average of about 9 percent to about 30 percent. This means that, in the poorest counties, the proportion of blacks living below the FPL is more than 3 times the state average. Appendix E provides scatter on percentage below poverty, county education score and adjacency.

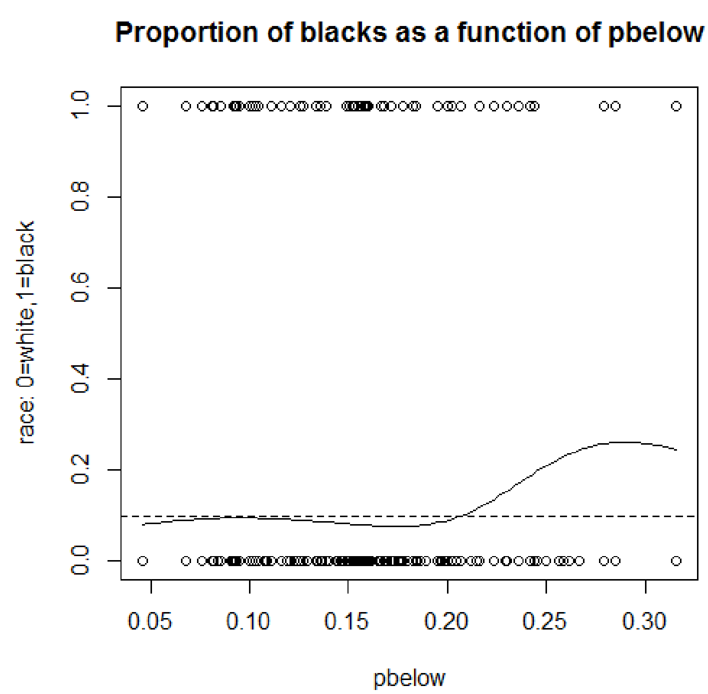


Figure 4.13. Distribution of Blacks below FDL

Distribution of Breast Cancer Incidence in Missouri

The data from MCR-ARC included 28, 536 women who were diagnosed early and late stage breast cancer between during the period of 2003 t0 2008. Percentages and rates per 100,000 female population for each year as well as both years were calculated using late stage only. The maps from Figures 4.14 to 4.19 show detailed results of proportion of women diagnosed with late stage breast cancer only.

As shown in Figure 4.14, the highest incidence of late stage female breast cancer per 100,000 in Missouri from 2003 to 2008 occurred in the following counties. Mercer 699.3, Chariton 597.0, Putnam 588.2, Cedar 500.8, Carter 499.3, Clark 497.2, Benton 485.3 and Carroll 475.6 counties. Numerous studies have indicated that geographical location is very important in the diagnosis and treatment of breast cancer. Using the RUCC two broad classification system all these counties are location in nonmetropolitan areas. Also, apart from Carroll County which is specifically located in nonmetropolitan urban area with a population between 2,500 and19,999 and adjacent to metropolitan area, the seven other counties are located in nonmetropolitan rural area with population less than 2,500 and not adjacent to metropolitan area. Again, after calculating percent diagnosed late, the result did not change much. The same counties that recorded the highest prevalence between 2003-2008 rates were among the top five counties with the highest percentage late stage diagnosis in the state (Figure 4.15).

Figure 4.14. Proportion of Breast Cancer Diagnosed Late per 100,000 Population

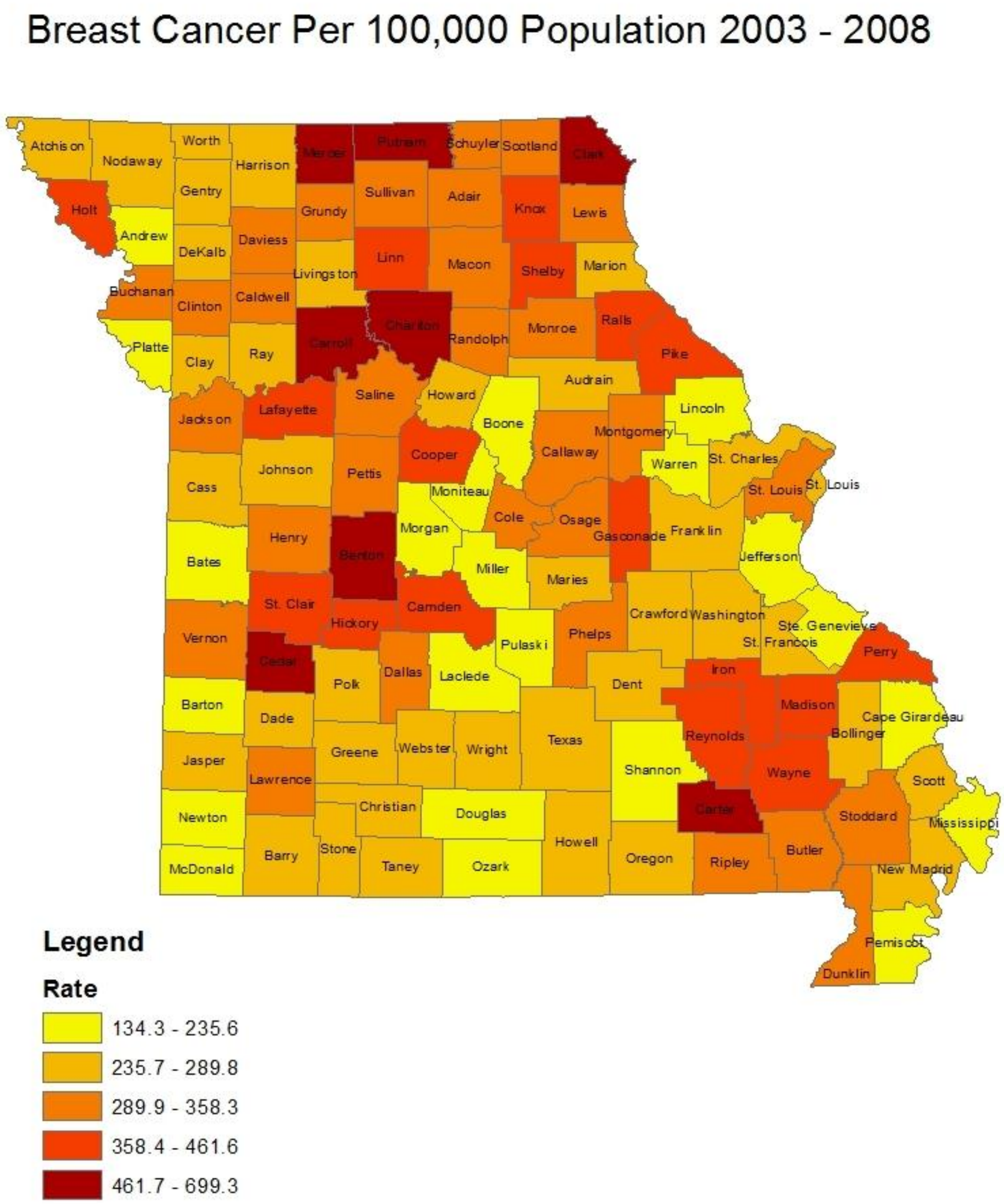
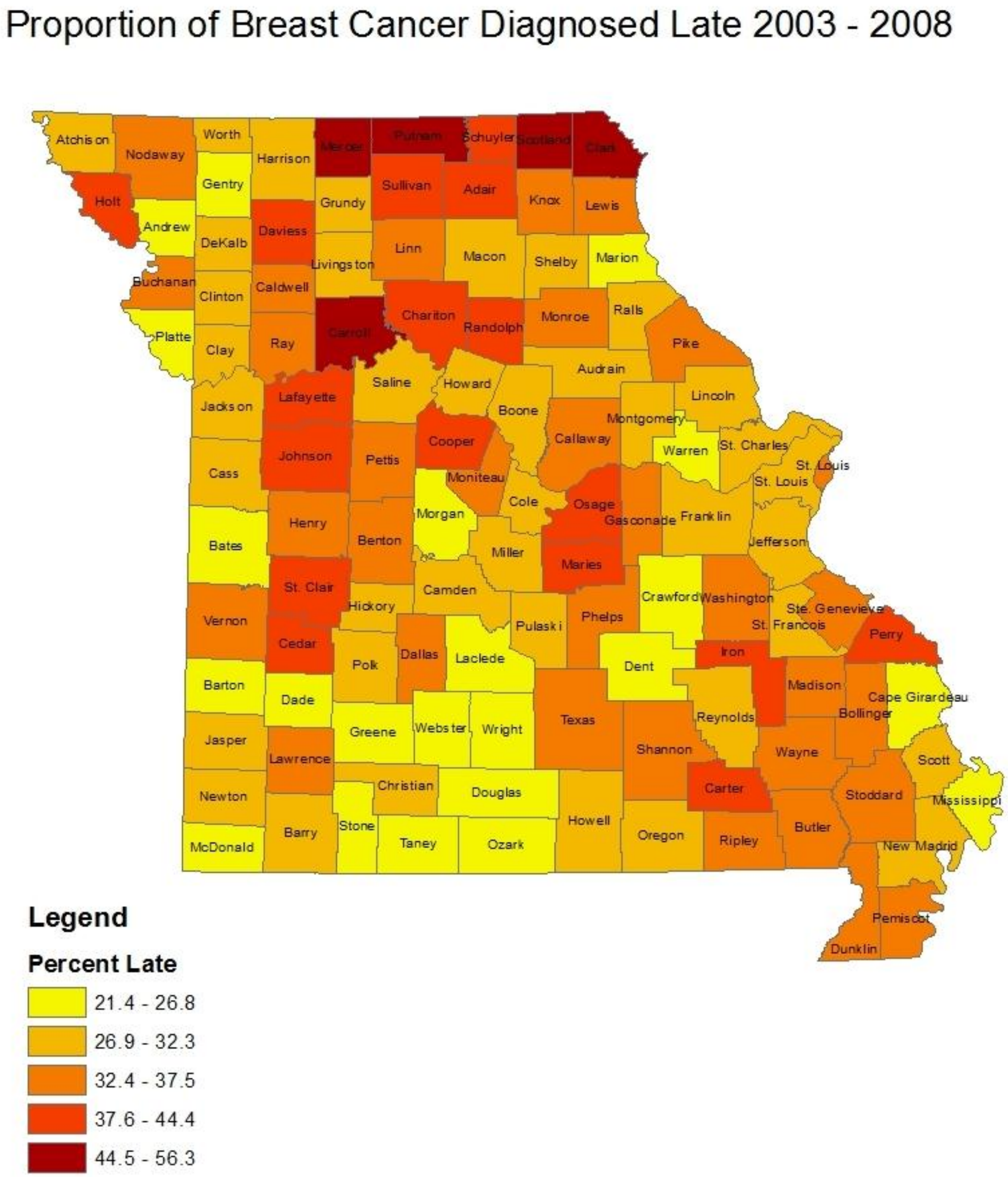


Figure 4.15. Proportion of Breast Cancer Diagnosed Late, 2003-2008



Racial disparities in access, diagnosis, treatment and survival of breast cancer among different racial and ethnic groups have been identified as a major factor impacting cancer incidence and mortality. As concluded McLafferty et al. (2011) rural-urban inequalities in risk are associated with differences in the demographic characteristics of area populations and differences in the social and spatial characteristics of the places in which they live. For instance, lower rate of breast cancer incidence is reported among minority women but the highest breast cancer mortality is seen among African American women than in white women. With this mind, percent late between the two different time periods was computed based on racial classification and county rural type. From 2003 to 2005 for white women in all 115 counties in Missouri, the highest percentage late stage cases resided in Clark, Monroe, Osage, Perry, Pemiscot, Chariton, Shannon, Adair, and Nodaway Counties with percentage ranging from 43.6 to 58.3 per female population (Figure 4.16).

Figure 4.16. Late Stage Diagnosis for White Women, 2003-2005

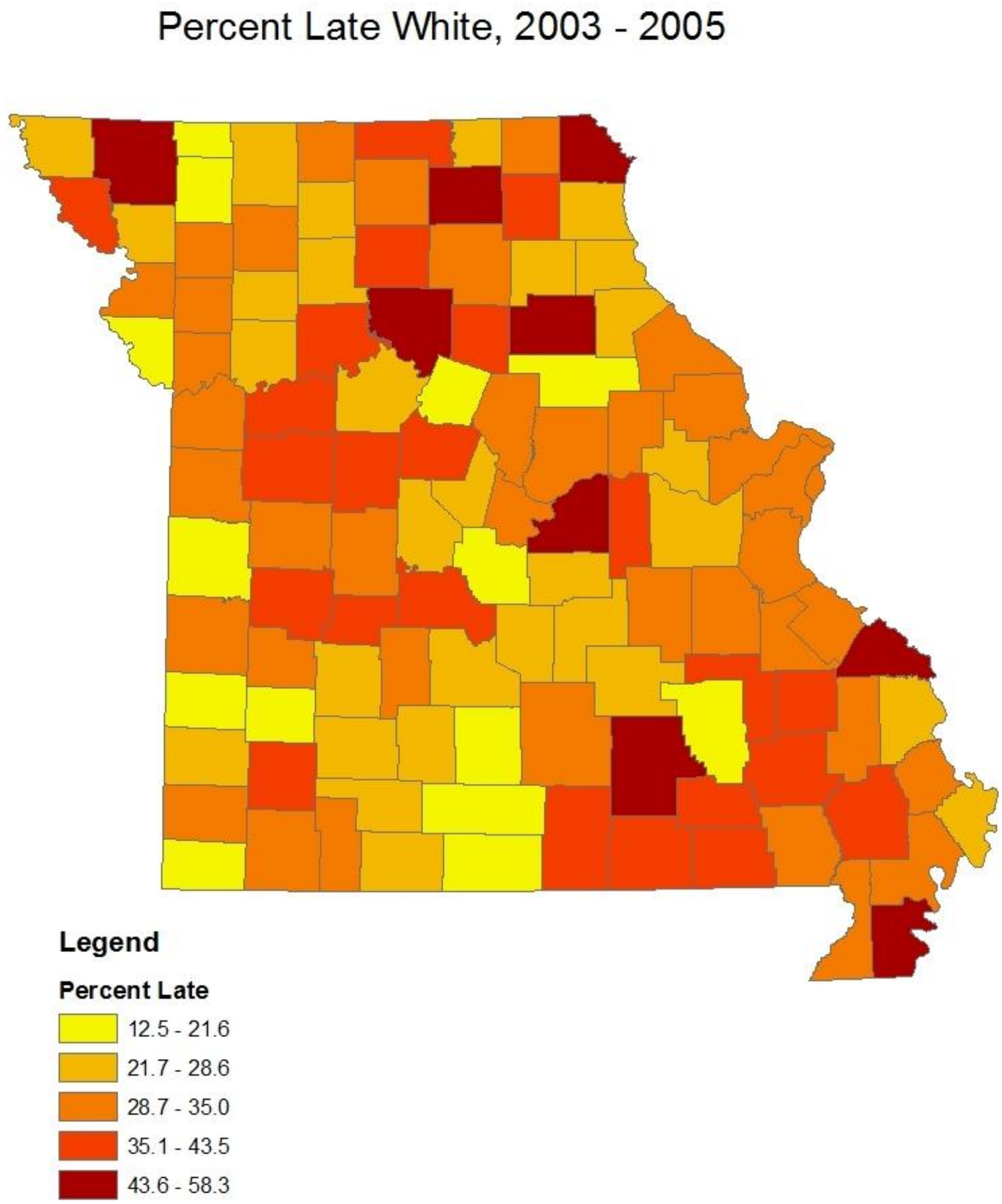
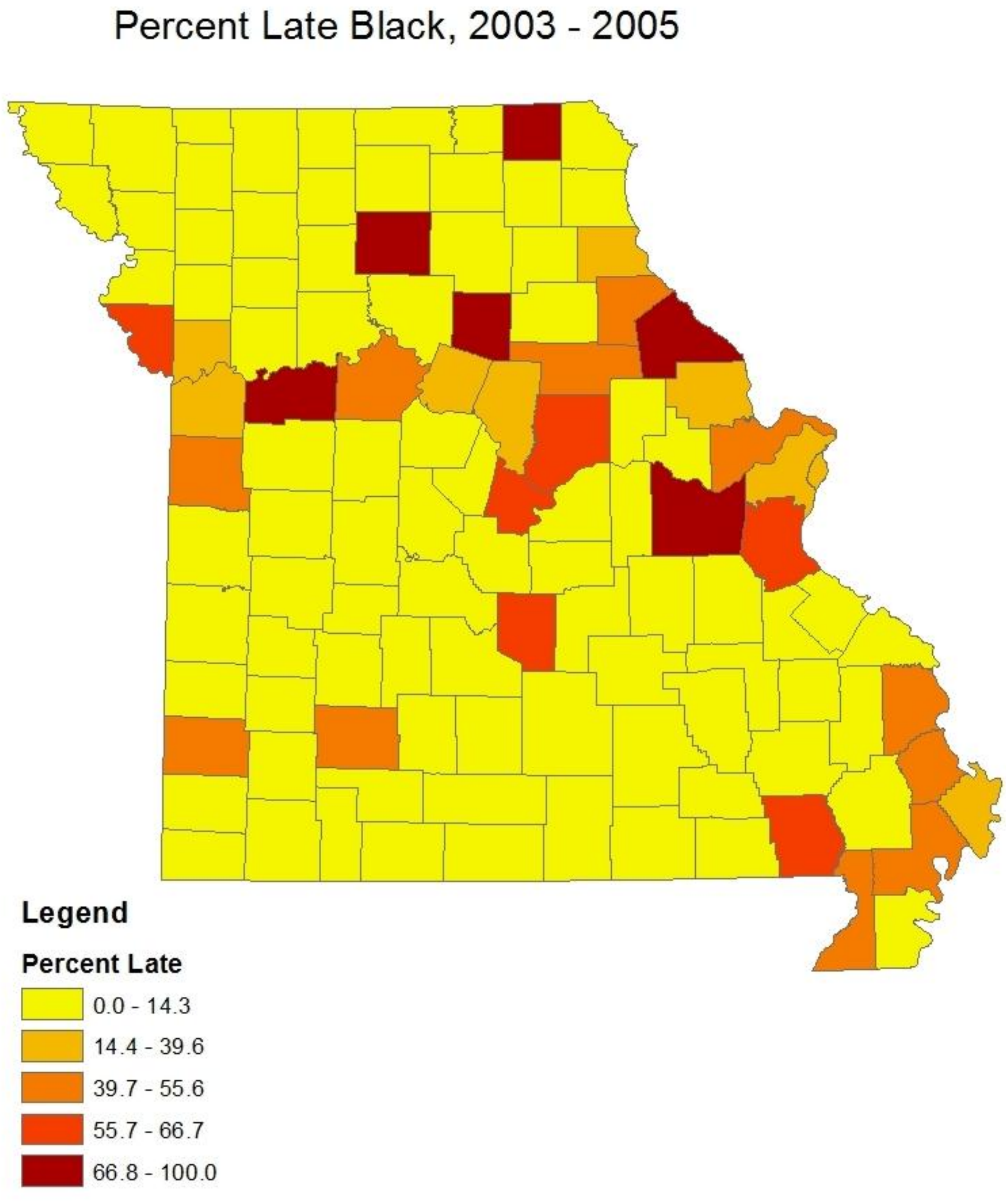


Figure 4.17. Late Stage Diagnosis for Black Women, 2003-2005



Among blacks, from 2003 to 2005, the highest late stage percentages of 66.8 and over were seen in Scotland, Lin, Pike, Lafayette, Franklin and Randolph Counties (Figure 3.17). There were no much differences in the pattern of percentage late distribution when considering rural-urban type during 2006 to 2008 for both white and black (Figures 4.18 & 4.19). Apart from Bates, Lincoln, Warren and Washington Counties which were among the highest incidence cases for blacks all the other counties were in nonmetropolitan rural areas (Figure 4.19).

Figure 4.18. Late Stage Diagnosis for White Women, 2006-2008

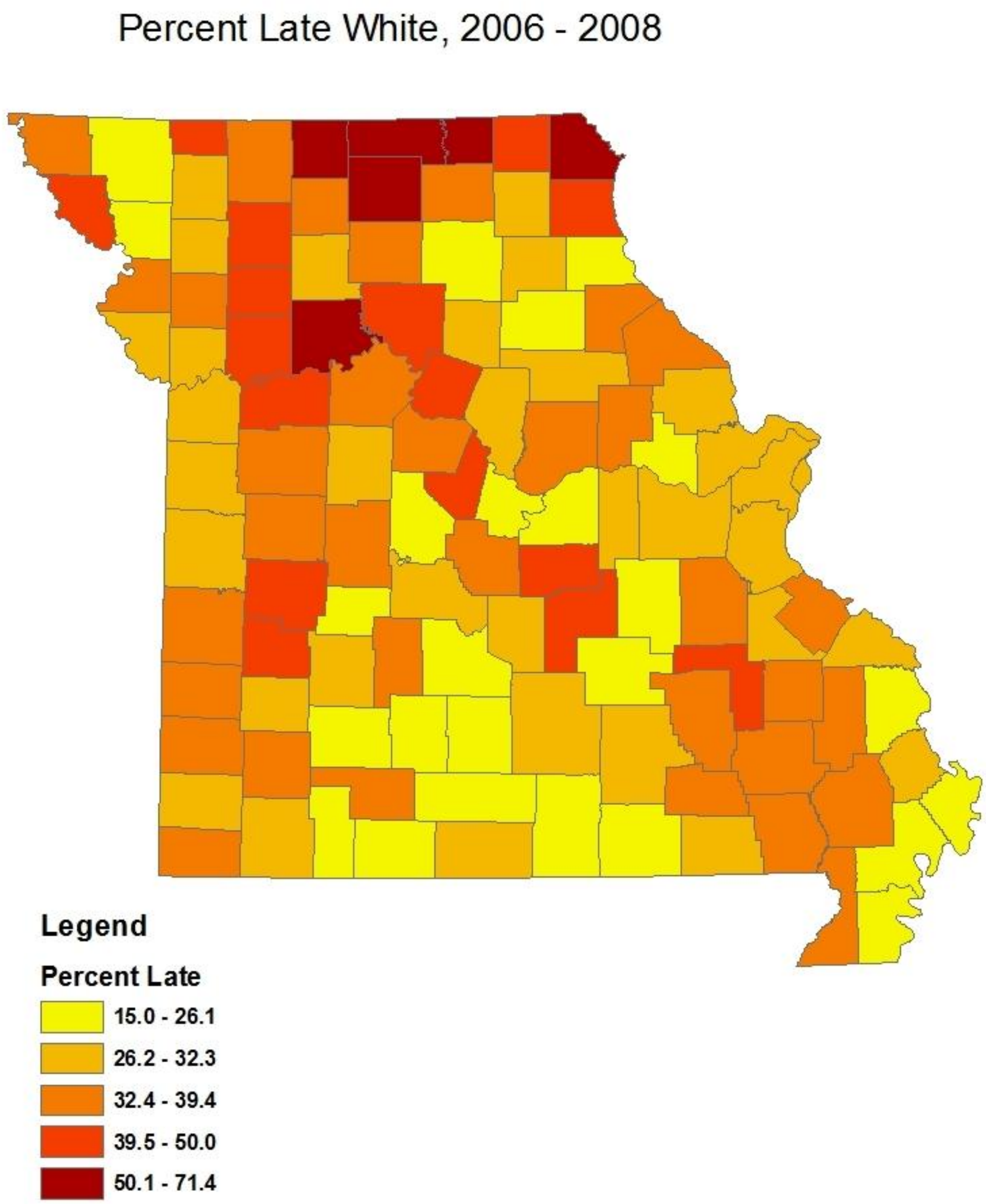
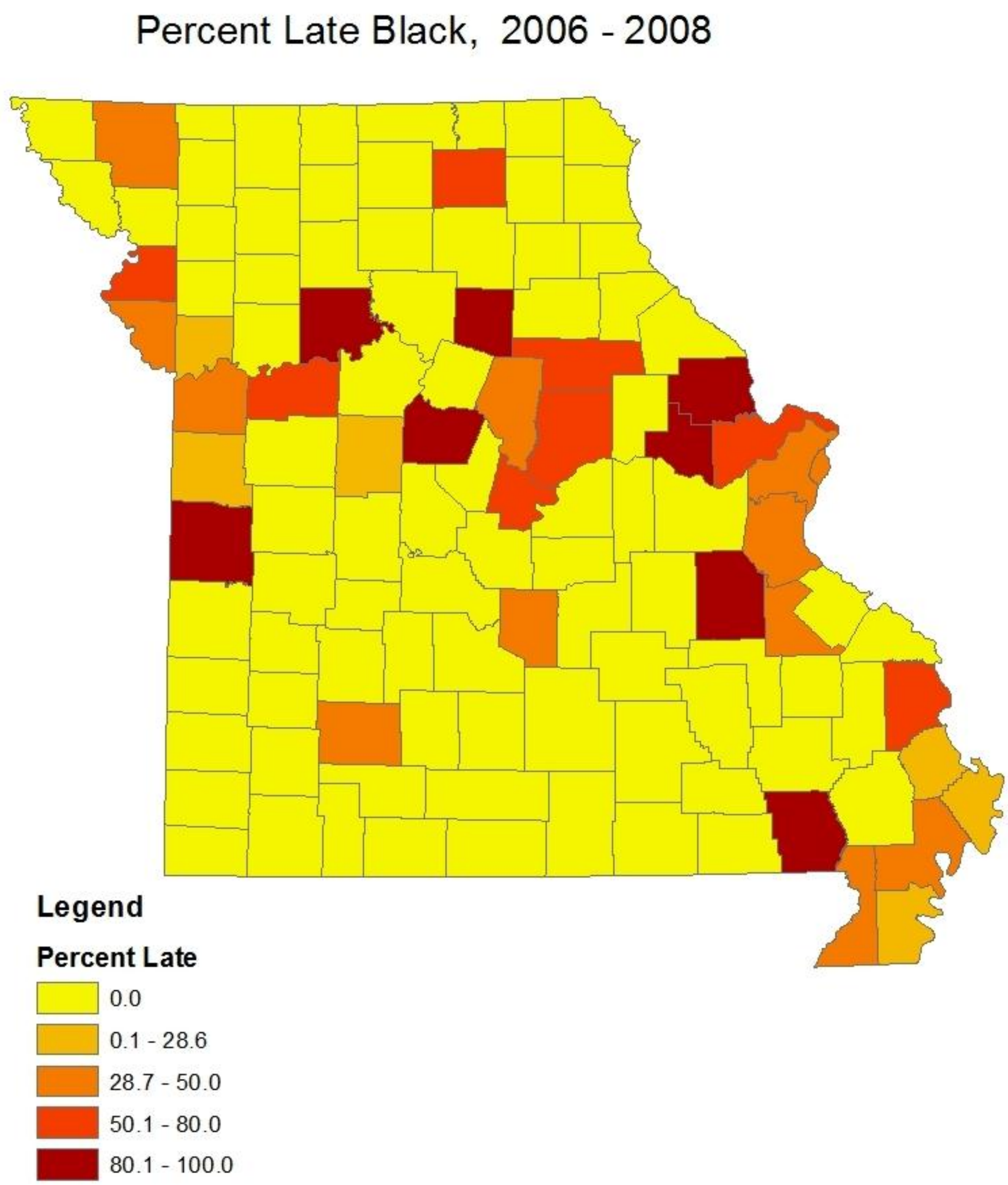


Figure 4.19. Late Stage Diagnosis for Black Women, 2006-2008



Distance Measure Description – Travel Time to the Nearest Medical Center in Missouri

According Penchansky et al. (1981) and Gold (1998) there are two dimensions of health care access. The first is economic access which refers to affordability and second geographic access which includes reasonable distance travel time to providers. Studies on geographic access also suggest that people usually seek health care in places that are nearer to them than those that are greater distance away (Gesler & Meade, 1998; Brustrom and Hunter, 2001). Similarly, Parkin, (1979), Williams, Schwartz, Newhouse and Bennett, (1983) also concluded that people may be discouraged from seeking early medical care if they are to travel lots of distance. However, there is hardly any agreement in the literature on the standard minimum travel distance that a person is require to access health care service. For a critical access hospital, the standard requirement is a distance of 35 mile. Based on this requirement, to explore the minimum and maximum distance travel time to any health care center in Missouri, several access time travel distances in minutes such as, 15, 30, 45 etc. were measured and compared.

The maps on Figure 4.20 through Figure 4.24 displays the minimum and maximum distance travel time using various time zone polygon for each Missouri County to access health care. The type of health care facilities include the *Show Me Healthy Women* (SMHW) mammography centers, hospitals, rural health clinics, critical access hospitals and federally qualified health centers.

Show Me Healthy Women (SMHW) is a free breast and cervical cancer screening program for the state of Missouri. The goal of the program is to reduce breast and cervical cancer mortality and morbidity by increasing availability of cancer screening for early detection of breast or cervical cancer among women in high-risk populations.

Currently, there are approximately 180 facilities throughout the state that provide these free cancer screenings. However, examination of the map tilted distance travel to mammography center (Figure 4.20) revealed that access to these services is not evenly distribution thought the state. For instance white St Louis City and County area has over twenty mammography centers, some countries do not access a single center. Using network analyst closest facility measure informs us that women living in Taney, Stone in the Southwest, Dunklin – Southeast, Nodaway, Worth – Northwest, Shannon and Dent in the Ozark regions among others have to travel over a distance of 45 minutes for mammography services.

Figure 4.20. Time Travel in Minutes to Mammography Center

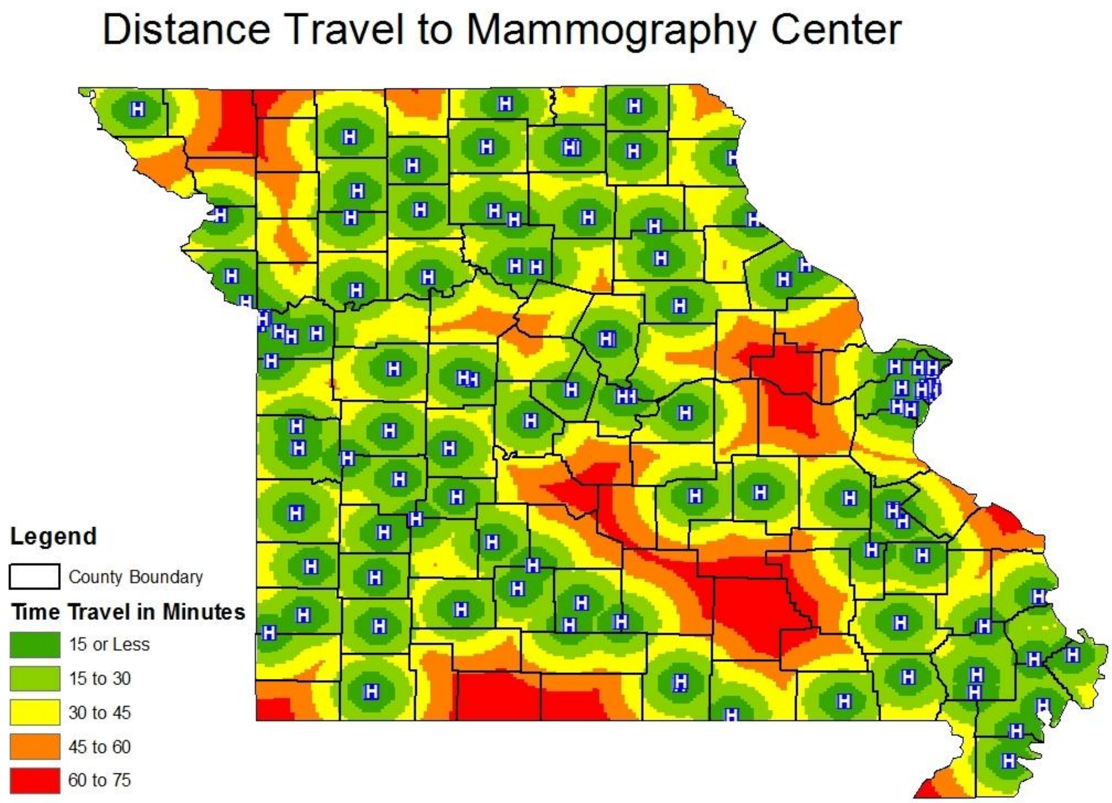


Figure 4.21. Time Travel to Hospitals in Missouri

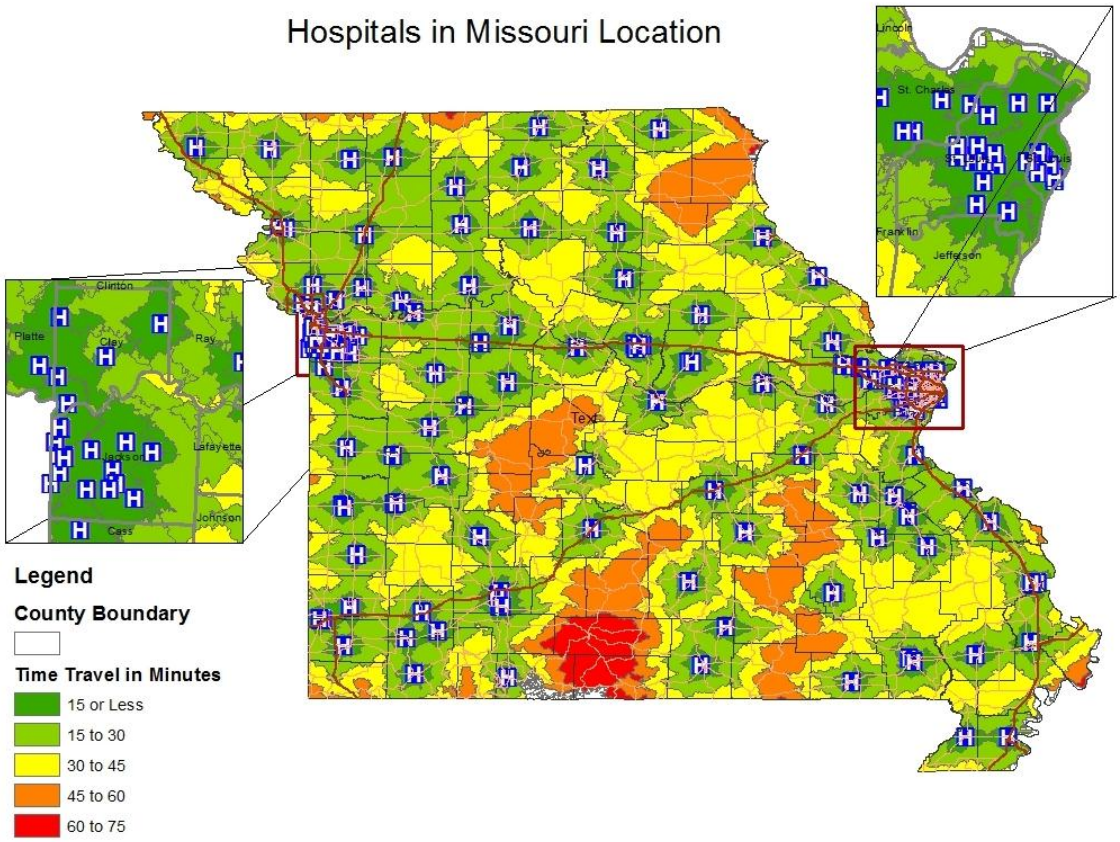


Figure 4.22. Time Travel in Minutes to Rural Clinics in Missouri

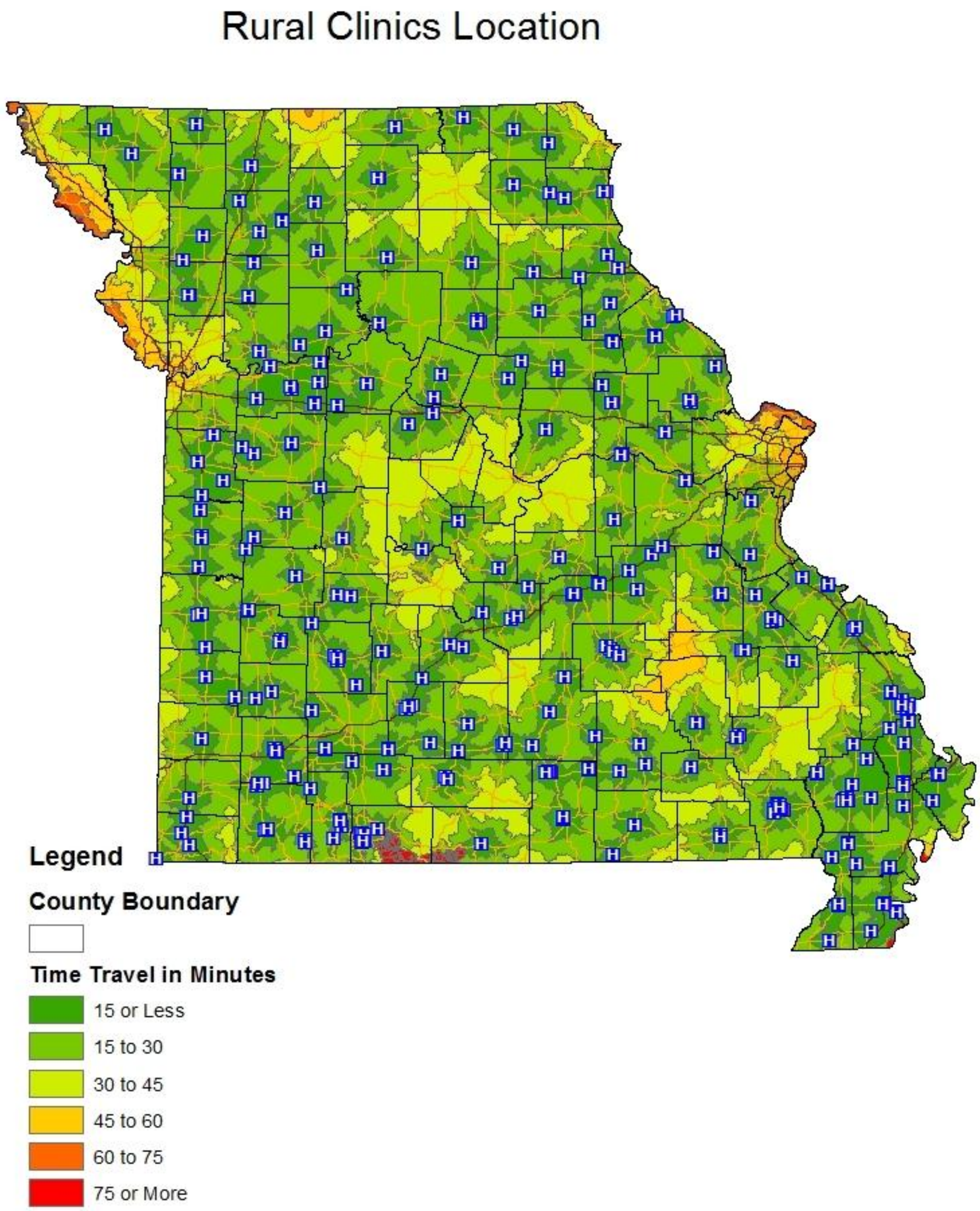


Figure 4.23. Time Travel in Minutes to Critical Access Hospital in Missouri

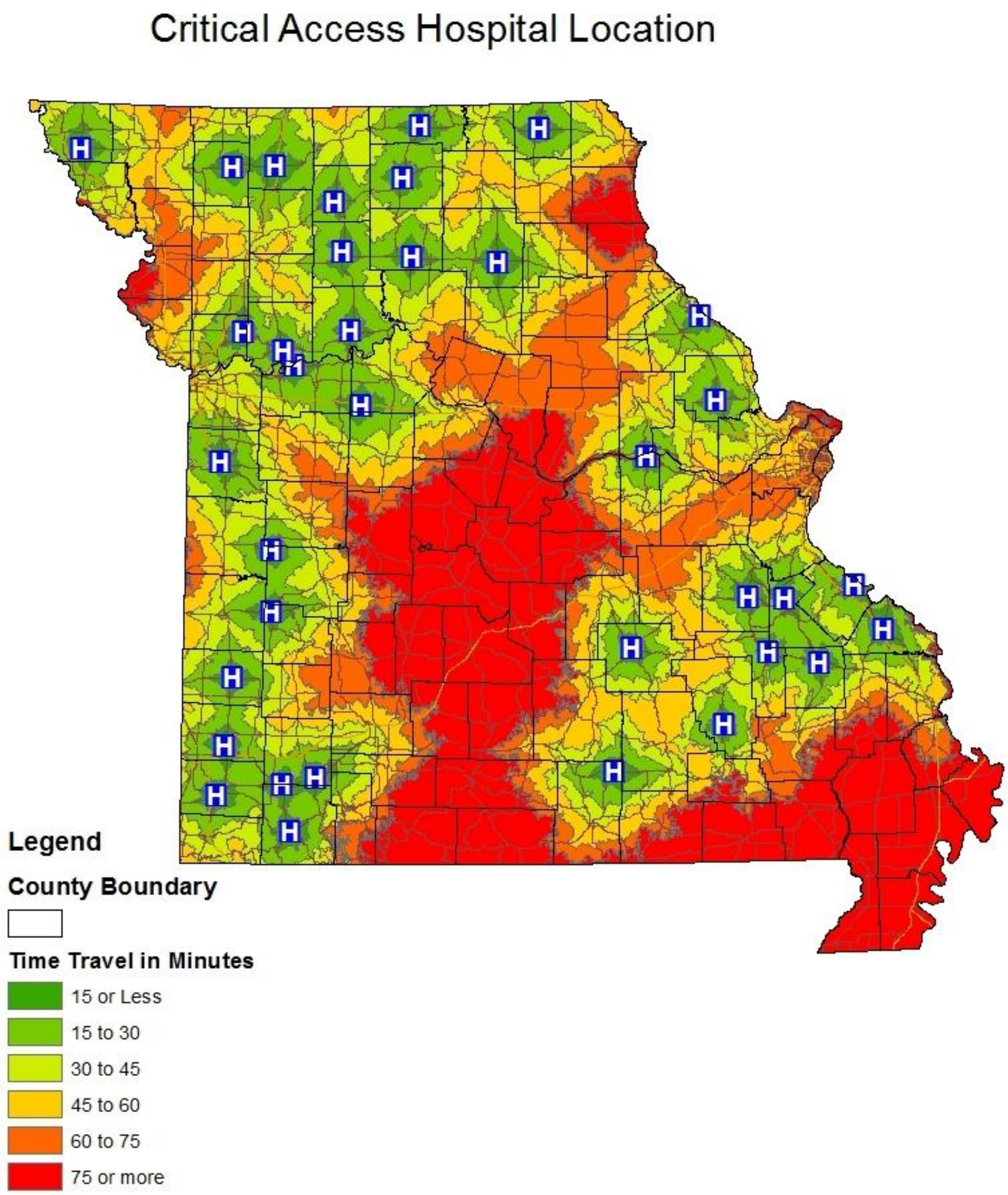
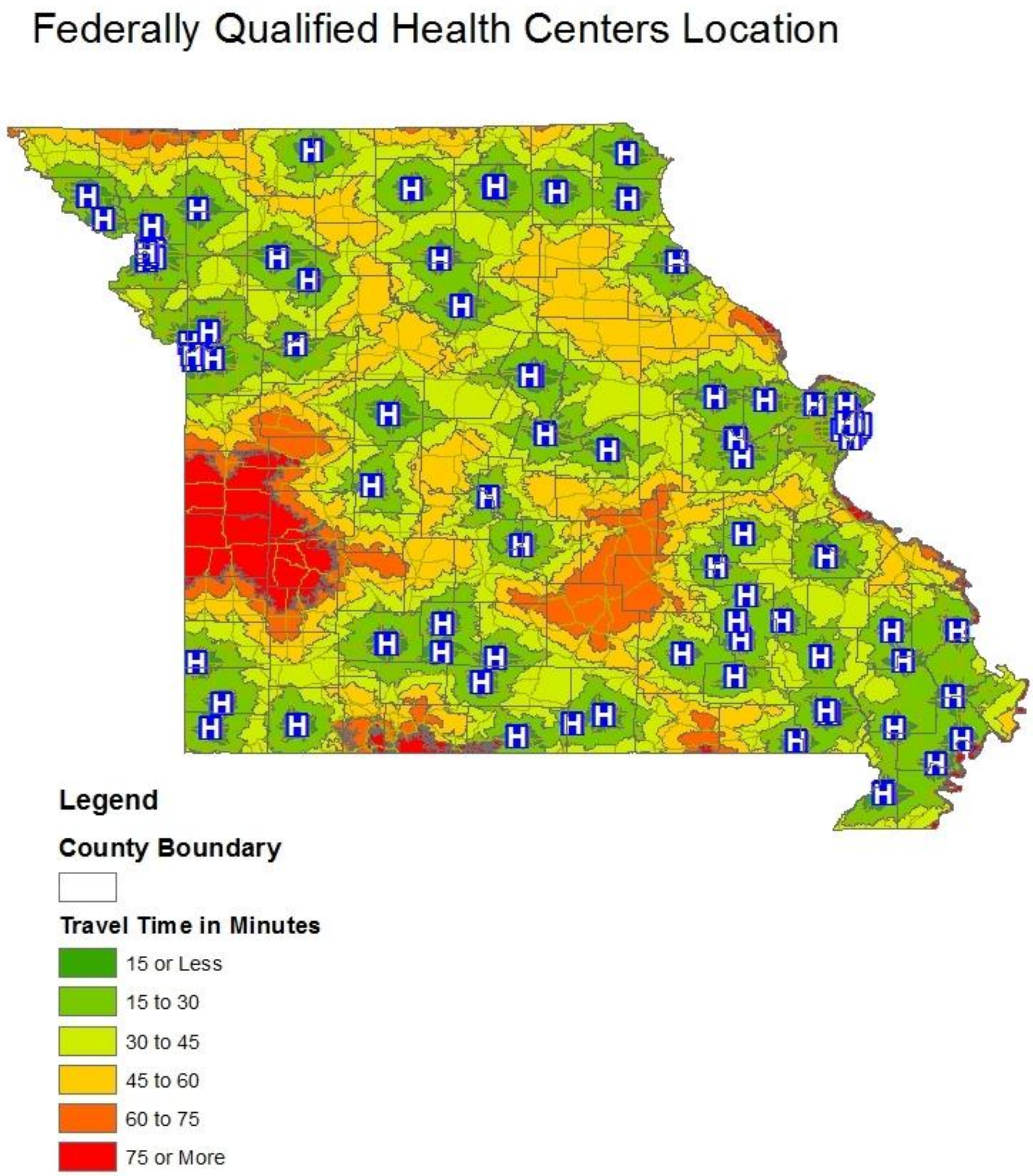


Figure 4.24. Time Travel in Minutes to FQHC in Missouri



Another important finding relates to Figures 4.23 and 4.24 above. Critical Access Hospitals (CAHs) and Federally Qualified Health Centers (FQHCs) are set up in the rural areas to help sustain the health of people there. Specifically, CAHs are designed to improve rural health care access and reduce hospital closures. Whereas FQHCs are to provide a detailed scope of primary health care as well as supportive services to all patients, regardless of their ability to pay. Unfortunately, most counties located in the Southern and South Central part of the state are required to travel more than 75 minutes one way for their medical care.

Logistic Regression to Predict Stage at Diagnosis

The maps on geographic distribution of breast cancer screening centers and hospitals in Missouri have shown large geographic differences. The question therefore is, does place of residence affect the stage at breast cancer diagnosis? To answer this question, logistic regression analyses were done in two main parts to contemplate the following hypotheses.

H_1 – Women with breast cancer in more remote non-metropolitan regions will, on average, be diagnosed at a more advanced stage than women in metropolitan areas over time.

H_2 – The negative effect of race, age, education and poverty on stage of breast cancer diagnosis will be increased by living in more remote non-metropolitan areas; i.e., a statistical interaction effect.

To explore the first hypothesis, a number of intuitive models were fitted to the data and this was followed by stepwise regression using the smallest AIC as the selection criterion. The details are given in appendix F. The final logistic regression model resulting from the stepwise regression selection process is presented below in Table 4.3.

Table 4.3.

Final Model from Stepwise Regression Selection

```
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.50803    0.15921   3.191  0.00142 **
age2        -0.44638    0.06431  -6.941 3.90e-12 ***
age3        -0.66576    0.06070 -10.968  < 2e-16 ***
age4        -0.80891    0.06032 -13.411  < 2e-16 ***
race         0.36765    0.04178   8.800  < 2e-16 ***
educ        -0.26301    0.05654  -4.652 3.29e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 35333  on 28535  degrees of freedom
Residual deviance: 34998  on 28530  degrees of freedom
AIC: 35010
Delta AIC = 335 df 5 p-value=0
```

Age was classified as a factor with four levels, race was coded as 0 for white and 1 for black, and education represents the county education score as described previously. The *p-values* in table 4.3 measure the strength of evidence against the null hypothesis that the associated regression coefficient is zero, given that all the other predictors are in the model. The smaller the *p-value* the stronger is the evidence against the null hypothesis, i.e. small *p-values* are strong. We now turn to the interpretation of the estimated regression coefficients in Table 4.3 above. The estimated coefficients associated with age are all negative with very small *p-values* suggesting that the evidence against the null hypothesis that these are zero is overwhelming. Thus, we considered age as an important predictor in the model. Also, the negative coefficients imply that the probability of late detection decreases with age. On the hand, the estimated coefficient with race is positive with a small *p-value*. Again, the evidence against the null hypothesis is strong, meaning that the probability of late detection is higher for blacks than whites. Finally the coefficient for education is negative also suggesting the probability of late detection reduces with county education score.

The odd ratios were calculated from the logistic regression as follows:

$$odds\ ratio\ for\ \frac{\text{Age } 40-49}{\text{Age } 18-39}:\ \ e^{0.44638}=1.563$$

$$odds\ ratio\ for\ \frac{\text{Age } 50-64}{\text{Age } 18-39}:\ \ e^{0.66576}=1.946$$

$$odds\ ratio\ for\ \frac{\text{Age } 65+}{\text{Age } 18-39}:\ \ e^{0.8089}=2.245$$

$$odd\ ratio\ for\ \frac{black}{white}:\ \ e^{0.36765}=1.444$$

$$odd\ ratio\ for\ \frac{Education\ score=2}{Education\ score=3}:\ \ e^{0.26301}$$
$$=1.300$$

Within the analysis of covariant model above, no significant interactions were observed. However, we suspected interaction was present but not detectable within the geometric constraints of this analysis of this covariant model. Hence, individual models including race and education were fitted for every age group. A brief summary of the model is given in Table 4.4 below.

Table 4.4.

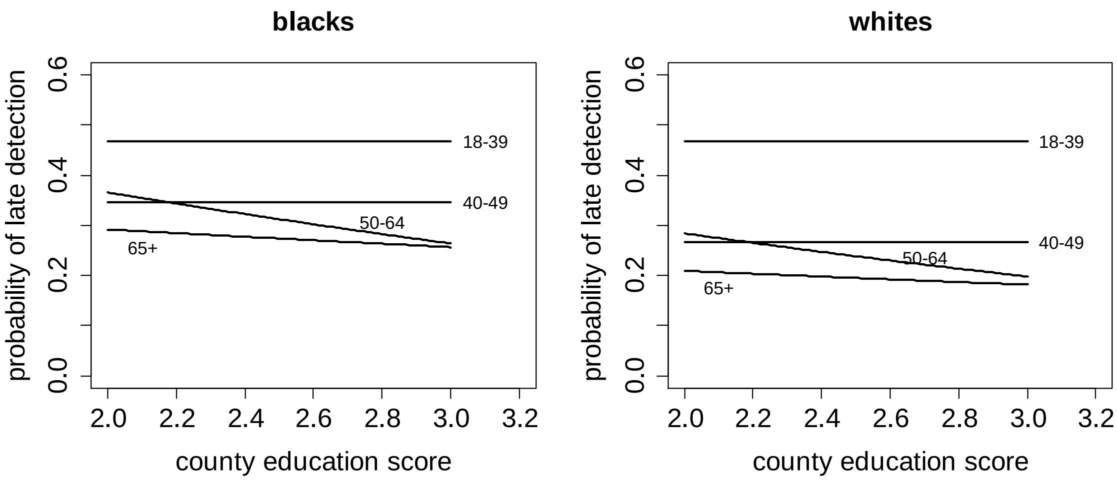
Fit for Race and Education for Each Level of Age

Fit	Predictors	Estimate	SE	Z-Values	P-Value
age1	Intercept	-0.10638	0.68647	-0.155	0.877
	Race	-0.06682	0.15815	-0.423	0.673
	Edu	0.02049	0.26059	0.079	0.937
age2	Intercept	-0.73894	0.35861	-2.061	0.0393 *
	Race	0.38789	0.09021	4.300	1.71e-05 ***
	Edu	-0.10817	0.13427	-0.806	0.4204
age3	Intercept	0.02074	0.24842	0.083	0.933
	Race	0.37159	0.06909	5.378	7.53e-08 ***
	Edu	-0.47265	0.09400	-5.028	4.96e-07 ***
age4	Intercept	-0.97665	0.23608	-4.137	3.52e-05 ***
	Race	0.44136	0.07011	6.295	3.07e-10 ***
	Edu	-0.17596	0.08801	1.999	0.0456 *

Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The model for age group 18-39 had no significant predictors and for this group the probability of late detection did not depend on race or education. The model of age group two which is 40-49, only race was important. For age group three both race and education were significant. Finally for age group four, race was the most significant predictor. These models were plotted separately for blacks and white in Figure 4.25 below; given the geometric picture of the differences and interactions that exist between the variables. Although no formal statistical tests were computed to compare these models, a large; number of observations in each of the models, gave us confidence in the representation, of the data.

Figure 4.25. Probability of Late Diagnosis by Race, Age and County Education Score



Considering the regression lines for 18-39 age groups, for both blacks and whites, the regression lines are horizontal at the same height, because race and education were not significant. For the 40-49 age groups, the regression lines were again horizontal because education was not significant but were at different heights because race was significant, and blacks are at higher risk for late detection than whites. For 50-64 age groups, both race and education were highly significant resulting in parallel lines in different heights. This means that, blacks at higher risk at late detection than whites. For 65 and older age group, race was highly significant as seen in the parallel line. Also, blacks appeared to be at higher risk for late detection than whites. Even though education is equally important for this age group, it was not as important as race. The negative slope implies that the risk of late detection decreases with increase with county education score or decreases as county education score increases. Appendix H provides detailed results on proportion of late detection by the four age groups.

Summary

Using data from the Missouri Cancer Registry and Research Center, this chapter discussed process of data analysis and findings on 28,536 women who were diagnosed with breast cancer from 2003 to 2008. The majority of the population was white. Using RUCC metro and nonmetro classification system, a new code system was derived using county size to assess geographic location impact on female breast cancer diagnosis in Missouri. Even though the proportion of black or Africa American diagnosed with breast cancer during the study period were less than 10 percent when percentage late stage diagnosis was calculated using the new six rural type groups, the total percentage of black diagnosed with late stage cancer far exceeded those of white in every rural type location throughout the state.

Also, county level educational score was computed using the number of persons in each county with no high school diploma, high school diploma, some college degree and bachelor and beyond as a proxy for education. Results indicated that county high educational score increases with county or population size. This means that more individuals in metropolitan or urban areas are more likely to be highly educated than those in rural area. Further, socioeconomic status is widely known to correlate health and well-being. Using boxplot the relationship between education and age, race, stage at diagnosis and spatial isolation was examined.

Studies have shown that geographic location affects access and utilization of health services. To explore the minimum and maximum distance travel time to any health care center in Missouri, several access time travel distances in minutes such as, 15, 30, 45 etc. were measured and compared. In spite of the numerous health care services in

Missouri, the network analysis closest facility travel time indicated that access to these services is not evenly distributed in the state. Counties located in the Southwest, Southeast and South Central for instance have to travel over 60 minutes one way for mammography and other medical care. The burden of long travel time for health care services could account for the high prevalence of late detection in these areas.

Finally logistic regression was used to explore the hypothesis on spatial isolation and stage at diagnosis. While there was no direct interaction among the important variables, the fitted logistic regression for each of the four age categories, indicated that race and county educational score are the important predictors for late stage breast cancer detection.

CHAPTER FIVE
DISCUSSION, CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

Discussion

The purpose of this research was to examine the role of spatial access to health care services on the probability of late detection of female breast cancer diagnosis in Missouri taking into account the access and distance to clinics and hospitals. There were two central research questions which were examined in this study. The first was to what extent does spatial geographic access to diagnostic facilities have on the stage at which breast cancer is diagnosed? The second question was to what extent are the effects of other social factors such as race, age and poverty associated with later diagnosis of breast cancer? Data for the study came from the Missouri Cancer Registry and Research Center. It comprised of 28,536 cases on women in Missouri who were diagnosed of breast cancer between 2003 and 2008. The findings from this study are important because the magnitude of these factors on late diagnosis of breast cancer was quantified. Also, these findings are essential because no previous studies have specifically looked at spatial isolation and its effect on female breast cancer diagnosis in this state.

Previous studies all over the world have indicated that GIS provides an effective way of assessing health care access and issues relating to accessibility in the community. Cromley et al. (1998); Wang et al. (2008); Owen et al. (2010); Oppong et al. (2005) have all applied the techniques in GIS to access to health services in various countries. For instance in Illinois, Wang and colleagues (2008) used spatial analysis methods to create a measure of spatial access to primary care and mammography clinics on late stage breast cancer diagnosis. The current study contributes to the use of GIS in addressing access to health care by looking how geographic location and distance travel time affect both early and late stage diagnosis of breast cancer.

Using ArcGIS 10 network analyst function, each county centroid was computed for all 115 counties in Missouri. The next step, was to determine network travel time to closest facilities using various distance travel polygons of 15, 30, 45 etc. in minutes. County centroids as origin and service areas as destinations were mapped together to calculate population proportion in each travel closest facility. GIS results on mammography center locations revealed that Nodaway, Warren, Franklin, McDonald, Taney, Ozark and Shannon Counties do not have a single screening center. Women in these counties have to travel a distance of over 60 minutes one way if they are to benefit from the *Show Me Healthy Women* free screening breast and cervical cancer services. This finding highlights the predicament women in these counties face regarding access to breast cancer screening services in Missouri.   In contrast, counties such as Boone, Jackson, St Louis City and County have an abundance of health care facilities within maximum of 30 minutes travel.

Disparities in breast cancer between white and black have been well documented over the years (MacKinnon et al., 2007; MacKinnon, Duncan & Huang et al., 2007). Black women are well known to have lower incidence and prevalence rates of breast cancer than white women, but they have a higher mortality rate and a lower survival rate (ACS, 2012; Komen for the Cure, 2012). Causes of these disparities have been linked to social, behavioral, and economic factors such as persistent inequalities in access to care, unhealthy environments, and racial discrimination (Campbell et al., 2000; Wang et al., 2008; Jordon, Roderick, Martin & Barnett 2004; Cromley & Cromley 2009; Peters et al., 2008). In the current study, detailed results revealed that the percentage of late breast cancer diagnosis for blacks were always higher than whites when comparing all rural-urban type of residence even though county-level educational attainment indicated that overall blacks have higher median educational scores than whites. This result was a rather surprising result that goes against human folk-law beliefs. A systematic review by Johnson, Elbert-Avila and Tulsky (2005), indicated that African Americans and Hispanics are prone to rely on spiritual help and prayer rather than formal health care infrastructure to cope with sickness, treatment options and the restoration of their health. Africa American women in particular, believe that only God has the power to heal and decide on life and death. As a result, a strong spiritual believe could serve as a hindrance in seeking a woman's medical treatment for their health. Considering the high level of educational score among black population in Missouri, it is likely that their strong faith and spiritually is contributing to later stage breast cancer diagnosis among them.

Various studies have found that place of residence and neighborhood socioeconomic characteristics, are associated with cancer outcomes and quality of life (Singh et al., 2003). Studies for several cancer sites have shown that individuals living in poor areas are more likely not to utilize cancer screening services and present at a late stage compared with individuals living in affluent areas (Henry, 2009). McLafferty et al. (2009), Wang et al. (2008), Campbell et al. (1991), and Campbell et al. (2000) reported that higher risk of late stage cancer diagnosis is prevalent among rural residents who face long distances in accessing cancer screening services. The maps corroborate these findings. However, due to the geometric constraints on the logistic regression, this finding was not observed.

Henry et al. (2011), Henry, Sherman & Roche (2009), Warner et al. (2010) found that geographic place of measures were weakly associated with differences in the risk of late stage breast cancer. The outcomes from the logistic regression in this study corroborate their findings that rural-urban residence and poverty level of a county did not have any significant effect on the stage of breast cancer diagnosis in Missouri even though the highest late stage cancer cases were in nonmetropolitan counties.

In this study, the most important predictors of breast cancer diagnosis were age, race and county level educational attainment. Comparing the stage of diagnosis in terms of race, and education score for each age category, gave the following results: For the 18-39 age category, neither race nor education was statistically significant, and the probability of late detection was constant at 48 percent. For the 40-49 age groups, the regression lines were again horizontal because education was not significant but were at different heights because race was significant, and blacks are at higher risk for late detection than whites. For 50-64 age groups, both race and education were highly significant resulting in parallel lines in different heights. This means that, blacks are at higher risk at late detection than whites. For 65 and older age group, race was highly significant as seen in the parallel line. Also, blacks appeared to be at higher risk for late detection than whites. Finally, even though education is equally important for this age group, it was not as important as race.

Conclusions

Access to health care services is an important health policy issue. Past studies have all recognized the multi-dimensional effect of access to health care services. Access and health outcomes usually vary by many factors including age, race, education as well as geographical location and other SES like poverty. Based on the GIS maps and the derived variables of county education and poverty, the distribution of cancer prevalence during the six year period, indicated that women living in areas with limited access to health care services are more likely to be diagnosed with late stage breast cancer. However, the fitted logistic regression models did not detect any relationship between geographical location and later stage at diagnosis. The maps are simple county averages and are therefore without any constraints and much more versatile. The logistic regression has strong geometric constraints. The two methodologies, therefore view the distributions of late stage from an entirely different perspective, and thus give us different picture. Thus, this is like viewing an object from different viewing points. The findings are not contradictory but they are different.

In conclusion, the logistic regression analyses did not support the contemplated research hypotheses. There were no identified relationship between geographical location, access and late stage breast cancer diagnosis in Missouri, as expressed by the derived variable of adjacent, county education score, poverty and county population variables such as county head of households, and female population head of households.

It was hypothesized that women with breast cancer in more remote non-metropolitan regions will, on average, be diagnosed at a more advanced stage than women in metropolitan areas. No statistically significant association was found between county level percent living below the poverty line, and female headed household and late detection as well in the logistic regression. Although findings from this study suggest no increased risk of breast cancer in nonmetropolitan areas in Missouri, based on the logistic regression results, the following conclusions were made.

1. Based on the GIS maps, women residing in a rural or nonmetropolitan area at a higher risk for late stage breast cancer diagnosis.
2. The GIS analysis indicated that county-level educational score correlates highly with poverty score. Thus, counties with high education score are less likely to be poor while, those with low county-level educational score are more likely to be poor.
3. The majority of women in rural Missouri counties does not have access to screening and other health care services and had to travel over 60 minutes one way for medical care. This travel burden resulted in a higher probability of late detection.

Secondly, even though the logistic regression model did not show any interaction among the predictors, and in particular no effect of distance and poverty on stage at diagnosis, the following conclusions can be made from the logistic regression results.

1. Among younger white and black women (18-39 age groups), the effect of race and county-level educational score on late detection was similar. However, for older group (40 years and older), the effect of race, and in particular the lack of education on late detection was greater among blacks than whites.

2. Overall, the age of a woman, race and county-level educational score of residence were the most statistically significant factors in predicting late stage cancer diagnosis among women in Missouri.

Implications, Future Research and Recommendations

Breast cancer remains the leading cause of mortality among women across all racial and ethnic groups in the United States. The GIS mapping system, has provided strong insight into the joint predicament that goes along with low county-level educational score, and high poverty rate in a manner that formal statistical models are not able to achieve in this study due to geometric constraints. The formal statistical model of logistic regression, using the derived variables in this study, seems unable to match the versatile simplicity of GIS mapping system and its color coding. Although the logistic regression detected little statistical significance among the important variables considered, there are important implications for future research and practice. The findings from this research have highlighted the complex nature of factors like geographical location, poverty, education, race and access to health services on breast cancer detection and treatment. The approach used in the current study has provided useful information on provision of health care access and accessibility regarding what health care services is needed, where, and to whom coverage is lacking. This analysis can serve as a guide to policy makers in the state of Missouri, about deliberation on health care resource allocation as well as prioritizing targets.

As the determinants of health such as environmental, socio-cultural and the physical environment differ greatly in space, so also, does people's health care needs differ from place to place. Geographic Information System can therefore serve as an essential tool for researchers, social scientists, health educationist, as well as for health planning, monitoring and evaluation of effectives of health programs to reduce premature disability and death due to breast cancer. There is therefore, a need for further research on the use of GIS to identify and measure spatial relationship between geographical location, environmental and social factors effect on breast cancer detection and diagnosis. In addition spatial access to primary health care services is critically important for early breast cancer detection. Lastly, due to the importance of socioeconomic factors on health and wellbeing, it is recommended for cancer registries to collect data on education and poverty.

Delimitations of the Study

The delimitation of this study is that it specifically applied to women with breast cancer in Missouri, who were diagnosed between 2003 and 2008 and whose cancer was reported to the Missouri Cancer Registry and Research Center. However, findings from this study can equally apply to women in other states and with similar characteristics and may have global applications.

Limitations of the Study

Just like any research work, this study was not without limitations. First, County was used as a unit of analysis rather than block or tract group. County-level unit covered a very large and often diverse area therefore less likely to provide accurate information on characteristics of each breast cancer patient in that county. Secondly, due to restrictions on cancer data usage and protection of patient privacy and confidentiality the four main distinctive stages at diagnosis (*in situ,* localized, regional and distant) were collapsed into two – early (*in situ* and localized) late (regional and distant). In reality these stages are completely different from each other when considering tumor size, treatment options, survival rate and quality of life of patients. Also, absence of data on education, income, and poverty on each patient was another major limitation. Finally, county centroid rather than actual address of residence was used to measure distance travel to a specific health care facility available in each county.

APPENDICES

A. COMMONLY USED RURAL DEFINITIONS

Definition	Definition Description	Geographic Unit Used
U.S. Census Bureau: Urban and Rural Areas	The Census Bureau's classification of rural consists of all territory, population, and housing units located outside of urbanized areas and urban clusters. Urbanized areas include populations of at least 50,000, and urban clusters include populations between 2,500 and 50,000. The core areas of both urbanized areas and urban clusters are defined based on population density of 1,000 per square mile and then certain blocks adjacent to them are added that have at least 500 persons per square mile.	Census Block and Block Groups
Economic Research Service, U.S. Department of Agriculture & WWAMI Rural Health Research Center: Rural-Urban Commuting Areas (RUCAs)	This classification scheme utilizes the U.S. Census Bureau's urbanized area and cluster definitions and work commuting information. The RUCA categories are based on the size of settlements and towns as delineated by the Census Bureau and the functional relationships between places as measured by tract level work commuting data. This taxonomy defines 33 categories of rural and urban census tracts.	Census Tract, ZIP Code approximation available
U.S. Office of Management and Budget (OMB): Core Based Statistical Areas (i.e., Metropolitan and Nonmetropolitan areas)	A metropolitan area must contain one or more central counties with urbanized areas. Nonmetropolitan counties are outside the boundaries of metropolitan areas and are subdivided into two types, micropolitan areas and noncore counties. Micropolitan areas are urban clusters of 10,000 or more persons.	County
Economic Research Service, U.S. Department of Agriculture: Rural-Urban Continuum Codes (Beale Codes)	This classification scheme distinguishes metropolitan counties by the population size of their metropolitan area, and nonmetropolitan counties by degree of urbanization and adjacency to a metropolitan area or areas. All counties and county equivalents are grouped according to their official OMB metropolitan-nonmetropolitan status and further subdivided into three metropolitan and	County

	six nonmetropolitan groupings.	
Economic Research Service, U.S. Department of Agriculture: Urban Influence Codes	This classification scheme subdivides the OMB metropolitan and nonmetropolitan categories into 2 metropolitan and 10 nonmetropolitan categories. Metropolitan counties are divided into two groups by the size of the metropolitan area. Nonmetropolitan-micropolitan counties are divided into three groups by their adjacency to metropolitan areas. Nonmetropolitan-noncore counties are divided into seven groups by their adjacency to metropolitan or micropolitan areas and whether they have their "own town" of at least 2,500 residents.	County
Office of Rural Health Policy, U.S. Department of Health and Human Services: RUCA Adjustment to OMB Metropolitan and Nonmetropolitan Definition	This method uses RUCAs 4-10 to identify small towns and rural areas within large metropolitan counties. In addition, census tracts within metropolitan areas with RUCA codes 2 and 3 that are larger than 400 square miles and have population density of less than 30 people per square mile are also considered rural.	Census Tract within OMB Metropolitan Counties

B. DEFINITION OF RUCC AND ALLOCATION OF RUCC TO MISSOURI COUNTIES

RUCC	Description	Missouri Counties
Metro Countries Classification		
1	Counties in metro areas of million population or more	
2	Counties in metro areas of 250,000 to 1 million population	
3	Counties in metro areas of fewer than 250,000 population	
Nonmetro counties Classification		
4	Urban population of 20,000+, adjacent to a metro area	
5	Urban population of 20,000+, not adjacent to a metro area	
6	Urban population of 2,500 to 19,999, adjacent to a metro are	
7	Urban population of 2,500 to 19,999, not adjacent to a metro area	
8	Completely rural or less than 2,500 urban population, adjacent to a metro area	
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area	

Source: USDA - 2003 Rural-Urban Continuum Code

C. STEPWISE REGRESSION ANALYSIS

```
tina=read.table("F:/fresen      2      cancer      120303      no
nines.txt",header=T)
attach(tina)
dim(tina)
[1] 29410    26
tina[1:10,1:10]
   ru urban popsize adjacent age stage race year  fips countypop
1   7     0  10,000        0   4     0    1 2005 29001    24,801
2   7     0  10,000        0   4     0    1 2008 29001    24,801
3   7     0  10,000        0   4     1    1 2005 29001    24,801
4   7     0  10,000        0   4     1    1 2005 29001    24,801
5   7     0  10,000        0   4     1    1 2008 29001    24,801
6   7     0  10,000        0   3     0    1 2008 29001    24,801
7   7     0  10,000        0   1     1    1 2005 29001    24,801
8   7     0  10,000        0   3     0    1 2005 29001    24,801
9   7     0  10,000        0   3     1    1 2005 29001    24,801
10  7     0  10,000        0   2     1    1 2005 29001    24,801
```

This section shows the barplot of age, stage, race, and year at diagnosis

```
par(mfrow=c(1,4))

barplot(table(stage),
ylim=c(0,22000),main="stage",cex.main=3,cex.names=2,cex.axis=2)
text(0.65,21000,"early",cex=2)
text(1.92,10200,"late",cex=2)


barplot(table(age),main="age",cex.main=3,cex.names=2,cex.axis=2,yli
m=c(0,15000))
text(0.66,2100,"18-39",cex=2)
text(1.86,5600,"40-49",cex=2)
text(3.1,11110,"50-64",cex=2)
text(4.3,13048,"64+",cex=2)

barplot(table(race),main="race",
cex.main=3,cex.names=2,cex.axis=2,ylim=c(0,30000))
text(0.62,28200,"white",cex=2)
text(1.89,5560,"black",cex=2)


barplot(table(year),
main="year",cex.main=3,cex.names=2,cex.axis=2,ylim=c(0,18000))
text(0.66,15076,"2003-2005",cex=2)
text(1.88,15816,"2006-2008",cex=2)
```

The following barplot codes shows diagram of rural urban code, metro vs nonmetro, the population of each county, adjacent/nonadjacent rurality type.

```
par(mfrow=c(1,2))
barplot(table(ru),main="rural urban code",cex.main=3,cex.axis=2,
cex.names=2,ylim=c(0,18000))
barplot(table(popsize), main="county pop",cex.main=3, cex.axis=2,
cex.names=2,ylim=c(0,18000),names="")
text(0.6,9500,paste("pop=2500"),srt=90,cex=2)
text(1.75,9500,paste("pop=10000"),srt=90,cex=2)
text(3.1,9500,paste("pop=20000"),srt=90,cex=2)
text(4.3,9500,paste("pop=100000"),srt=90,cex=2)
text(5.3,9500,paste("pop=250000"),srt=90,cex=2)
text(6.6,9500,paste("pop=1000000"),srt=90,cex=2)


barplot(table(urban), main="",cex.main=3, cex.axis=2,
cex.names=2,ylim=c(0,25000),names=c("nonmetro","metro"))
barplot(table(adjacent), main="",cex.main=3,
cex.axis=2,cex.names=2,ylim=c(0,25000),names=c("adjacent","nonadj
acent"))
```

Cross tab of stage taking into adjacent and urban type

```
xtabs(stage~adjacent+urban)
        urban
adjacent     0     1
       0  2631 11781
       1  2300     0
```

To convert popsize to an ordered categorical variable with 2500 as category 1 up to a million as category 6. Procedure as follows in R.

```
size=1*(popsize==2500)+2*(popsize==10000)+3*(popsize==20000)+4*(p
opsize==100000)+5*(popsize==250000)+6*(popsize==1000000)
table(size)
size
    1     2     3     4     5     6
 1659  4171  1600  2640  2088 16378

text(0.67,2386,"2500")
text(1.90,4782,"10000")
text(3.017,2386,"20000")
text(4.220,3384,"100000")
text(5.480,2885,"250000")
text(6.740,17012,"1000000")

  nohsd hsdip someco bach
1  1060  3274   4485 2177
2  1060  3274   4485 2177
3  1060  3274   4485 2177
```

To compute a measure of education, we use weighed average of the 4 educational variables in R as follows; This was done to find education of each person

```
edu=(1*nohsd+2*hsdip+3*someco+4*bach)/(nohsd+hsdip+someco+bach)

par(mfrow=c(1,2))

size=1*(popsize==2500)+2*(popsize==10000)+3*(popsize==20000)+4*(p
opsize==100000)+5*(popsize==250000)+6*(popsize==1000000)
edu1=mean(edu[size==1])
edu2=mean(edu[size==2])
edu3=mean(edu[size==3])
edu4=mean(edu[size==4])
edu5=mean(edu[size==5])
edu6=mean(edu[size==6])
edu.means=c(edu1,edu2,edu3,edu4,edu5,edu6)
x.vals=1:6
hist(edu,prob=T,nclass=12,main="Histogram  of  Education  Score",
cex.main=1.5, cex.axis=1.5,cex.lab=1.5)
lines(density(edu,bw=0.1))
```

```
plot(size,edu,main="Regression of Education Score on Size",
cex.main=1.5, cex.axis=1.5,cex.lab=1.5)
lines(x.vals,edu.means,col="blue",lwd=2)
```

Violin plot of education and size of the county

```
plot(size,edu)

# Violin Plots
library(vioplot)
x1 <- mtcars$mpg[mtcars$cyl==4]
x2 <- mtcars$mpg[mtcars$cyl==6]
x3 <- mtcars$mpg[mtcars$cyl==8]
vioplot(x1, x2, x3, names=c("4 cyl", "6 cyl", "8 cyl"),
col="gold")
title("Violin Plots of Miles Per Gallon")

x1=edu[size==1]
x2=edu[size==2]
x3=edu[size==3]
x4=edu[size==4]
x5=edu[size==5]
x6=edu[size==6]
vioplot(x1,x2,x3,x4,x5,x6,col="gray")
```

Histogram of county size and educational distribution of each county

```
par(mfrow=c(1,6))
hist(x1,xlab="county size 1", main="Dist of edu")
hist(x2,xlab="county size 2", main="Dist of edu")
hist(x3,xlab="county size 3", main="Dist of edu")
hist(x4,xlab="county size 4", main="Dist of edu")
hist(x5,xlab="county size 5", main="Dist of edu")
hist(x6,xlab="county size 6", main="Dist of edu")
xbar1=mean(edu[size==1])
xbar2=mean(edu[size==2])
xbar3=mean(edu[size==3])
xbar4=mean(edu[size==4])
xbar5=mean(edu[size==5])
xbar6=mean(edu[size==6])
xvals=1:6
xbar=c(xbar1,xbar2,xbar3,xbar4,xbar5,xbar6)
lines(xvals,xbar,col="blue")

par(mfrow=c(1,5))
```

#Code to compute education and race of each patient

```
boxplot(edu~race,col="gray",names=c("white","black"),main="Educat
ion Boxplot by Race",cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
```

#Education and Adjacent (type of rurality)

```
boxplot(edu~adjacent,col="gray",names=c("adjacent","non-adjacent"),main="Boxplot Education by Adjacent", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
```

#Education by the four main 4 Age Groups

```
boxplot(edu~age,col="gray",names=c("18-39","40-49","50-64","65+"),main="Education Boxplot by Age", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
```

#Education by Stage at Diagnosis

```
boxplot(edu~stage,col="gray",names=c("early","late"),main="Education Boxplot by Stage", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)

stagebar1=mean(stage[age==1])
stagebar2=mean(stage[age==2])
stagebar3=mean(stage[age==3])
stagebar4=mean(stage[age==4])
stagebar=c(stagebar1,stagebar2,stagebar3,stagebar4)
ages=c(1,2,3,4)

text(0.65,0.49,"48%")
text(1.82,0.38,"36%")
text(3.03,0.34,"31%")
text(4.22,0.31,"28%")
```

Boxplot and Violin plots of education by race, stage, age, and adjacent and education

```
x1=edu[size==1]
x2=edu[size==2]
x3=edu[size==3]
x4=edu[size==4]
x5=edu[size==5]
x6=edu[size==6]
x=c(x1,x2,x3,x4,x5,x6)

par(mfrow=c(1,5))
boxplot(edu~race,col="gray",names=c("white","black"),main="Edu by Race", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
boxplot(edu~stage,col="gray",names=c("early","late"),main="Edu by Stage", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
boxplot(edu~age,col="gray",names=c("18-39","40-49","50-64","65+"),main="Edu by Age", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
boxplot(edu~adjacent,col="gray",names=c("adjacent","non-adjacent"),main="Edu by Adjacent", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
vioplot(x1,x2,x3,x4,x5,x6,col="gray", cex.lab=1.5,cex.axis=1.5,cex.main=1.5)
title(main="Violpot:Edu by County Size")
```

Percent Diagnosed Late by Age category of the patient

```
barplot(stagebar,ylim=c(0,0.5),col="gray",names=c("18-39","40-
49","50-64","65+"),main="Diagnosed Late by Age")
text(0.65,0.49,"48%")
text(1.82,0.38,"36%")
text(3.03,0.34,"31%")
text(4.22,0.31,"28%")
```

Barplot of Percent Diagnosed at Late Stage by Age of the Patient

```
barplot(stagebar,ylim=c(0,0.5),col="gray",names=c("18-39","40-
49","50-64","65+"),main="Diagnosed Late by Age")
text(0.65,0.49,"48%")
text(1.82,0.38,"36%")
text(3.03,0.34,"31%")
text(4.22,0.31,"28%")
```

Stage by Adjacent

```
p.late.adj=sum(stage[adjacent==0])/length(stage[adjacent==0])
p.late.non=sum(stage[adjacent==1])/length(stage[adjacent==1])
p=c(p.late.adj,p.late.non)
barplot(p,ylim=c(0,0.8),col="gray",names=c("adjacent","non
adjacent"),main="Percentage Diagnosed Late by Adjacent")
text (0.65,0.74,"69%")
text (1.85,0.35,"31%")

stagebar1=mean(stage[adjacent==0])
stagebar2=mean(stage[adjacent==1])
stagebar=c(stagebar1,stagebar2)
adjacent=c(0,1)
barplot(p,ylim=c(0,0.8),col="gray",names=c("adjacent","non
adjacent"),main="Percentage Diagnosed Late by Adjacent")

text(0.65,0.49,"48%")
text(1.82,0.38,"36%")
text(3.03,0.34,"31%")
text(4.22,0.31,"28%")

boxplot(stage~adjacent)

educ=(1*nohsd+2*hsdip+3*someco+4*bach)/(nohsd+hsdip+someco+bach)
```

Kernel Smooth of Stage by Education

```
fit.2=ksmooth(edu,stage,kernel="normal",bandwidth=0.2)
lines(fit.2)
plot(edu,stage,main="Late Diagnosis by Education")
lines(fit.2)
```

Bargraph of Stage by Race.

```
stage.white=sum(stage[race==1])/length(stage[race==1])
stage.black=sum(stage[race==2])/length(stage[race==2])
p.race=c(stage.white,stage.black)
```

```
barplot(p.race,col="gray",names=c("white","black"))

barplot(p.race,col="gray",names=c("white","black"))
barplot(p.race,ylim=c(0,0.5),col="gray",names=c("white","black"),
main="Percent Late by Race")
text (0.65,0.32,"30%")
text (1.88,0.41,"39%")

table (race)
race
    1     2
25743  2793
p.below=popbelow/(popbelow+popabove)
hist(p.below,prob=T)
boxplot(p.below)

Below Poverty and stage at Diagnosis (Late)
plot(p.below,stage,main="Stage by proportion below poverty line")
fit.below.0.05=ksmooth(p.below,stage,kernel="normal",bandwidth=0.
05)
lines(fit.below.0.05)
```

Model Fit Trails

1. **Stage taking into account age + race + family**

```
fit.trial.1=glm(stage~age+race,family=binomial)
summary(fit.trial.1)  ----(disregard age not considered as a
factor)
```

2. **Fit 2---Age as a factor**

```
age=as.factor(age)
fit.trial.2=glm(stage~age+race,family=binomial)
summary(fit.trial.2)
fit.trial.2=glm(stage~age+race,family=binomial)
summary(fit.trial.2)
```

3. **Fit 3----Add population below the poverty line**

```
fit.trial.3=glm(stage~age+race+p.below,family=binomial)
summary(fit.trial.3)
```

4. **Fit 4**

```
fit.trial.4=glm(stage~age+race+p.below+edu+adjacent,family=binomi
al)
summary(fit.trial.4)

Call:
glm(formula = stage ~ age + race + p.below + edu + adjacent,
    family = binomial)
```

5. **Fit 5**

```
fit.trial.5=glm(stage~age+race+edu,family=binomial)
```

```
summary(fit.trial.5)
Error in summary(fit.trial5) : object 'fit.trial5' not found
summary(fit.trial.5)
```

6. Fit 6

```
fit.trial.6=glm(stage~age+race+edu+p.head,family=binomial)
summary(fit6)
Error in summary(fit6) : object 'fit6' not found

fit.trial.6)
```

SECOND ORDER MODELS

```
tina=read.table("G:/fresen cancer 120303 no nines.txt",header=T)
attach(tina)
educ=(1*nohsd+2*hsdip+3*someco+4*bach)/(nohsd+hsdip+someco+bach)
p.below=popbelow/(popbelow+popabove)
age=as.factor(age)
race=as.factor(race)
adjacent=as.factor(adjacent)
```

library(MASS) **------**

```
scope=list(upper=~(age+race+pbelow+edu+adjacent+headf)^2,lower=~1
,data=tina)
fit.0=glm(stage~1,family=binomial,data=tina) # null
fit.f=stepAIC(fit.0,scope,direction="both")
```

library(MASS)

```
scope=list(upper=~(age+race+pbelow+edu+adjacent+headf)^2,lower=~1
,data=tina)
 fit.0=glm(stage~1,family=binomial,data=tina) # null
 fit.f=stepAIC(fit.0,scope,direction="both")
```

Start: AIC=35335.06
stage ~ 1

	Df	Deviance	AIC
+ age	1	35107	35111
+ race	1	35249	35253
+ pbelow	1	35324	35328
+ edu	1	35324	35328
+ headf	1	35329	35333
<none>		35333	35335
+ adjacent	1	35333	35337

Step: AIC=35110.62
stage ~ age

	Df	Deviance	AIC
+ race	1	35039	35045
+ pbelow	1	35093	35099
+ edu	1	35093	35099
+ headf	1	35102	35108

```
<none>            35107 35111
+ adjacent  1     35106 35112
- age       1     35333 35335
```

Step: AIC=35044.89
stage ~ age + race

```
           Df Deviance   AIC
+ edu       1    35017 35025
+ pbelow    1    35031 35039
+ age:race  1    35036 35044
+ adjacent  1    35036 35044
<none>           35039 35045
+ headf     1    35038 35046
- race      1    35107 35111
- age       1    35249 35253
```

Step: AIC=35024.7
stage ~ age + race + edu

```
           Df Deviance   AIC
+ age:race  1    35014 35024
<none>           35017 35025
+ race:edu  1    35016 35026
+ headf     1    35017 35027
+ age:edu   1    35017 35027
+ pbelow    1    35017 35027
+ adjacent  1    35017 35027
- edu       1    35039 35045
- race      1    35093 35099
- age       1    35232 35238
```

Step: AIC=35023.54
stage ~ age + race + edu + age:race

```
           Df Deviance   AIC
<none>           35014 35024
- age:race  1    35017 35025
+ race:edu  1    35013 35025
+ age:edu   1    35013 35025
+ headf     1    35013 35025
+ pbelow    1    35014 35026
+ adjacent  1    35014 35026
- edu       1    35036 35044
> help("stepAIC")
```

Start: AIC=35335.06
stage ~ 1

```
           Df Deviance   AIC
+ age       1    35107 35111
+ race      1    35249 35253
```

```
+ p.below   1    35324 35328
+ educ      1    35324 35328
<none>           35333 35335
+ adjacent  1    35333 35337
```

Step: AIC=35110.62
stage ~ age

```
           Df Deviance   AIC
+ race      1    35039 35045
+ p.below   1    35093 35099
+ educ      1    35093 35099
<none>           35107 35111
+ adjacent  1    35106 35112
- age       1    35333 35335
```

Step: AIC=35044.89
stage ~ age + race

```
           Df Deviance   AIC
+ educ      1    35017 35025
+ p.below   1    35031 35039
+ adjacent  1    35036 35044
<none>           35039 35045
- race      1    35107 35111
- age       1    35249 35253
```

Step: AIC=35024.7
stage ~ age + race + educ

```
           Df Deviance   AIC
<none>           35017 35025
+ p.below   1    35017 35027
+ adjacent  1    35017 35027
- educ      1    35039 35045
- race      1    35093 35099
- age       1    35232 35238
```

SECOND ORDER MODELS
Start: AIC=35335.06
stage ~ 1

```
           Df Deviance   AIC
+ age       1    35107 35111
+ race      1    35249 35253
+ p.below   1    35324 35328
+ educ      1    35324 35328
<none>           35333 35335
+ adjacent  1    35333 35337
```

Step: AIC=35110.62
stage ~ age

```
           Df Deviance   AIC
+ race      1    35039 35045
+ p.below   1    35093 35099
+ educ      1    35093 35099
<none>           35107 35111
+ adjacent  1    35106 35112
- age       1    35333 35335
```

Step: AIC=35044.89
stage ~ age + race

```
           Df Deviance   AIC
+ educ      1    35017 35025
+ p.below   1    35031 35039
+ age:race  1    35036 35044
+ adjacent  1    35036 35044
<none>           35039 35045
- race      1    35107 35111
- age       1    35249 35253
```

Step: AIC=35024.7
stage ~ age + race + educ

```
            Df Deviance   AIC
+ age:race   1    35014 35024
<none>            35017 35025
+ race:educ  1    35016 35026
+ age:educ   1    35017 35027
+ p.below    1    35017 35027
+ adjacent   1    35017 35027
- educ       1    35039 35045
- race       1    35093 35099
- age        1    35232 35238
```

Step: AIC=35023.54
stage ~ age + race + educ + age:race

```
            Df Deviance   AIC
<none>            35014 35024
- age:race   1    35017 35025
+ race:educ  1    35013 35025
+ age:educ   1    35013 35025
+ p.below    1    35014 35026
+ adjacent   1    35014 35026
- educ       1    35036 35044
```

Start: AIC=35335.06
stage ~ 1

```
           Df Deviance   AIC
+ age       1    35107 35111
+ race      1    35249 35253
+ p.below   1    35324 35328
+ educ      1    35324 35328
<none>           35333 35335
+ adjacent  1    35333 35337

Step:  AIC=35110.62
stage ~ age

           Df Deviance   AIC
+ race      1    35039 35045
+ p.below   1    35093 35099
+ educ      1    35093 35099
<none>           35107 35111
+ adjacent  1    35106 35112
- age       1    35333 35335

Step:  AIC=35044.89
stage ~ age + race

           Df Deviance   AIC
+ educ      1    35017 35025
+ p.below   1    35031 35039
+ adjacent  1    35036 35044
<none>           35039 35045
- race      1    35107 35111
- age       1    35249 35253

Step:  AIC=35024.7
stage ~ age + race + educ

           Df Deviance   AIC
<none>           35017 35025
+ p.below   1    35017 35027
+ adjacent  1    35017 35027
- educ      1    35039 35045
- race      1    35093 35099
- age       1    35232 35238

fit.are=glm(stage~age+race+educ,family=binomial)
summary(fit.are)
Call:
glm(formula = stage ~ age + race + educ, family = binomial)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2955  -0.8550  -0.8049   1.4399   1.6628

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept)  0.50803    0.15921   3.191  0.00142 **
age2        -0.44638    0.06431  -6.941 3.90e-12 ***
age3        -0.66576    0.06070 -10.968  < 2e-16 ***
age4        -0.80891    0.06032 -13.411  < 2e-16 ***
race2        0.36765    0.04178   8.800  < 2e-16 ***
educ        -0.26301    0.05654  -4.652 3.29e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 35333  on 28535  degrees of freedom
Residual deviance: 34998  on 28530  degrees of freedom
AIC: 35010
Delta AIC = 335 df 5 p-value=0
```

Interpretation
Diagnostics
Graphs

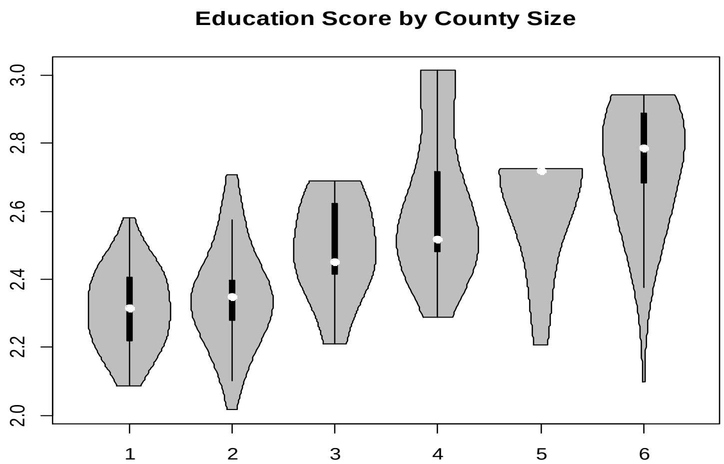
```
fit.arse=gam(stage~age+race+s(educ),family=binomial)
> summary(fit.arse)

Call: gam(formula = stage ~ age + race + s(educ), family =
binomial)
Deviance Residuals:
    Min      1Q  Median      3Q     Max
-1.2906 -0.8629 -0.8118  1.4300  1.6694

(Dispersion Parameter for binomial family taken to be 1)

    Null Deviance: 35333.06 on 28535 degrees of freedom
Residual Deviance: 34994.86 on 28527 degrees of freedom
AIC: 35012.86

Number of Local Scoring Iterations: 4

DF for Terms and Chi-squares for Nonparametric Effects

            Df Npar Df Npar Chisq P(Chi)
(Intercept)  1
age          3
race         1
s(educ)      1       3     2.6707 0.4452
```

Interpretation of fit.are

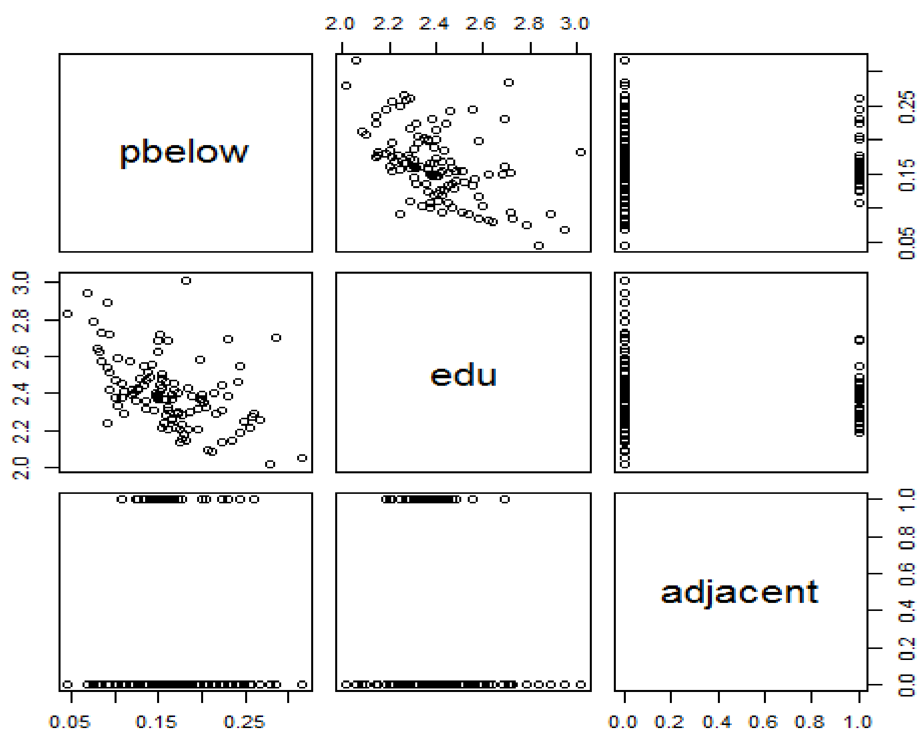
```
e=seq(2,3,length=100)
p.1.w=1/(1+exp(-(0.50803-0.26301*e)))
p.2.w=1/(1+exp(-(0.50803-0.44638-0.26301*e)))
p.3.w=1/(1+exp(-(0.50803-0.66576-0.26301*e)))
p.4.w=1/(1+exp(-(0.50803-0.80891-0.26301*e)))
```

```
p.1.b=1/(1+exp(-(0.50803+0.36765-0.26301*e)))
p.2.b=1/(1+exp(-(0.50803-0.44638+0.36765-0.26301*e)))
p.3.b=1/(1+exp(-(0.50803-0.66576+0.36765-0.26301*e)))
p.4.b=1/(1+exp(-(0.50803-0.80891+0.36765-0.26301*e)))
par(mfrow=c(1,2))

plot(e,p.1.b,type="l",lwd=3,ylim=c(0,1),xlim=c(2,3.2),xlab="count
y education measure",ylab="Probability of late
detection",main="blacks")
lines(e,p.2.b, lwd=3)
lines(e,p.3.b, lwd=3)
lines(e,p.4.b,lwd=3)
text(3.1,0.518,"18-39")
text(3.1,0.414,"40-49")
text(3.1,0.360,"50-64")
text(3.1,0.322,"64+")

plot(e,p.1.w,type="l",lwd=3,ylim=c(0,1),xlim=c(2,3.2),xlab="count
y education measure",ylab="",main="whites")
lines(e,p.2.w,lwd=3)
lines(e,p.3.w,lwd=3)
lines(e,p.4.w,lwd=3)
text(3.1,0.436,"18-39")
text(3.1,0.325,"40-49")
text(3.1,0.284,"50-64")
text(3.1,0.252,"64+")

plot(edu,headc)
plot(edu,headf)
cor(headf,headc)
[1] 0.9977235
```

D. VIOLIN PLOT OF EDUCATION SCORE BY COUNTRY SIZE



E. SCATTER PLOT PERCENT BELOW POVERTY, EDUCATION AND ADJACENT

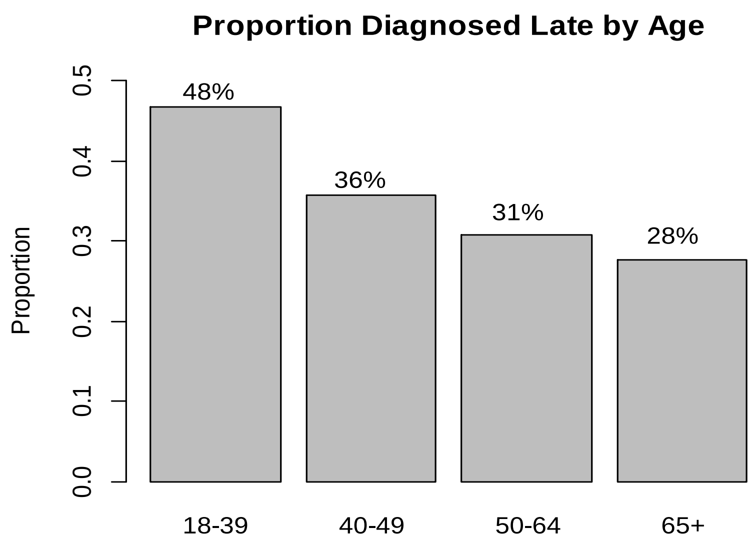


F. STEPWISE LOGISRIC REGRESSION FOR AIC SELECTION

	Predictors	Estimate	SE	Z-Value	P-Value
Model I	Intercept	-0.52736	0.07427	-7.201	1.24e-12 ***
	Age 2	-0.44849	0.06429	-6.976	3.03e-12 ***
	Age 3	-0.66348	0.06067	-10.936	< 2e-16 ***
	Age 4	-0.80359	0.06028	-13.331	< 2e-16 ***
	Race	0.34390	0.04144	8.298	< 2e-16 ***
Model II	Intercept	0.59965	0.07889	-7.601	2.93e-14 ***
	Age 2	-0.44685	0.06430	-6950	3.65e-12 ***
	Age 3	-0.66459	0.06068	-10.952	< 2e-16 ***
	Age 4	-0.80721	0.06030	-13.386	< 2e-16 ***
	Race	0.33100	0.04171	7.935	2.10e-15 ***
	Pbelow	0.65253	0.23969	2.722	0.00648 **
Model III	Intercept	0.16781	0.21817	0.769	0.441796
	Age 2	-0.44645	0.06431	-6.942	3.87e-12 ***
	Age 3	-0.66573	0.06070	-10.968	< 2e-16 ***
	Age 4	-0.80876	0.06032	-13.407	< 2e-16 ***
	Race	0.36954	0.04298	8.598	< 2e-16 ***
	Pbelow	-0.05712	0.30544	-0.187	0.851653
	Edu	-0.27132	0.07193	-3.772	0.000162 ***
Model IV	Intercept	0.172811	0.230954	0.748	0.454310
	Age 2	-0.446455	0.064315	-6.942	3.87e-12 ***
	Age 3	-0.665697	0.060701	-10.967	< 2e-16 ***
	Age 4	-0.808710	0.060329	-13.404	< 2e-16 ***
	Race	0.369378	0.043047	8.581	< 2e-16 ***
	Pbelow	-0.057983	0.305711	-0.190	0.849570
	Edu	-0.272995	0.076253	-3.580	0.000343 ***
	Adjacent	-0.002807	0.042510	-0.066	0.947354

Significant codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

G. PROPORTION DIAGNOSED LATE BY AGE



BIBLIOGRAPHY

Aagon, R., Mogan, J., Wong, J. H. & Lum, S. (2011). Potential impact of USPSTF recommendation on early diagnosis of breast cancer. *Annals of Surgical Oncology, 18*, 3137-3142.

Abeloff M.D., Wolff A.C., Wood W.C. et al. Cancer of the breast: In: Abeloff M.D., ed. Clincial Oncology. 3rd ed. Philadelphia: Elsevier Churchill Livingstone; (2004): 2369-2450.

Aday, L. A. & Andersen, R. (1974). A framework for the study of access to medical care. *Health Services Research,* 9, 208–20.

Adler, N. E., Boyce, T. & Chesney, M. A. et al., (1994). Socioeconomic status and health: the challenge of the gradient. *Am Psychol,* 49, 15–24.

Adler, N. E. & Newman, K. 2002. Socioeconomic disparities in health: Pathways and policies. *Health Affairs,* 21 (2), 6076.

Agresti, A. (2002). *Categorical data analysis.* Hoboken, New Jersey: John Wiley & Sons Inc.

Altekruse, S. F., Kosary, C. L., Krapcho, M. et al, eds. (2010). Surveillance, Epidemiology, and End Results Cancer Statistics Review, 1975-2007.

Alvarez, P., L., L., Guerra-Yi, M., E., Faes, C., Alvarez, Y., G & Molenberghs, G. (2009). Spatial analysis of breast and cervical cancer incidence in small geographical areas in Cuba, 1999–2003. *European Journal of Cancer and Prevention,* 18, 395-403.

American Cancer Society. (2012). Cancer facts and figures 2012. Atlanta, GA: American Cancer Society, Inc.

American Cancer Society & Livestrong (2010). About the global cost of cancer report. Retrieved from http://www.cancer.org/acs/groups/content/@internationalaffairs/documents/document/acspc-026203.pdf

American Joint Committee on Cancer & American Cancer Society. (2009). *Cancer Staging Manual* (7th ed). Chicago, IL: Springer

Amornsiripanitch, N., Shaolin, H., Campa, M. J., Frank, M. M., Gottlin, E. B. & Patz, E. F. (2010). Complete factor H autoantibodies are associated with early stage NSCLC. *Clinical Cancer Research*, 16 (12), 3226-32231.

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: Does it matter? *Journal of Health and Social Behavior,* 36, 1-10.

Ansari, A., Carson, N. J., Ackland, M. J., Vaughan, L. & Serraglio, A. (2003). A public health model of the social determinants of health. *Social and Preventive Medicine,* 48 (4), 242-251.

Arcury, T. A., Gesler, W. M. & Preisser, J.S. (2005). The effects of geography and spatial behaviour on health care utilization among the residents of a rural region. Health Services Research 40 (1), 135–55.

Baldwin, L. M., Cai, Y., Larson, E. H., Dobie, S. A., Wright, G. E., Goodman, D. C., Hart, L. G. (2008). Access to cancer services for rural colorectal cancer patients. *The Journal of Rural Health, 24*(4), 390-399.

Barry, J. & Breen, N. (2005). The importance of place of residence in predicting late-stage diagnosis of breast or cervical cancer. *Health and Place,* 11, 15-29.

Beedasy, J., Ramloll, R., Stamm, B. H. (2008). Terrorism. In Gary Goreham (*Ed*.), Encyclopedia of Rural America: The Land and People. Vol 2. (2nd ed, pp. 970-973). New York: Grey House Publishing, Inc.

Bethesda, MD: National Cancer Institute. Retrieved from http://seer.cancer.gov/csr/1975_2007/

Bettencourt, B. A., Schlegel, R. J., Talley, A. E. & Molix, L. A. (2007). The breast cancer experience of rural women: a literature review. *Psycho-oncology,* 16 (10), 875-887

Blackman, D. J. & Masi, C. M. (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.

Blazer, D., Landerman, L., Fillenbaum, G. (1995). Health services access and use among older adults in North Carolina: urban versus rural residents. *American Journal of Public Health,* 85 (10), 1384-90.

Bradley, C. J., Given, C. W. & Roberts, C. (2002). Race, socioeconomic status and breast cancer treatment and survival. *Journal of National Cancer Institute,* 94 (7), 490-496.

Brameld, K. J., & Holman, C. D. J. (2006). The effect of locational disadvantage on hospital utilisation and outcomes in Western Australia. *Health and Place,* 12, 490-502.

Booth, C. M., Gavin, L., Zhang-Salomons, J & Mackillop, W. J. (2010). The impact of socioeconomic status on stage of cancer at diagnosis and survival. *Cancer,*4160-4167.

Breast Cancer Organization. (2011). Stages of breast cancer. Retrieved from http://www.breastcancer.org/symptoms/diagnosis/staging.jsp

Brewster, D. H., Thomson, C.S., Hole, D. J., Black, R. J., Stroner, P. L. & Gillis, C. R. (2001). Relation between socioeconomic status and tumour stage in patients with breast, colorectal, ovarian, and lung cancer: results from 4 national, population based studies. *BMJ,* 322, 830-831.

Brown, D. L. & Schafft, K. A. (2011). Rural people and communities in the 21st century: Resilence and transformation. Cambridge, UK: Polity Press.

Brustrom, J. E. & Hunter, D. C. (2001). Going the distance: How far will women travel to undergo free mammography. *Military Medicine,* 166 (4), 347-349.

Bushy, A. (2000). Special and at-risk populations. In Bushy, A. (Ed.), Orientation to nursing in the rural community (pp. 73-85.). Thousand Oaks, CA: Sage Publications, Inc.

Byers, T. E., Wolf, H. J. & Bauer, K. R. et al., (2008). The impact of socioeconomic status on survival after cancer in the United States: findings from the National Program of Cancer Registries Patterns of Care Study. *Cancer,* 113, 582-591.

Campbell, B., Staley, D. & Matas, M. (1991). Who misses appointments? An empirical analysis. *Cancer Journal of Psychiatry,* 36, 223-225.

Campbell, N. C., Elliot, A. M., Sharp, L., Ritchie, L.D., Cassidy, J. & Little J. (2000). Rural factors and survival from cancer: Analysis of Scottish cancer registrations. *British Journal of Cancer*, 82(11):1863-1866.

Campbell, R.T., Li, X., Dolecek, T.A., Barrett, R.E., Weaver, K.E. & Warnecke. (2009). Economic, racial and ethnic disparities in breast cancer in the US: Towards a more comprehensive mode. *Health and place*, 15, 870-879.

CDC (2011). National Breast and Cervical Cancer Early Detection Program. Retrieved from http://www.cdc.gov/cancer/nbccedp/legislation/law.htm

Celaya, M. O,, Rees, J. R., Gibson, J. J., Riddle, B. L. & Greenberg, E. R. (2006). Travel distance and season of diagnosis affect treatment choices for women with early-stage breast cancer in a predominantly rural population (United States). *Cancer Causes Contro*l, 17, 851-856.

Chan, L., Hart, L. G., & Goodman, D. C. (2006). Geographic access to health care for rural Medicare beneficiaries. *The Journal of Rural Health, 22*(2), 140-146.

Clegg, L. X., Reichman, M. E., & Miller, B. A. et al., (2009). Impact of socioeconomic status on cancer incidence and stage at diagnosis: Selected findings from the surveillance, epidemiology, and end results: National Longitudinal Mortality Study. *Cancer Causes Control,*20, 417-435.

Coldman, A., Phillips, N., Warren, L. & Kan, L. (2007). Breast cancer mortality after screening mammography in British Columbia women. *Int J Cancer,* 120, 1076-1080.

Cohen, M. (1994). Impact of poverty on women's health. *Canadian Family Physician,* 40, 949-958.

Crandall, M. S. & B. A. Weber (2004) Local Social and Economic Conditions, Spatial Concentrations of Poverty, and Poverty Dynamics. *American Journal of Agricultural Economics,* 86**,** 1276-1281.

Cromley, R.G. & Cromley, E. K. (2009). Chloropleth map legend design for visulizting community health disparities. *International Journal of Health Geographics.* Vol: 8:52

Davis, C., Girgis, A., Williams, P., & Beeney, L. (1998). Needs assessment of rural and remote women travelling to the city for breast cancer treatment. *Australian and New Zealand Journal of Public Health, 22*(5), 525-527.

Devasa, S. S. & Diamond, E. L. (1980). Association of breast cancer and cervical cancer incidence with income and education among whites and black. *Journal of National Cancer Institute,* 65, 515-528.

DHSS (2010). Physicians in short supply in rural Missouri areas. Retrieved from http://www.columbiamissourian.com/stories/2010/01/05/health-care

DHSS (2011). Show me healthy women. Retrieved April 7, 2011, from http://health.mo.gov/living/healthcondiseases/chronic/showmehealthywomen/pdf/ providerManual.pdf

Duffy, S. W., Tabar, L., Chen, H. H., et al. (2002). The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. *Cancer,* 95, 458-469.

Eberhardt, M.S., Ingram, D.D., & Makus, D.M., et al. (2001). Urban and rural health chartbook. Health, United States, 2001. Hyattsville, MD: National Center for Health Statistics.

Eggebeen, D. J., A. R. Snyder, W M. (1996). Single father families in demographic perspective. *Journal of Family Issues,* 17, 441–465.

Elliott, T. E., Elliott, B. A., Renier, C.M. & Haller, I.V. (2004). Rural-urban differences in cancer care: Results from the Lake Superior Rural Cancer Care Project. *Minnesota Medicine,* 87 (9), 44-50.

Evans, W. P., Poston, G. & Poston, C. (2012) Breast cancer screening: Success and challenges. CA *Cancer Journal for Clinicians,* 62 (1), 5-9.

Fein, O. (1995). The influence of social class on health status: American and British research on health inequalities. *Journal General Internal Medicine,* 10, 577–86.

Feinstein, J. S. (1993). The relationship between socioeconomic status and health: a review of the literature. *Milbank Quarterly,* 71, 279–322.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C. & Parkin, D.M. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer,* 127, 2893-2917.

Field, K. S. & Briggs, D. (2001). Socio-economic and locational determinants of accessibility and utilization of primary health-care. *Health and Social Care in the Community, 9* (5), 294–308.

Flora, C.B. & Flora, J. L. (2008). *Rural communities: Legacy and change (3rd ed.).* Boulder, CO: Westview Press

Folland, S., Goodman, A. C., & Stano, M. (2001). The economics of health and health care. Upper Saddle River, N.J: Prentice-Hall, Inc.

Fortney, J., Rost, K., Zhang, M., Warren, J. (1999). The impact of geographic accessibility on the intensity and quality of depression treatment. *Medical Care,* 37, 884–893

Fryer, G.E., Drisko, J., Krugman, R.D., Vojir, C.P., Prochazka, A., Miyoshi, T.J., et al. (1999). Multi-method assessment of access to primary medical care in rural Colorado. Journal of Rural Health, 15, 113-121.

Gold, M. 1998. “Beyond coverage and supply: Measuring access to health care in today’s market. *Health Services Research,* 33 (3 Part II), 625-652.

Gorey, K. M., Holoway, E. J., Laukkanen, E., Fehrienger, G. & Richter, N. L. (1998). Association between socioeconomic status and cancer incidence in Toronto, Ontario: Possible confounding of cancer mortality by incidence and survival.

Cancer Prevention and Control, 2, 236-241.

Gray, R. E., James, P., Manthorne, J., Gould, J., & Fitch, M. I. (2004). A consultation with Canadian rural women with breast cancer. *Health Expectations, 7* (1), 40-50.

Greenlee, R. T., & Howe, H. L. (2009). County-level poverty and distant stage cancer in the United States. *Cancer Causes and Control, 20*(6), 989-1000.

Health Management Associate Inc., (2011) in Issues in Missouri health care: The state of health in Missouri: Coverage, access and health status. Retrieved from http://www.mffh.org/mm/files/issue_brief_02_Diagnosis_What_Data_says.pdf

Health People 2020 (2011). Cancer. Retrieved from http://www.healthypeople.gov/2020/topicsobjectives2020/overview.aspx?topicid=5

Hendrick, R. E. & Helve, M. A. (2011). United States Preventive Services Task Force screening mammography recommendation science ignored. American *Journal of Roentgenology,* 196, W112-W116.

Henry, K. A., Boscoe, F. P., Johnson, C. J., Goldber, D. W., Sherman, R. & Cockburn M. (2011). Breast cancer stage at diagnosis: Is travel time important? *Journal of Community Health,* 36, 933-942.

Henry, K. A., Sherman, R. Roche, L. A. (2009). Colorectal cancer stage at diagnosis and socioeconomic characteristics in New Jersey. *Health and Place*, 15, 505-513.

Huang, B., Dignan, D., Han, D. & Johnson, O. (2009). Does distance matter" Distance to mammography facilities ad stage at diagnosis of breast cancer in Kentucky. *Journal of Rural Health,* 366-371.

Institute of Medicine (2002). Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: National Academy of Sciences.

Isserman, A. M. (2005). In the national interest: Defining rural and urban correctly in research and public policy. *International Regional Science Review, 28*(4), 465.

Jack, R. H., Gulliford, M. C., Ferguson, J., & Møller, H. (2006). Explaining inequalities in access to treatment in lung cancer. *Journal of Evaluation in Clinical Practice, 12*(5), 573-582

Johnson, K.S., Elbert-Avila, K.I., and Tulsky, J.A. (2005). The influence of spiritual beliefs and practices on the treatment preferences of African Americans: A review of the literature. *Journal of the American Geriatric Society,* 53 (4), 711-719.

Joliffe, D. (2005). Rural poverty at a glance. Washington, DC: USDA Economic Research Service. Retrieved from http://www.ers.usda.gov/publications/rdrr100/rdrr100.htm

Jolliffe, D. (2004). Rural poverty at a glance (RDRR No. 100). Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Jones, A., P., Haynes, R., Sauerzapf, V., Crawford, S., M., Zhao, H. & Forman, D. (2008). Travel times to health care and survival from cancer in Northern England. *European Journal of Cancer, 44*, 269-274.

Jordan, H., Roderick, P., Martin, D. & Barnett, S. (2004). Distance, rurality and the need for care: access to health services in South West England. *International Journal of Health Geographics*, 3: 21

Katirai, M. (2012). Access to health care and colorectal cancer in Kentucky. *International Journal of Humanities and Social Science,* 1 (18).

Kerlikowske, K., Grady, D., Rubin, S.M., Sandrock, C., Ernster, V. (1995). Efficacy of screening mammography: A meta-analysis. *JAMA,* 273, 149-154.

Khan, M. A., Walley, J. D., Witter, S. N., Imran, A. & Safdar, N. (2002). Costs and cost-effectiveness of different DOTS strategies for treatment of tuberculosis in Pakistan. *Health Policy and Planning,* 17, 178–86.

Kheireland, E. A. H., Boggs, J. M. E., Curran, C., Glynn, R., Dooley, C. & Sweeney, K. et al., (2011). Younger age as a prognostic indicator in breast cancer: A cohort study. *BMC Cancer,* 11, 383.

Komen for the Cure (2012). Racial facts for life: Racial and ethnic differences. Retrieved from http://ww5.komen.org/uploadedFiles/Content_Binaries/806-373a.pdf

Kopans, D. B. (2011). Just the facts: Mammography saves lives with little if any radiation risk to the nature breast. *Health Physics,* 101 (5) 578-582.

Larson, S.L. and, Fleishman, J. (2003). Rural-urban differences in usual source of care and ambulatory service use: analyses of national data using urban influence codes. *Medical Care,* 41 (7), Supplement: III-65-III-74.

Lee, C. H., Dershaw, D. D., Kopans, D., Evans, P., Monsees, B., Monticciolo, D., et al. (2010). Breast cancer screening with imaging: recommendations from the Society of Breast Imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *Journal of the American College of Radiology,* 7, 18–27.

Lee, P. & Paxman, D. (1997). Reinventing public health. *Annual Review of Public Health, 18*, 135.

Liburd, L. C., Giles, H. W., Mensah, G. A. (2006). Looking through a glass, darkly: eliminating health disparities. *Preventing Chronic Disease Public Health Research,* Practice, and Policy, 3 (3), 13.

Lin, G., Allan, E.D., & Penning, J.M. (2002). Examining distance effects on hospitalization using GIS: A study of three health regions in British Columbia, Canada. *Environment and Planning,* 34, 2037-2053.

Liu, M., J., Hawk, H., A., Getshman, S., T., Smith, S., M., Karacek, R., Woodfood., M., L. & Ayanian, J., Z., (2005). The effects of National Breast and Cervical Cancer Early Detection Program on social disparities in breast cancer diagnosis and treatment in Massachusetts. *Cancer Causes and Control,* 16, 17-33.

Lobao, L., Rulli, J., & Brown, L. (1999). Macrolevel theory and local-level inequality: Industrial structure, institutional arrangements, and the political economy of redistribution, 1970 and 1990. *Annals of the Association of American Geographers, 89*(4), 571-601.

Lovett, A., Haynes, R., Sunnenberg, G., & Gale, S. (2002). Car travel time and accessibility by bus to general practitioner services: A study using patient registers and GIS. *Social Science & Medicine*, 55, 97-111.

Makinen, M., Waters, H., Rauch, M. et al., (2000). Inequalities in health care use and expenditures: empirical data from eight developing countries and countries in transition. *Bulletin of the World Health Organization,* 78, 55–74.

Mackilliop, W.J., Zhang-Salomons, J, Boyd, C. J. & Groome, P. A. (2000). Association between community income and cancer incidence in Canada and the United States. *Cancer,* 89, 901-912.

Mackinnon, J., Duncan, R., Huang, Y., Lee, D., Fleming, L., Voti, L., Rudolph, M. & Wilkinson, J. (2007). Detecting an association between socioeconomic status and late stage breast cancer using spatial analysis and area-based measures. *Cancer Epidemiology, Biomarkers & Prevention,* 16 (4), 756-762.

Maheswaran, R., Pearson, T., Jordan, H. & Black, D. (2006). Socio-economic deprivation, travel distance, location of service, and uptake of breast cancer screening in North Derbyshire, UK. *J Epidemiol Commun Health*, 60, 208-212.

McCormack, V.A., & Boffetta, P. (2011). Today's lifestyles, tomorrow's cancers: Trends in lifestyles risk factors for cancer in low and middle income countries. *Annals of Oncology,* Retrieved from http://annonc.oxfordjournals.org/content/early/2011/03/03/annonc.mdq763.full.pdf

McElroy, J., Remington, P., Gangnon, R., Hariharan, L. & Andersen, L. (2006). Identifying geographic disparities in the early detection of breast cancer using a geographic information system. *Preventing Chronic Disease* 3(1), 1-8.

McLafferty, S., & Wang, F. (2009). Rural reversal? *Cancer,* 115, 2755–2764.

Mechanic, D. (1978). Medical Sociology: A comprehensive text (2nd ed.). New York, NY: Free Press.

Meliker, J. R., Jacquez, G., M., Goovaerts, P., Copeland, G. & Yassine, M. (2009). Spatial cluster analysis of early stage breast cancer: A method for public health practice using cancer registry data. *Cancer Causes Control,* 20, 1061-1069.

Merkin, S. S., Stevenson, L., & Powe, N. (2002). Geographic socioeconomic status, race, and advanced stage breast cancer in New York City. *American Journal of Public Health*, 92 (1), 64-70.

Meersman, S. C., Breen, N., Pickle, L. W., Meissner, H. I., & Simon, P. (2009). Access to mammography screening in a large urban population: A multi- level analysis. *Cancer Causes and Control,* 20, 1469–1482.

Missouri Census Data Center (2010). Ten things to know about urban vs. rural. Retrieved from http://mcdc.missouri.edu/TenThings/urbanrural.shtml

Missouri DHSS & Department of Economic Development (2010). Rural health plan. Retrieved from http://health.mo.gov/living/families/ruralhealth/pdf/ruralhealthplan.pdf

Missouri DHSS (2007). Community Data Profiles, 2007 County-Level Study

Missouri DHSS (2010). The burden of cancer in Missouri: A comprehensive analysis and plan. Retrieved from http://health.mo.gov/living/healthcondiseases/chronic/chronicdisease/cancerburdenreport.pdf

Mitchell, A. (1999). *The ESRI guide to GIS analysis*: ESRI Press.

Mojtabia, R., Olfson, M. & Mechanic, D. (2002) Perceived need and help seeking in adults with mood, anxiety, or substance use disorders. *Archives of General Psychiatry,* 59, 77-84.

Muller, I., Smith, T., Mellor, S., Rare, L. & Genton, B. (1998). The effect of distance from home on attendance at a small rural health centre in Papua New Guinea. *International Journal of Epidemiology,* 27, 878-884.

Nelson, H. D., Tyne, K., Naik, A., Bougatsos, C., Chan, B. K., Humphrey, L. (2009). U.S. Preventive Services Task Force. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med,* 151, 727-737.

NoorAli, R., Luby, S. & Rahbar, M. H. (1999). Does use of a government service depend on distance from the health facility? *Health Policy Plan,* 14, 191-7.

Noor, A. M., Alegana, V. A., Gething, P. G. & Snow, R. W. (2009). A spatial national health facility database for public health section planning in Kenya in 2008. *International Journal of Health Geographics,* 8, 13.

Office of Management and Budget. (2000). Standards for defining metropolitan and micropolitan statistical areas: Notice. *Federal Register* 65: 82228-38.

Onega, T., Duell, E. J., Shi, X., Wang, D., Demidenko, E. & Goodman, D. (2008). Geographic access to cancer care in the U.S. *Cancer,* 112 (4), 909-918.

Onwujekew, B. S. C. & Uzochukwu, O. E. (2004). Socio-economic differences and health seeking behaviour for the diagnosis and treatment of malaria: a case study of four local government areas operating the Bamako initiative programme in south-east Nigeria. *International Journal of Equity Health*, 3: 6.

Oppong, J. R. & Hodgso, M.J. (2005). Spatial Accessibility to Health care facilities in Suhum District, Ghana. *The Professional Geographer*, 46(2), 1999-209.

Otto, S. J., Fracheboud, J., Looman, C. W. N., Broeders, M. J. M., Boer, R., Hendriks, J. N. H. C. L. et al., (2003). Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet*, 361, 411– 417; 2003.

Owen, K. K., Obregón, E. J., & Jacobsen, K. H. (2010). A geographic analysis of access to health services in rural Guatemala. [10.1016/j.inhe.2010.03.002]. *International Health, 2*(2), 143-149

Parsons, T. (1951). The social system. Glencoe, IL: Free Press

Peipins, L. A., Graham, S., Young, B., Lewis, B., Foster, S., Flanagan, B & Dent, A. (2011). Time and distance barriers to mammography facilities in the Atlanta Metropolitan Area. *Journal of Community Health,* 36, 675-683.

Penchansky, R. & Thomas, J. W. (1981). The Concept of Access. *Medical Care,* 19(2), 127-140.

Perry, B. & Gesler, W. (2000). Physical access to primary heath in Andean Bolivia. *Social Science and Medicine*, 50 (9), 1177-1188

Peters. H. D., Garg, A., Bloom, G., Walker, D.G., Brieger, W. R. & Rahman, H. R. (2008). Poverty and access to health care in developing countries, Annual N.Y. Academy of Sciences. 1136:161-171

Pierce, R. P., Williamson, H. A. & Kruse, R. L. (1998). Distance, use of resources, and mortality among rural Missouri residents with acute myocardial infarction. *Journal of Rural Health,* 14 (1), 28–35.

Punglia, R. S., Weeks, J. C., Neville, B. A., Earle, C.C. (2000). Effect of distance to radiation treatment facility on use of radiation therapy after mastectomy in elderly women. *International Journal Radiat Oncol, Biol, Phys*., 66:56-63.

Rahman, S. & Smith, D. K. (2000). Use of location-allocation models in health service development planning in developing nations. *European Journal of Operational Research,* 53(3), 437-452.

Rupasingha, A. & S. J. Goetz (2007) Social and political forces as determinants of Poverty: A spatial analysis. *Journal of Socio-Economics,* 36**,** 650-671.

Schuler, K. M., Chu, P. W. & Smith-Bindman, R. (2008). Factors associated with mammography utilization: A systematic quantitative review of the literature. *Journal of Women's Health,* 17(9), 1477-1498.

Siegel, R., Ward, E., Brawley, O., & Jemal, A. (2011). Cancer statistics, 2011: The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA: A Cancer Journal for Clinicians*, caac. 20121v20121.

Siminoff, L., Ross, L. (2005). Access and equity to cancer care in the USA: A review and assessment. *Postgraduate Medical Journal*, 81, 674-679.

Singh, G. K. (2011). Rural–Urban trends and patterns in cervical cancer mortality, incidence, stage, and survival in the United States, 1950–2008. *Journal of Community Health*, 1-7.

Slayden, G. (2003). Using geographic codes to import data into MapPoint. Retrieved from http://msdn.microsoft.com/en-us/library/ms980061.aspx

Slifkin, R. (2002). Developing policies responsive to barriers to health care among rural residents: What do we need to know? *Journal of Rural Health,* 18(S), 233-41.

Smith-Bibdman, R., Miglioretti, D. L., Lurie, N., Abraham, L., Barbash, R. B., Kerlikowske, K. (2006). Does utilization of screening mammography explain racial and ethnic differences in breast cancer? *Annals of Internal Medicine,* 144(8), 541-554.

Smith, R. A., Cokkinides, V., & Brawley, O. W. (2009). Cancer screening in the United States, 2009: A review of current American Cancer Society guidelines and issues in cancer screening. *CA: A Cancer Journal for Clinicians*, 59, 27–41.

Snyder, A. R., McLaughlin, D. K. & Findeis, J. (2006). Household composition and poverty among female-headed households with children: Differences by race and residence. *Rural Sociology*, 71, 597-624.

Snyder, A. R. & McKaughlin, D. K. (2004). Female-headed families and poverty in rural America. *Rural Sociology, 69* (1(, 127-149.

Thomson, C. S., Hole, D. J., Twelves, C. J., Brewster, D. H. & Black, R. J. (2001). Prognostic factors in women with breast cancer: distribution by socioeconomic status and effect on differences in survival. *Journal of Epidemiology and Community Health,* 55, 308-315.

Ulubasoglu, M. A. & Cardak, B. (2005). International comparisons of rural-urban educational attainment: Data and determinants, *Economics Series*, 1.

U.S. Census Bureau. (2010). Rural America continuing to shrink. Retrieved from http://www.star-telegram.com/2011/07/27/3251216/rural-america- continuing-to-shrink.html

U.S. Census Bureau. (2010). Income, poverty and health insurance coverage in the United States: 2009. Retrieved from http://www.census.gov/newsroom/releases/archives/income_wealth/cb10-144.html

U.S. Census Bureau (2000). Socioeconomic differences in household.

U.S. Department of Agriculture and Economic Research Services. (2005). Rural poverty at a glance. Rural Development Research Report Number 100, March 2005. U.S. Department of Health and Human Services. (2003) Agency for Health care Research and Quality. Screening for cervical cancer. Washington, DC: U.S. Government Printing Office

U.S. Department of Health and Human Services. (2010). Healthy People 2020 framework. Retrieved from http://healthypeople.gov/2020/consortium/HP2020Framework.pdf

U.S. Department of Health and Human Services (2000) Healthy People 2010: Vol I. Understanding and Improving Health (Vol I) (2nd Ed). http://www.healthypeople.gov/Document/tableofcontents.htm#volume1.

U.S. Preventive Services Task Force. (2002). screening for breast cancer: United States Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine*, 137(Part 1), 344–346.

Voss, P., D. D. Long, R. B. Hammer & S. Friedman (2006) County child poverty rates in the US: A spatial regression approach. *Population Research and Policy Review,* 25**,** 369-391.

Waldorf, B. (2007). What is rural and what is urban in Indiana. *West Lafayette: Purdue Center for Regional Development, PCRD*.

Wang, F., McLafferty, S., Escamilla, V. & Luo, L. (2008). Late-stage breast cancer diagnosis and health care access in Illinois. *Professional Geographer,* 60 (1), 54–69.

Warner, E.T. & Gomez, S. L. (2010). Impact of neighborhood racial composition and metropolitan residential segregation on disparities in breast cancer stage at diagnosis and survival between black and white women in California. *Journal of Community Health,* 35 (4), 398-408.

Webster, T. F., Hoffman, K., Weinberg, J., Vieira, V. & Aschengrau, A. (2008). Community and individual level socioeconomic status and breast cancer risk: Multilevel modeling on Cape Cod, Massachusetts. *Environmental Health Perspective,* 116 (8), 1125-1129.

Williams, A.P., Schwartz, W.B., Newhouse, J.P., Bennett, B.W. (1983). How many miles to the doctor? *New England Journal of Medicine,* 309 (16), 958-963.

Woods, L. M., Rachet, B. & Coleman, M. P. (2006). Origins of socio-economic inequalities in cancer survival: a review. Ann Oncology, 7, 5-19.

Yost, K., Perkins, C., Cohen, R., Morris, C. & Wright, W. (2001). Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer causes Control,* 12, 703-711.

Young, A. F., Dobson, A. & Byles, J. E. (2001). Determinants of general practitioner use among women in Australia. *Social Science and Medicine*, 53 (12), 1641–51.

Zenk, S. N., Tarlov, E., & Sun, J. (2006). Spatial equity in facilities providing low- or no-fee screening mammography in Chicago neighborhoods. *J Urban Health,* 83, 195–210.

VITA

Faustine Williams was born in Accra, Ghana on September 26, 1973. She went to Albert Academy and Victory Commercial College all located in Madina, Accra, Ghana for her high school. After high school, she worked with World Vision International in Ho from 1992-1997 as a Secretary and Receptionist. She later went to the Accra School of Hygiene and pursued a diploma in Environmental Health. She worked with the Ministry of Health and the Ministry of Local Government and Rural Development from 2000-2001 as an Environmental Health Officer, in Accra. Aiming for greater heights, in her final year at School of Hygiene she registered for the General Certificate of Education Advanced-Level Examination (GCE A-level) and passed with distinction in 1999.

She received her bachelor degree with a First Class Honors in Information Studies with a minor in Psychology from the University of Ghana, Legon in May 2003. Faustine worked with Third World Network-Africa (TWN-Africa) in Accra as Information Analyst. While with TWN-Africa, she developed and implemented online library database in Microsoft Access and also trained and provided technical support for information users. In 2005, she was awarded the American Association of University Women (AAUW) International Fellowship Scholarship to pursue her master's degree in Health Informatics at the University of Missouri, Columbia. While pursuing her Ph.D. studies in Rural Sociology at the same university, she also completed a master's degree in Public Health in May 2011. Her Ph.D. emphasis area is in Community Studies, Development and Informatics and completed in July 2012. Her doctoral research examined the Spatial Cluster Analysis of Female Breast Cancer Diagnosis in Missouri: Using GIS and Spatial Analyst Functions.