2008-06-03

A Spatial Analysis of Colorectal Cancer in Miami-Dade County

Monique Hernandez

University of Miami, m.hernandez35@umiami.edu

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation

https://scholarlyrepository.miami.edu/oa_dissertations/99

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
A SPATIAL ANALYSIS OF COLORECTAL CANCER IN MIAMI-DADE COUNTY

By

Monique Hernandez

A DISSERTATION

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Coral Gables, Florida

May 2008
This dissertation explores the spatial patterns and place-based characteristics of colorectal cancer (CRC) late stage incidence and CRC-specific mortality in Miami-Dade County. Because CRC is the second leading cause of death among all cancers and is almost 90 percent preventable through medical screenings, investigations of CRC disparities across groups and communities are extremely relevant in the fight against cancer. This paper analyzes the geographic distribution of CRC cases in Miami-Dade County between two periods, 1988-1992 and 1998-2002 to: a) identify significant “hot spots” or clusters of disease; b) investigate associations of CRC patterns with neighborhood level characteristics such as socio-economic status, race/ethnicity, and poverty; and c) explore the policy implications of the spatial trends identified for the disease, with particular reference to the Welfare Reform Act of 1996. This dissertation analyzes data from the Florida Cancer Data Registry and tract level U.S. Census data, to identify the spatial distribution of CRC and study its relation to place-based variables using Geographic Information Systems (GIS) and spatial statistical modeling. Identifying spatial clusters of disease can assist in targeting public health interventions and improving social service delivery, particularly for uninsured populations. Identifying communities facing greater obstacles to screenings and quality medical care through the use of spatial analysis is an effort to mitigate these barriers while simultaneously
providing empirically based evidence linking neighborhood-level social and economic conditions to health disparities.
DEDICATION

I would like to dedicate this dissertation to my family, whose encouragement of higher learning and constant support throughout my academic pursuits has made this journey positive and worthwhile.
ACKNOWLEDGEMENTS

I would like to acknowledge and thank my mentor, Dr. Rinku Roy Chowdhury, for her guidance, wisdom, and dedication to this dissertation. I would also like to thank my committee members, Dr. Daniel A. Griffith, Dr. Lora E. Fleming, and Dr. Elizabeth Aranda for their service to my Interdisciplinary Ph.D. program. I would like to give special thanks to Dr. Jill MacKinnon at the Florida Cancer Data System and Dr. Amie L. Nielson in the Department of Sociology, for their contributions to this dissertation. Finally, I want to extend my deepest gratitude to Greg Reece, for his extraordinary proofreading abilities, tireless dedication, and boundless emotional support. Without the support of these people, this dissertation would not have been made possible. I am eternally grateful.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.0</td>
<td>Overview of Colorectal Cancer</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Health Inequalities: Theoretical and Methodological Approaches</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Contributions of This Dissertation</td>
<td>10</td>
</tr>
<tr>
<td>1.3</td>
<td>Research Site: Miami-Dade County</td>
<td>11</td>
</tr>
<tr>
<td>1.4</td>
<td>Research Problem and Significance</td>
<td>13</td>
</tr>
<tr>
<td>1.5</td>
<td>Structure of the Dissertation</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>18</td>
</tr>
<tr>
<td>2.0</td>
<td>Overview</td>
<td>18</td>
</tr>
<tr>
<td>2.1</td>
<td>Descriptive Epidemiology of CRC</td>
<td>19</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Individual Risk Factors</td>
<td>19</td>
</tr>
<tr>
<td>2.1.2</td>
<td>Epidemiology by Race, Ethnicity and Gender</td>
<td>22</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Descriptive Epidemiology by Socioeconomic Status and Geographic Area</td>
<td>26</td>
</tr>
<tr>
<td>2.2</td>
<td>Health Inequalities Research and Theoretical Approaches</td>
<td>28</td>
</tr>
<tr>
<td>2.3</td>
<td>Miami, CRC Patterns, and Health Insurance</td>
<td>34</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Miami-Dade County</td>
<td>34</td>
</tr>
<tr>
<td>2.3.2</td>
<td>CRC Patterns in Florida and Miami-Dade County</td>
<td>38</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Health Insurance, Cancer Screening and Miami-Dade County</td>
<td>40</td>
</tr>
<tr>
<td>2.4</td>
<td>Spatial Analysis of Health</td>
<td>42</td>
</tr>
</tbody>
</table>
2.4.1 Spatial Representation and Analysis of Public Health Data ................. 42
2.4.2 Area-Based Social and Economic Measures ........................................ 44
2.4.3 Limitations and Considerations ............................................................ 45
2.5 Summary ................................................................................................. 47

3 WELFARE REFORM EFFECTS ON ELDERLY IMMIGRANTS .............. 48
3.0 Overview ................................................................................................. 48
3.1 New Public Assistance Eligibility Criteria Under PRWORA ................. 49
3.2 PRWORA and Elderly Immigrants .......................................................... 51
3.3 PRWORA and Elderly Legal Immigrants in Florida and Miami ............... 55
3.4 Conclusion ............................................................................................... 58

4 RESEARCH DESIGN AND METHODS ..................................................... 60
4.0 Overview of Study Design ..................................................................... 60
4.1 Data Sources ........................................................................................... 61
4.1.1 The Florida Cancer Data System (FCDS) .......................................... 61
4.1.2 U.S. Census Bureau Variables .......................................................... 65
4.2 Cluster Analysis ....................................................................................... 66
4.3 Spatial Regression Analysis .................................................................... 70
4.3.1 Dependent Variables ......................................................................... 71
4.3.2 Independent Variables ....................................................................... 71
4.3.3 Negative Binomial Regression: Models and Hypotheses ................... 72
4.3.4 Limitations of the Negative Binomial Model with the Spatial Lag Term ......................................................................................... 76

5 RESULTS .................................................................................................. 77
5.0 Overview of Results .............................................................................. 77
5.1 Descriptive Statistics of CRC Patient Data

5.1.1 1988-1992 (T1) Cancer Cases By Race/Ethnicity and Gender in Miami-Dade County

5.1.2 1998-2002 (T2) Cancer Cases By Race/Ethnicity and Gender in Miami-Dade County

5.1.3 Comparison of Incidence, Late Stage, and Mortality from T1 to T2

5.2 Results of Spatial Clustering Analysis for T1 and T2

5.2.1 Study Hypothesis A1

5.2.2 Study Hypothesis A2

5.2.3 Study Hypothesis B1

5.2.4 Study Hypothesis B2

5.2.5 Descriptive Statistics for Significant CRC Clusters

5.3 Negative Binomial Spatial Regression Analysis

5.3.1 Study Hypotheses C1 and C2

5.3.2 Study Hypotheses D1 and D2

5.4 Discussion of Results

5.4.1 Patterns of Late Stage CRC Incidence and CRC-Specific Mortality in Miami-Dade County

5.4.2 Significance of Cluster Analysis Results

5.4.3 Discussion of Spatial Regression

5.5 Limitations

5.6 Areas for Future Research

6 CONCLUSION

WORKS CITED
LIST OF FIGURES

Figure 1.1: Conceptual Diagram of Research Approach ...................................................5

Figure 4.1: CRC Incidence by Age Category .................................................................68

Figure 5.1: CRC Incidence Rates by Age Group and Race/Ethnicity, 1988-1992 ..........80

Figure 5.2: CRC Late Stage Rates by Age Group and Race/Ethnicity, 1988-1992 .......80

Figure 5.3: CRC Mortality Rates by Age Group and Race/Ethnicity, 1988-1992 ..........81

Figure 5.4: CRC Incidence Rates by Age Group and Race/Ethnicity, 1998-2002 ..........84

Figure 5.5: CRC Late Stage Rates by Age Group and Race/Ethnicity, 1998-2002 .......84

Figure 5.6: CRC Mortality Rates by Age Group and Race/Ethnicity, 1998-2002 ..........84

Figure 5.7: Map of Clusters of Higher than Expected Late-Stage CRC Cases in Miami-Dade, 1988-1992 ................................................................................................................90

Figure 5.8: Map of Clusters of Higher than Expected CRC Mortality Cases in Miami-Dade, 1988-1992 ................................................................................................................93

Figure 5.9: Map of Clusters of Higher than Expected Late-Stage CRC Cases in Miami-Dade, 1998-2002 ................................................................................................................96

Figure 5.10: Map of Clusters of Higher than Expected CRC Mortality Cases in Miami-Dade, 1998-2002 ................................................................................................................99
### LIST OF TABLES

Table 2.1: U.S. CRC Incidence and Mortality by Race, Ethnicity, and Gender ..........23

Table 2.2: Percent CRC by Stage, Five-Year Survival Rates, Race, Ethnicity, and Gender for Nine SEER Registries, 1996 – 2003 .................................................................25

Table 4.1: International Classification of Diseases for Colorectal Cancer .............62

Table 4.2: File Layout of FCDS Data Provided Under January 23, 2008 Release H06105, Amendment #1 .............................................................................................................64

Table 4.3: Variables Extracted from U.S. Census Data .........................................66

Table 4.4: Number of CRC Late Stage and Mortality Cases by Study Period and Age Category .........................................................................................................................68

Table 5.1: CRC Stage Distribution by Race/Ethnicity and Gender, 1988-1992 ...........79

Table 5.2: Age-Adjusted CRC Incidence, Late Stage, and Mortality Rates per 100,000 People by Race/Ethnicity, 1988-1992 ...............................................................82

Table 5.3: CRC Stage Distribution by Race/Ethnicity and Gender, 1998-2002 ...........83

Table 5.4: Age-Adjusted CRC Incidence, Late Stage, and Mortality Rates per 100,000 People by Race/Ethnicity, 1998-2002 ...............................................................86

Table 5.5: Cluster Analysis Results for Late Stage CRC Incidence Cases, Age 50 and Older, 1988-1992 .................................................................................................89

Table 5.6: Cluster Analysis Results for CRC-Specific Mortality, Age 50 and Older, 1988-1992 .................................................................................................92

Table 5.7: Cluster Analysis Results for Late Stage CRC Incidence Cases, Age 50 and older, 1998-2002 .................................................................................................95

Table 5.8: Cluster Analysis Results for Late Stage CRC Incidence Cases, Age 50 and older, 1998-2002 .................................................................................................98

Table 5.9: Selected Population Characteristics of Statistically Significant Clusters .....100

Table 5.10: Descriptive Statistics of T1 Model Variables (n=264 tracts) .................104

Table 5.11: Zero Order Correlations, T1 CRC Cases with 1990 Census Variables ......104

Table 5.12: Descriptive Statistics of T2 Model Variables (n=345 tracts) .................104

Table 5.13: Zero Order Correlations, T2 CRC Cases with 2000 Census Variables ......105
Table 5.14: Negative Binomial Regression, Late Stage CRC Incidence, 1988-1992 .....107
Table 5.15: Negative Binomial Regression, CRC Mortality, 1988-1992.........................107
Table 5.16: VIFs, Late Stage CRC Incidence, 1988-1992 .............................................108
Table 5.17: VIFs, CRC Mortality, 1988-1992................................................................108
Table 5.18: Negative Binomial Regression, Late Stage CRC Incidence, 1998-2002 ....109
Table 5.19: Negative Binomial Regression, CRC Mortality, 1998-2002.......................110
Table 5.20: VIFs, Late Stage CRC Incidence, 1998-2002 .............................................110
Table 5.21: VIFs, CRC Mortality, 1998-2002................................................................110
CHAPTER 1
INTRODUCTION

1.0 Overview of Colorectal Cancer

This country faces enormous public health challenges due to rising rates of cancer, even of one of the most preventable types, colorectal cancer (CRC). In the past century, national health patterns in the United States (U.S.) have demonstrated an epidemiological shift marked by decreased infectious disease prevalence and increased prevalence of chronic diseases. Improvements in water and sanitation, higher standards of living, and widespread vaccination efforts have diminished infectious disease rates, while chronic diseases such as cancer, diabetes, and heart disease have become the leading causes of morbidity and mortality in the U.S. (CDC, 2001, 2007a). Largely due to increasing life expectancy and changes in lifestyle habits, chronic disease now accounts for seventy percent of all mortality cases in the country (NCI, 2007a), and cancer is the second leading cause of all deaths in the U.S., causing over half a million deaths in 2007. This trend will only continue.

The National Cancer Institute (2007a) estimates that 41 percent of men and women born today will be diagnosed with cancer within their lifetime. Edwards and colleagues (2002) project that the prevalence of cancer will double by the year 2050 and that individuals who are 75 years of age or older will suffer 42 percent of all cancer cases in that year. Current statistics related to CRC are particularly discouraging in light of the fact that CRC is one of the most preventable cancer types through early detection. The colon and rectum together represent the third leading site of cancer incidence, and CRC is the second leading cause of cancer-specific mortality in this country (NCI, 2007a). These}

1
statistics alone demonstrate the pressing need for research, health surveillance, and improvements in preventive care and treatment specifically targeted at reducing the incidence of and mortality rates due to CRC.

Although increased public attention and screenings have reduced CRC morbidity and mortality over the past 10 years, significant population disparities in late stage cancer and cancer mortality rates remain (ACS, 2008). Colorectal cancer incidence is distributed unevenly across population groups, showing marked differences when differentiated by race and ethnicity, socioeconomic status, and residential location. A study of cancer trends between 1975 and 2002 (Irby, 2006) reports that racial disparities in late stage CRC incidence were greatest between blacks and whites, with blacks having up to a 30 percent greater likelihood for advanced stage CRC than non-Hispanic whites. Additionally, trends of CRC-specific mortality rates over the same time period depict an overall decline for whites, but an increase for blacks. Stefenidis and colleagues (2006) found similar differences in CRC rates between Hispanic and non-Hispanic whites in Texas; they found that Hispanics were diagnosed with CRC at younger ages, had more occurrences of late stage cancers, and had poorer survival rates. Studies throughout the literature report similar findings of racial and ethnic CRC outcome differences (Chen, 1997; Fazio, 2005; Mandelblatt, 1996; Marcella, 2001; Ries, 2001; Singh et al., 2003).

Many of these studies also found socioeconomic differences between CRC and health outcomes. According to studies conducted by Singh and colleagues (2002; 2003), people residing in high poverty areas were as much as 12 percent more likely to experience CRC mortality than people in low poverty areas. The same studies found that significant differences in CRC mortality based on race alone disappeared when poverty
was taken into account. These findings suggest that socioeconomic differences at the neighborhood level may, in fact, explain the racial and ethnic variations in CRC health outcomes.

Similarly, health care coverage significantly affects the likelihood of late stage CRC incidence and CRC-related mortality, making access to health insurance a primary goal for CRC prevention strategies (Breen et al., 2001; CDC, 2001; Fazio, 2005; Mandel, 1993; Matthews et al., 2005). Diagnostic medical screenings can prevent up to 90 percent of colorectal cancer cases; however, less than 50 percent of those who satisfy the recommended screening criteria are actually screened, a gap that researchers have attributed largely to lack of health insurance (ACS, 2008; Breen et al., 2001; Matthews et al., 2005; Nash et al., 2006; Weir et al., 2003; Winawer et al., 2003; Zoorob et al., 2001).

In 2005, a total of 47 million Americans were uninsured, an increase of 7 million people since 2000 (Denavas-Walt, 2006). For the “near elderly” (ages 55 to 64), the age range at which CRC screenings typically begin, insurance figures depict a steady decline in health care coverage, with a million fewer people being insured in 2005 than in 2000 (Fronstin, 2007). Health insurance coverage is particularly low for Hispanics. Nationally, Hispanics are three times more likely than non-Hispanic whites to be uninsured (Bond, 2002). Similarly, a third of all foreign-born immigrants are uninsured, compared to 13.4 percent of native born citizens who do not have insurance.

Lack of health care coverage warrants particular attention in intervention programs because it is one of the main risk factors for advanced stage CRC and resultant death. While studies that identify CRC outcome disparities provide important insight into the social and economic factors of health, professionals in the health policy field
must use such scientific findings to evaluate health policies and to develop disease intervention programs that target high risk population groups (NCI, 2007b). Studies that combine policy evaluation and health risk research prove particularly helpful to such efforts. For example, studies that have analyzed the impact of new restrictions on health coverage resulting from federal welfare reform in 1996 have identified heightened vulnerabilities among non-citizen immigrants (Binstock & Jean-Baptiste, 1999; Prentice, 2005; Weil & Finegold, 2002). In addition to the effects on health care coverage of individuals, welfare reform has burdened local-level social service organizations, such as community clinics and non-profit organizations, by increasing demand for these alternative sources of aid (Coburn, 2000; Navarro & Shi, 2001; Szreter & Woolcock, 2004). Such evidence-based research provides important information for health policy development and points to key concerns that require the attention of public health officials. In that light, this dissertation analyzes health inequalities in CRC incidence and outcomes in Miami-Dade County in an effort to identify health policy priorities for the early detection and treatment of this most preventable cancer.

1.1 Health Inequalities: Theoretical and Methodological Approaches

Health inequalities research traditionally has relied on medical and social science disciplines that evaluate disease risk factors ranging from individual health behaviors to broader environmental exposures. Inevitably, the particular theoretical and methodological approach applied in health research itself influences findings and explanations of disease pathways. To minimize the preconceptions introduced by selection of a particular methodology, this dissertation combines multiple health research
approaches. Specifically, it employs aspects of individual, societal, and spatial approaches at their point of intersection. Figure 1.1 is a diagram representing each approach and where they overlap. This section describes each of these components and their individual contributions to the health field.

**Figure 1.1: Conceptual Diagram of Research Approach**

During the mid-twentieth century in the U.S., the field of modern epidemiology concentrated research efforts on chronic disease control by seeking to understand the role of individual behaviors in disease risk (Susser & Susser, 1996). Commonly employing observational cohort and case-control study designs, epidemiological researchers control for individual exposures and disease outcomes by comparing the outcomes of two distinct groups, such as individuals exposed to a risk versus those not exposed or those with a disease versus those without, respectively (McLaren & Hawe, 2005). In this way, epidemiological studies can evaluate the risk of developing a disease given a certain
exposure. Epidemiologists Doll and Hill (1950) produced one of the most important studies that linked the smoking of tobacco with lung cancer and heart disease. Their findings led to wide examination of personal health habits and behaviors, and perpetuated the belief that individuals were primarily responsible for their health outcomes. While modern epidemiology has generated important research during the emerging chronic disease epidemic, many scholars have criticized the discipline for its overemphasis on individual behaviors and for the lack of analysis of social and economic factors affecting individual choices (Schwartz et al., 1999; Susser & Susser, 1996).

Moreover, although case-control and cohort study designs produce immensely detailed data of individual characteristics, conducting such a study often requires a significant amount of resources – including time and money – and in some cases require access to or creation of extensive medical records. For instance, the Framingham Heart Study followed 5,200 men and women in a longitudinal cohort research design, producing detailed accounts of each subject and facilitating analyses of behavioral and group characteristics in cardiovascular disease outcomes (NHLBI, 2007). However, the study required extensive follow-up procedures and administrative resources and took a total of 50 years to conduct. The Health Insurance Portability and Accountability Act (Health Insurance Portability and Accountability Act, 1996) can further complicate access to data, particularly to medical records, as it instituted strict privacy compliance standards and put into place institutional review boards charged with ensuring that studies comply with those requirements.

The limited scope of individual and behavioral explanations in health studies and the resource requirements of large-scale cohort and case control studies prompted the
resurgent use of ecological methods and population-based studies to examine the effects of societal structural factors on health across population groups. While ecological studies have been used for a long time in epidemiology, in the early twentieth century, industrialized countries began to focus increasingly on individual-level health risks (Susser & Susser, 1996). In the 1980s, seminal studies, such as the Black Report in the United Kingdom (Marmot et al., 1991; Townsend & Davidson, 1982), underscored the importance of ecological analysis by reporting significant associations that connected social class to life expectancy and chronic disease outcomes. Such population-based ecological studies renewed interest in research into the social and economic determinants of health inequalities, and precipitated theoretical arguments as to the social structural pathways involved in population health (Braveman et al., 2005; Krieger & Aierler, 1996; Krieger et al., 1993; Schwartz et al., 1999; Shy, 1997; Susser & Susser, 1996; Terris, 1987; Wingo et al., 2004).

Researchers interested in the social determinants of health, including those drawing from disciplines such as medical sociology and social epidemiology, commonly study the effects of income inequality, material and social resources, and racial and ethnic inequalities on health outcomes (Bartly et al., 1998; Macintyre, 1997; Popay et al., 1998). Health researchers may be concerned specifically with the influence of relative income deprivation and related psychosocial effects on morbidity and mortality (e.g., Kaplan & Keil, 1993; Kennedy et al., 1998; Wilkinson, 1996), or focus on how the devolution of government-provided social services impacts health (e.g., Coburn, 2004; Muntaner et al., 2002; Navarro & Shi, 2001). Such analyses, which commonly utilize ecological study designs that employ group-level analysis, play an important role in health research, as
they aim to uncover the contexts affecting health, such as built and natural environments, social relations, and economic conditions (McLaren & Hawe, 2005). Such studies also typically require fewer resources and primarily draw from secondary data sets with more socioeconomic variables than traditional cohort and case-control studies are able to collect (Krieger, 2001; Krieger et al., 2003).

Despite efforts by health inequalities researchers to depart from exclusively individual-level explanations and to understand the social determinants of health, scholars including Krieger (1993), Kaufman (1999), and Macintyre and colleagues (2007; 2002) have pointed out that explanations of the spatial processes and pathways to health outcomes are absent from many health inequalities studies. In contrast, disciplinary subfields, such as medical/health geography and spatial epidemiology, actively integrate spatial processes and area features into research on designs. Two approaches predominate in geographical analysis of health. Medical geographers typically study the distribution of health events in geographic space (e.g., the distance from and physical access to health care facilities) and the relationship between area characteristics and disease incidence (Gesler, 2003; Meade, 2000). Studies that employ disease mapping, for instance, serve important public health surveillance needs by identifying clusters of high risk populations. Likewise, spatial regression analysis measures the relationships of health events, social and economic characteristics, and their geographical distributions (Waller & Gotway, 2004). Health geographers emphasize, on the other hand, the qualitative aspects of an area, or ‘place,’ as defined by socio-cultural histories and lay perceptions of health (Kearns, 1993; Kearns & Joseph, 1993; Moon, 1995). These studies generally are more qualitative and anthropological.
Current research increasingly highlights the contextual features of areas, examining both the material infrastructure and collective social functioning (Kearns, 1993; Macintyre & Ellaway, 2003; Macintyre et al., 2002; Meade, 2000). Davidson and colleagues argue that “space and society are mutually constructive; that people’s identities, attitudes and behaviors are shaped by, and in turn shape, the places in which they live” (Davidson et al., 2008:168). Research that analyzes the historic effects of racial discrimination on residential segregation and concentrated poverty provides examples of how space and society are mutually constructed (Massey & Denton, 1993; Rosenberg & Wilson, 2000; Sloggett & Joshi, 1994). Social patterns and processes such as these are examples of the contextual effects of health, and demonstrate why these effects warrant special consideration in the analysis of health inequalities.

While research that incorporates the spatial and contextual features of health generally builds on existing health inequalities studies, many scholars criticize the lack of supporting theory in the literature and point to the ecological fallacy, which occurs when conclusions drawn at the group-level are applied to individuals (Eyles, 1993; Gesler, 1991; Kearns & Joseph, 1993; Litva & Eyles, 1995; Macintyre et al., 2002; McGlashan, 1972). However, health researchers increasingly are making connections to social theory, political economy, and spatial processes when analyzing health patterns (Dycke, 1995; Gesler & Kearns, 2002; Macintyre et al., 2002; Mohan, 1998).

Even in this era of increasingly interdisciplinary research, the integration of theoretical and methodological approaches is a challenge in health studies, a task made particularly difficult by the many complexities inherent in population health outcomes. The study of health outcomes in population groups are typically characterized: 1) at an
individual level, by focusing on health behaviors and biological factors, 2) at a societal level, by highlighting different health outcomes across populations by race or class, and 3) geographically, by their spatial distributions across areas to delineate communities at highest risk for disease. This dissertation attempts to fill existing gaps in health inequalities research by purposefully integrating these approaches to health outcomes, thus providing a more comprehensive framework within which the causal mechanisms of health, and linkages to CRC in particular, can be examined across these factors. As Kawachi and Berkman articulate (Kawachi & Berkman, 2003), while the identification of links between neighborhoods characteristics and health are not new, the inter-disciplinary integration concepts, methods, and evidence is just now emerging.

1.2 Contributions of This Dissertation

This dissertation examines the health inequalities of colorectal cancer through multiple theoretical and methodological approaches to study the spatial distribution of CRC outcomes in Miami-Dade County. The methodological approach combines ecological and spatial analytical techniques. The integration of these two techniques permits exploration of the contextual as well as spatial processes behind CRC incidence. This dissertation considers a number of relevant questions regarding CRC outcomes and neighborhood context. Specifically, are there spatial clusters of higher than expected CRC late stage incidence and CRC mortality? What associations exist between the neighborhood-level social and economic characteristics and individual-level CRC outcomes and their spatial distributions in Miami-Dade County? Additionally, how have recent changes in federal and state welfare policies affected legal immigrant’s access to
public aid services such as Medicaid, and how do regional-level policies influence neighborhood differences in CRC outcomes? To date, few studies have focused on the relationship between immigration status and access to federally-provided health care services among one of the most disadvantaged population groups, elderly legal immigrants. This study examines how policy decisions at the federal and state level influence immigrants’ access to government-provided health insurance, and highlights the increasingly exclusionary treatment of immigrants based on citizenship status.

1.3 Research Site: Miami-Dade County

Miami-Dade County is a diverse metropolitan area made up of 2.4 million people, of which over 50 percent are foreign-born residents. A majority (61 percent) of the County residents are Hispanic, another 20 percent are black, and the remaining 19 percent are non-Hispanic white (Census, 2000b). Cubans constitute the largest Hispanic sub-group, making the county distinct from the Hispanic immigrant communities in New York, Chicago, and Los Angeles. Additionally, the County is home to a significant Caribbean and South American black population.

Over the past 25 years, Miami-Dade County has experienced tremendous growth in business, economic development, and wealth. The benefits of this growth, however, are disproportionately spread, leaving many people struggling economically (Nijman, 1997). Socioeconomic inequalities in Miami are correlated with race and ethnicity, with blacks suffering the greatest proportion of economic, social, and political disadvantage (Nijman, 1997). The introduction of Cuban and Haitian immigrants into Southern Florida in the 1980s drastically changed racial dynamics and population demographics in
Miami. The Cuban community, having the advantage of government conferred refugee status, became increasingly entrepreneurial, politically engaged, and exclusive in hiring practices, eliminating many occupational opportunities for both the U.S. and Caribbean-born black communities (Portes & Stepick, 1993). Conversely, Haitian immigrants were denied the status and support that were afforded those with political refugee status and have faced greater constraints and more limited opportunities than did Cuban refugees.

In addition to these challenges, current health and health care trends in Miami-Dade County indicate increasing indigent and at-risk populations. In 1996, changes in the U.S. federal welfare policy put many citizens and non-naturalized immigrants at risk by instituting new eligibility requirements for welfare programs, such as Medicaid, Temporary Aid for Needy Families, and Supplementary Security Income (Fix & Passel, 1999). Federal welfare reform threatened legal non-citizen immigrants in particular. Non-citizen immigrants account for 50 percent of all foreign-born residents in Miami-Dade County (Census, 2000b).

Health care coverage is directly linked to socioeconomic status in Miami-Dade County and is also a major concern since over 28 percent of the county’s population lacked health insurance in 2004 (DeNavas-Walt et al., 2006; DHSRMP, 2005). In the County, cancer ranks second in cause of death after heart disease, and, among types of cancer, only lung cancer causes more deaths than CRC (FCDS, 2005). While screening is the primary means for the early detection and prevention of CRC, low rates of compliance with CRC screening, combined with limited access to preventive care, characterize the population of Florida and of Miami-Dade County. In 2003, only 40 percent of the Florida population aged 50 and over had ever undergone a colonoscopy or
sigmoidoscopy screening for CRC, even though, starting at age 50, everyone is recommended to have a sigmoidoscopy every five years and a colonoscopy every 10 years (FDH, 2007).

The social, political, and health conditions outlined above make Miami-Dade County particularly vulnerable to growing inequalities in CRC late stage and mortality outcomes. Almost 90 percent of all colorectal cancers are preventable through early detection from regular screenings (ACS, 2008). In this respect, access to affordable health care becomes essential for CRC prevention and screening, but is increasingly difficult to obtain due to reductions in employment-based health care and government-provided health insurance.

1.4 Research Problem and Significance

Drawing from theoretical and methodological approaches in medical geography, social epidemiology, and medical sociology, this study contributes practical public health surveillance information, in addition to discussing the structural pathways involved in CRC incidence given the social conditions that exist in Miami-Dade County.

This dissertation includes a spatial analysis of CRC late stage incidence and CRC-specific mortality among Miami-Dade County residents who were diagnosed during the years 1988 to 1992 and 1998 to 2002. It specifically employs spatial cluster analysis to identify communities with higher than expected rates of CRC, as well as spatial statistical regression analysis to measure the impact of social inequalities on CRC. Colorectal cancer outcomes are examined based on axes of difference defined by socio-economic status (SES), race and ethnicity. Utilizing cross-sectional data at two distinct time
periods, this study evaluates temporal changes in CRC outcomes with the goal of gauging the interplay between health policy changes (such as welfare reform or changes in screening guidelines) and health outcomes. The specific objectives of this dissertation are to:

- identify colorectal cancer “hot spots,” i.e., global clusters of elevated rates of late stage CRC incidence and CRC-specific mortality in Miami-Dade County, for two time periods: 1988-1992 and 1998-2002;
- employ spatial regression analysis to measure tract-level associations between CRC outcomes and social and economic variables for each time period; and,
- analyze the policy implications of the 1996 Federal welfare reform and its effects on legal elderly non-citizens.

The purpose of this research is to highlight the spatial dimensions of CRC outcomes, link it with neighborhood-level social and economic characteristics, and identify important relationships that operate above individual-based explanations. Furthermore, in analyzing the spatial and structural patterns of CRC outcomes, this dissertation integrates multiple theoretical approaches and methodologies and underscores the significance of interdisciplinary research in the health field, with the specific purpose of examining how inequalities in health are a product of many processes. These processes directly and indirectly influence vulnerabilities people face in disease prevention and health management. It is essential that these vulnerabilities are identified and mediated at the individual, neighborhood, and national scales. Dunn and colleagues (2006) are particularly effective in describing the importance of this approach.
If the systematic patterning of an individual’s health status is shaped partly by the contexts, places, and locations in which individuals live, in addition to their own individual attributes, this may open up new avenues of intervention, over the long run, that differ fundamentally from traditional individually based interventions (pp 238).

Based on *a priori* findings in the health literature, the hypotheses of this dissertation are as follows:

1. **Extant patterns**: Late stage CRC incidence and CRC-specific mortality display spatial concentrations after controlling for age and population size.

2. **Neighborhood social and economic features**: Late stage CRC incidence and CRC-specific mortality display associations with neighborhood-level socioeconomic factors and racial and ethnic compositions.

3. **Policy effects**: National welfare policy reform in 1996 has placed those now ineligible for government-provided health care at greater risk for negative CRC outcomes. In particular, elderly non-citizens and recently immigrated foreign-born residents are at greater risk due to these policy changes.

The research and findings produced by this dissertation have direct implications for the intervention and prevention of CRC among vulnerable populations in Miami-Dade County, and can be applied to comparable metropolitan areas. Moreover, this research promotes theoretical considerations of social and economic structures and their pathways to health. The integration of multidisciplinary methods and theory provide a more comprehensive approach to uncovering complex and multi-scalar processes involved in health inequalities than any single method could achieve. Finally, this dissertation underscores the importance of linkages between science and policy, and their influence on health outcomes.
1.5 Structure of the Dissertation

Chapter 2 reviews the literature on colorectal cancer epidemiology, the demographic, socioeconomic, and spatial distributions of CRC, and analytical and methodological approaches in health inequalities studies. The chapter includes an account of the socio-political history of Miami-Dade County and its relevance to the county’s current health inequalities, with a specific focus on CRC differences and access to health care challenges.

Chapter 3 provides a description of the Personal Responsibility and Work Opportunity Act of 1996 (PRWORA) and the resulting restrictions on access to public assistance it placed on legal immigrants in the U.S. Additionally, it discusses the Act’s direct and indirect impacts on legal immigrants’ loss of Medicaid coverage and on the care needs of elderly non-citizens in particular. It ends by suggesting the need for greater attention to the specific health vulnerabilities of this population group and the potential effects on CRC-related mortality for these elderly non-citizens.

Chapter 4 describes the set of methodological approaches utilized in this dissertation, as well as the procedures used to employ them. Specifically, this chapter provides detailed description of the data sources, model variables, Kulldorff’s spatial scan cluster method, the negative binomial regression method, design limitations, and the research hypotheses set forth in this dissertation.

Chapter 5 presents the descriptive statistics of late stage CRC incidence and CRC-specific mortality cases for the two time periods and identifies the census variables used in the analyses. The chapter also describes the results of the cluster analysis for late stage CRC incidence and CRC mortality and the results of the negative binomial
regression for both time periods. It concludes with a discussion of study limitations and recommendations for areas of future research.

Chapter 6 discusses the results obtained from empirical analyses while highlighting possible explanations. It concludes with a discussion of the significance of the findings for public health intervention programs and for health inequality research.
CHAPTER 2
LITERATURE REVIEW

2.0 Overview

CRC is the second most deadly cancer in the United States. The National Cancer Institute (2008) estimates that in the year 2008, approximately 148,810 new CRC cases will develop and 49,960 CRC-caused deaths will occur. The most successful preventive measure against CRC, well exceeding the primary preventive value of a healthy diet, is regular medical screenings and early detection of CRC of people 50 years of age and older (ACS, 2008; Nakajima, 2003; Winawer et al., 2003). While CRC can be prevented in as many as 90 percent of cases (ACS, 2008), current incidence rates far exceed potential levels. As a result, health research has increasingly examined CRC outcome differences across groups to determine population-level risk factors. Literature on CRC inequalities has identified significant risk factors related to health behaviors, socioeconomic status, access to health insurance, race and ethnicity, and geographic location. This chapter reviews the descriptive epidemiology of CRC and summarizes the major CRC risk factors, paying special attention to dominant approaches in health inequalities research. This chapter also describes the social and economic characteristics of Miami-Dade County and presents the current status of health insurance coverage and CRC incidence in the county. A discussion of the spatial methods commonly employed in geographical analysis of public health data is included at the end of the chapter.
2.1 Descriptive Epidemiology of CRC

2.1.1 Individual Risk Factors

Cancer is caused by a genetic mutation of a single cell that proliferates into a mass or tumor that invades the body (Hans-Olov et al., 2002). It is a family of diseases that includes over 100 different forms, and cancer can emanate from every cell type. Colorectal cancer originates in the body’s large intestine and rectum, and 70 to 95 percent of CRC cases are due to adenomatous polyps (adenocarcinomas) that develop in the inner lining of the colon (Jackson-Thompson et al., 2006; Rudy & Zdon, 2000). The histogenesis, or behavior, of the adenomatous polyps from which adenocarcinoma originates is often cancer forming, and the polyps can be removed upon detection to prevent malignancy.

The development of cancerous tissue in the large intestine is linked to a number of individual risk factors. Increasing age has the greatest effect. CRC occurs 50 times more often in people aged 60 to 79 years than in those aged less than 40 years (Jackson-Thompson et al., 2006). In addition to the effects of aging, diet and lifestyle factors influence the risk of CRC. A number of studies report an association between ingestion of red meat and CRC incidence that can be linked to excess mutagenic bile acids in the colon (Franco et al., 2005; Potter et al., 1993; Steinmetz & Potter, 1996; Van Munster & Nagengast, 1993). Findings also support that the intake of fruits and vegetables, possibly due to their high fiber content, reduces the risk of cancer by accelerating the transit of waste matter through the colon (Burkitt, 1971; Friedenreich et al., 1994; Potter et al., 1993; Young-In, 2000). Many of these associations, however, are inconclusive or contradicted by conflicting study results. A study by Giovannucci and colleagues (1992),
for example, showed a 50 percent reduction in risk through the ingestion of fiber, while other cohort studies and randomized trials reported no significant association between fiber intake and the risk of developing CRC (Alberts et al., 2000; Fuchs et al., 1999; Michels et al., 2000; Schatzkin et al., 2000).

Additional behavioral risk factors of CRC include alcohol and tobacco use and physical inactivity. Studies consistently report significant associations of CRC with alcohol and tobacco use (Franco et al., 2005; Reid et al., 2003; Slattery et al., 2000; Zisman et al., 2006), with some reports of a twofold increase in risk. However, the magnitude of these effects varies widely throughout the literature, mostly as a result of measurement differences, duration, and frequency of use. More consistent are the findings on the relationship between physical activity and colon cancer, with higher risk related to lower levels of physical activity (Martinez et al., 1997; Potter, 1997; Potter & Hunter, 2002). Research explanations for this relationship point to the positive effects of physical activity on accelerating passage of colonic material, stimulating immune system responses, and reducing levels of harmful glucose and triacylglycerols in the intestines (Giovannucci, 1995; McMichael & Potter, 1985; Potter et al., 1993).

Family history of CRC is also associated with a significantly increased risk. The number of family occurrences and the degree of the familial relationship affect the magnitude of risk, with CRC occurrence among first degree relatives conferring the greatest risk. The two most common hereditary types of genetic syndromes are familial adenomatous polyposis and hereditary nonpolyposis or Lynch Syndrome, which typically develop at younger ages than the more common adenomatous polyps and are associated with significantly increased risk of CRC (Cao et al., 2002). However, only about 15
percent of cases are due to inherited mutations (Jackson-Thompson et al., 2006). Therefore, while medical screening guidelines do not typically recommend screenings until age 50, they can begin as soon as age 45 if patients have a family history of CRC (Rex et al., 2000; Zoorob et al., 2001).

Health studies confirm that early cancer detection through medical screenings significantly reduces CRC risk (Cotterchio et al., 2005; Segnan et al., 2005). These screening modalities commonly include the Fecal Occult Blood Test (FOBT), sigmoidoscopy, and colonoscopy. While trends show overall improvements in screening compliance over the last decade (Breen et al., 2001), many studies report that compliance remains a significant problem, with less than 50 percent of people at average risk having ever received any type of CRC screening (CDC, 2001; Neilson & Whynes, 1995). Many researchers attribute these trends to behavioral factors, such as the effects of the ‘Health Belief Model,’ which explains health-related actions according to individual beliefs and attitudes (Conner & Norman, 1996). However, these explanations underestimate the importance of access to health care and insurance on compliance levels, an issue addressed in this chapter.

While the individual risk factors are important to consider for early prevention and behavior modification, apart from ensuring screening compliance, the extent to which they actually contribute to CRC risk is debatable. Colorectal cancer can take up to 15 years or more to develop, a latency period during which many variables could contribute risk. Studies of CRC risk thus face the difficult task of controlling for all possible relevant risks that may develop over the course of a person’s life. Additionally, access to adequate medical and risk factor data will be limited by this long latency period.
As a result, epidemiologic studies of CRC are typically designed as retrospective studies that rely heavily on imperfect patient recall or incomplete patient medical records that introduce bias or confounding variables. Conversely, prospective cohort studies and randomized clinical trials have the advantage of collecting information as they occur, enabling temporal associations between exposures and outcomes. However, the latency period of CRC requires such studies to run for a number of years, rather than producing immediate results that can be used to improve CRC surveillance, prevention, and treatment options. The biggest prospective CRC study in the United States to date is the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer screening trial (NCI, 2007b), which began in 1992 and will run through 2016. Results from this trial should further clarify the effect of individual behaviors on CRC patterns, but the study will not be completed for another eight years.

### 2.1.2 Epidemiology by Race, Ethnicity and Gender

Colorectal cancer risk varies significantly by race and ethnicity, socioeconomic status, and gender (Erban, 2001; Hardcastle, 1996; Nakajima, 2003). Table 2.1 lists CRC age-adjusted incidence rates by gender, race and ethnicity from the years 2000 to 2004 as reported by the Surveillance Epidemiology and End Results (SEER) registries (NCI, 2008). All rates reported in this dissertation are age-adjusted to the 2000 standard million population and represent number of cases per 100,000 people. During that period, the age-adjusted CRC incidence rate was 51.6, with a rate of 60.8 for men and 44.6 for women. Among males, American Indian/Alaskan Natives had the lowest incidence rate (42.1), followed by Hispanics (47.5), Asian/Pacific Islanders (49.7), whites (60.4), and
blacks (72.6). The male CRC mortality rate averages 24, affecting male Asian/Pacific Islanders the least at a rate of 10.3 and black males the most at 33.4. Overall, female CRC incidence and mortality rates are lower than male rates (44.6 and 16.4 for female incidence and mortality), but present the same racial and ethnic patterning (NCI, 2008). During the same time period of 2000 to 2004, incidence rates were lowest among female Hispanics (32.9) and highest among black females (55). Mortality rates were also lowest among female Hispanics and Asian/Pacific Islanders (11.1 and 10.3) and highest among black females (22.9).

Table 2.1: U.S. CRC Incidence and Mortality by Race, Ethnicity, and Gender

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>Incidence Rates</th>
<th>Mortality Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>All Races</td>
<td>60.8</td>
<td>44.6</td>
</tr>
<tr>
<td>White</td>
<td>60.4</td>
<td>44</td>
</tr>
<tr>
<td>Black</td>
<td>72.6</td>
<td>55</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>49.7</td>
<td>35.3</td>
</tr>
<tr>
<td>American Indian/Alaskan Native</td>
<td>42.1</td>
<td>39.6</td>
</tr>
<tr>
<td>Hispanic</td>
<td>47.5</td>
<td>32.9</td>
</tr>
</tbody>
</table>

*Based on cases diagnosed in 2000-2004 from 17 SEER Geographic Areas*
*Rates represent cases per 100,000 people*

Survival estimates, which compare survival of people with a disease to survival of those without, are important indicators of cancer effects on the population. Survival rates are represented as the percentage of people who survive for a specific time period, usually five years, from the point of diagnosis or the start of treatment (NCI, 2008). Five-year survival estimates calculated from CRC cases registered through SEER for the years 1996 to 2003 show that men and women with CRC have a 64 percent survival rate. Reported survival rates by race, ethnicity and gender show a five-year survival rate of 65
percent for both white men and women and 55 percent for both black men and women. Group specific rates range from 66 percent for white men, 64 percent for white women, 56 percent for black men, and 54 percent for black women.

Five-year CRC survival rates also vary by stage, or severity, of cancer (Fazio, 2005). Cancer stage categories consist of in situ or localized, regional, and distant stages. In situ and localized cancers are characterized by cancer cells that are confined within the primary cancer site, which, in the case of CRC, is the large intestine. Localized stage cancers typically have better prognoses than regional or distant stage cancers that have spread to surrounding lymph nodes and other parts of the body (Mandelblatt, 1996; Potter & Hunter, 2002; Singh et al., 2003; Wu, 2001). In the nine SEER registries for CRC cases diagnosed between 1996 to 2003, roughly 40 percent of all cases were diagnosed as localized, 37 percent as regional, and 19 percent as distant stage (NCI, 2008). The stage of four percent of cases was unknown. The corresponding proportions of people who survived five years from the time of CRC diagnosis were 90 percent for localized, 68 percent for regional, 10 percent for distant stage, and 36 percent for unstaged CRCs. The steep decline in the survival of individuals with advanced stages of CRC underscores the importance of early detection for survival. Studies also show that among men and women, blacks are more likely than whites to be diagnosed with more advanced stages of CRC, and that they have lower five-year survival rates at every stage (Freeman, 2002; Irby, 2006; Ries, 2001; Wu, 2001). Table 2.2 presents survival and stage statistics by age, race and ethnicity for SEER-registered CRC cases diagnosed between 1996 and 2003.
Table 2.2: Percent CRC by Stage, Five-Year Survival Rates, Race, Ethnicity, and Gender for Nine SEER Registries, 1996 – 2003

<table>
<thead>
<tr>
<th>Stage</th>
<th>Localized</th>
<th>Regional</th>
<th>Distant</th>
<th>Unstaged</th>
<th>All Stages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41</td>
<td>36</td>
<td>19</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>Female</td>
<td>39</td>
<td>38</td>
<td>18</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>35</td>
<td>34</td>
<td>25</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Female</td>
<td>35</td>
<td>35</td>
<td>24</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>5-Year Survival</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>90</td>
<td>69</td>
<td>10</td>
<td>42</td>
<td>65</td>
</tr>
<tr>
<td>Female</td>
<td>91</td>
<td>70</td>
<td>11</td>
<td>31</td>
<td>65</td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>85</td>
<td>63</td>
<td>9</td>
<td>40</td>
<td>55</td>
</tr>
<tr>
<td>Female</td>
<td>84</td>
<td>63</td>
<td>8</td>
<td>40</td>
<td>55</td>
</tr>
</tbody>
</table>

Cancer tumor subatomic site locations influence both stage and survival estimates, but vary by race and ethnicity and gender. Cancer tumors can occur in different sections, or sub-sites, of the colon and rectum. The proximal colon, or the right colon, forms the beginning of the large intestine and is primarily made up of the ascending colon and transverse colon. Making up the remaining parts, the distal colon is the section nearest to the rectum and is known as the left colon. A study conducted by Demers et al. (1997) of the CRC cases diagnosed during 1973 and 1994 from the Metropolitan Detroit SEER site shows that five-year survival rates were generally lower among cases that originated in the right colon than they were for the left colon, and that blacks had greater cases of right colon cancers than whites. An Illinois tumor registry study of CRC cases diagnosed during 1986 and 1991 confirms these patterns (Nelson et al., 1997). That study reports that, among males, whites were 26 times more likely than blacks to develop CRC in the left colon, whereas blacks were 19 percent more likely than whites to develop CRC in the right colon and suffered more advanced stage CRCs. Similar epidemiological findings are reported by a number of studies that analyze subsite

Temporal trends in CRC rates also exhibit differences by race. Irby and colleagues (2006) found that between the years 1975 and 2002, CRC mortality rates decreased for whites and increased among blacks. While more recently, CRC mortality rates are declining across racial groups, SEER estimates report that between 1995 and 2004, the Average Percent Change (APC) in CRC incidence among whites decreased by 1.5 percent but by only 0.4 percent in blacks (NCI, 2008). The APC for CRC-specific mortality rates reflect similar differences, with a 2.5 percent decrease for whites and a 1.5 percent decrease for blacks. Colorectal cancer disparities in incidence, staging, mortality, and temporal trends indicate that the degree of risk for developing CRC is not uniform across race and ethnicity or gender, but that both black males and females have a substantially higher risk of advanced CRC and CRC-caused mortality.

2.1.3 Descriptive Epidemiology by Socioeconomic Status and Geographic Area

As a major and independent risk factor, SES significantly influences levels of advanced stage CRC and mortality, and many studies find that SES accounts for differences by race and ethnicity (Fazio, 2005; Mandelblatt, 1996; Marcella, 2001; Powell-Griner, 1999). The National Cancer Institute’s (NCI) extensive U.S. monograph series on SES and cancer examines the trends in rates of CRC between 1975 and 1999 using cases from the SEER registries (Singh et al., 2003). Using area poverty levels to represent SES groups, men and women from areas with the highest rates of poverty experienced slower declines in CRC-related mortality than did those in areas with the
lowest poverty levels. By the late 1990s, men and women living in high poverty areas had, respectively, a 12 and 7 percent greater rate of CRC mortality than did those in low poverty areas. Similarly, the study identified greater rates of advanced stage CRC in areas with higher poverty levels. Some study findings report even more exaggerated effects of area poverty on CRC outcomes. In the state of New York, for example, Mandelblatt and colleagues (1996) found that individuals living in the lowest SES areas were 45 percent more likely to be diagnosed with advanced CRC than persons living in the highest SES areas.

Studies that examine differences in regional patterns of CRC rates cite associations to demographic characteristics, SES, race and ethnicity, and individual health behaviors (Mandelblatt, 1996). In a national study of CRC spatial patterns between the years 1998 and 2001 by Lai and colleagues (2006), the Middle Atlantic states such as New Jersey, New York, and Pennsylvania had the highest incidence rates (62.4 per 100,000 people), while the Mountain states of Arizona, Colorado, Idaho, Montana, New Mexico, Utah, and Wyoming recorded the lowest rates (46.4 per 100,000 people). Higher proportions of older age groups, blacks, and males accounted for these regional differences, as did higher levels of undereducated and inactive populations. Other studies of CRC spatial variations at the regional and local scales identify clusters of higher than expected CRC rates associated with high percent poverty, working class, male gender, and black race (Dechello & Sheehan, 2007; Hsu et al., 2006; Ika et al., 2000; Jacquez, 2004; Krieger & Aierler, 1996; Pollack et al., 2006).

Unfortunately, many studies of geographic variations are purely spatial and control for few variables like age, race, or gender (Dechello & Sheehan, 2007; Pollack et
al., 2006). In these cases, results can be the effect of exogenous variables not controlled for in the study. Regardless, disease mapping of regional rates is important for public health surveillance purposes, and while it may not be possible to draw associations between demographic factors and risk, the reporting of descriptive cancer statistics is the first step toward cancer control.

2.2 Health Inequalities Research and Theoretical Approaches

Social and economic contexts significantly shape health-related behaviors, such as diet, physical activity, and tobacco use (Lantz et al., 1998). Research findings, for example, increasingly show how material and social inequalities affect risk for disease by guiding individual-level choices (Schwartz et al., 1999; Susser & Susser, 1996). Investigations of the relationship between structural patterns and health status employ a range of theories that frame research questions. This section briefly reviews some of the dominant theoretical and methodological approaches in health inequalities research.

A vast body of literature that has emerged over the last thirty years consistently shows the significant effects of social and economic status on health outcomes (Bartley, 2004; Marmot & Wilkinson, 2006). A seminal study, the Black Report (Townsend & Davidson, 1982), lead to the resurgence of health disparities research by providing overwhelming evidence that social class largely determines mortality. Since then, a number of population health studies have identified similar associations, linking premature mortality and illnesses (such as cardiovascular disease and cancer) to social and economic factors (CDC, 2005; Kaplan & Keil, 1993; Kivamaki et al., 2007; Lantz et
al., 1998; Lynch et al., 1997; Najman et al., 2006; Singh et al., 2002; Wardle & Steptoe, 2003). Health research predominantly measures health inequalities through differences in socioeconomic status.

SES is a multidimensional construct that refers to education, income, and occupation, and it is typically regarded as a marker of social class (Braveman et al., 2005; Lahelma, 2001). The level of SES defines the particular set of qualities possessed by individuals, the occupational opportunities available to them, and potential earning capacities. A study by Ross and Mirwosky (1999) finds that higher occupational status leads to a greater sense of personal autonomy, personal control, and social support. Lower SES, on the other hand, often leads to a greater risk of unemployment, lower income potential, social isolation, and poverty (Bartley, 2004; Budrys, 2003). Many studies show a direct relationships between risky health behaviors, such as smoking and poor eating habits, and lower SES, and these studies identify qualities of social class that affect a person’s ability to maintain good health or deal with sickness (CDC, 2005; Hardey, 1998; Marmot & Wilkenson, 2006; Slattery et al., 2000). The poor also experience more prolonged illnesses, face significant limitations on daily activities, and incur greater health care costs than the non-poor (Wagstaff, 2002).

Despite the findings of the Black Report and other studies, some scholars claim the absolute level of SES is not significant, but rather that the spread of wealth in society, or the relative deprivation, is most important in health outcomes (Marmot & Wilkenson, 2006). Wilkenson (1996) argues that the effects of relative deprivation supersede absolute wealth differences as the main contributing factor to health disparities in societies. Studies that look at the spread of wealth and the degree of income inequality
between and within countries identify significant associations with self-rated health, mortality rates, life expectancy, and chronic illness (Baum, 2004; Crepaz & Crepaz, 2004; Kaplan & Keil, 1993; Kennedy et al., 1998). These studies emphasize the psychosocial effects of wealth disparities on health status and downplay the role of material conditions (such as housing, transportation, and economic resources). Cassel (1976) describes how psychosocial stressors directly impact neuroendocrine production (which supports the body’s immune system) and indirectly impact emotional stability and the conduct of health damaging behaviors. While wealth disparities research provides evidence of the significance of psychosocial effects on health, the majority of studies heavily relies on large-scale, country-to-country ecological data, and do not typically analyze outcomes at smaller units.

Another method for analyzing these issues is the political economy of health approach. Political economists such as Coburn, Navarro, and Poland (Coburn, 2004; Kelleher, 2002; Navarro & Shi, 2001; Poland et al., 1998) argue that studies of wealth disparities overlook the importance of the social context of income inequality. These authors also claim that income, as a variable, is not used in these studies as a proxy to describe more complex social conditions. In particular, they argue that class-based inequalities are at the heart of health disparities, produced by the decline of welfare mechanisms, class and racial conflict, and the deterioration of social cohesion. Moreover, they point to the rising rate of unemployment, growing wage disparities, and devolution of supportive social services as major obstacles to health management. The political economy of health approach provides a greater contextualization of the political processes involved in health outcomes and is more fundamentally engaged with
components of the social structure. While this approach provides in-depth analysis of the structural causal pathways of health, it most often concentrates on describing macro-level processes, without much analysis at the local level. Additionally, this approach often fails to consider geographical processes that affect health, such as migration, residential patterns, and spatial distribution of resources (Macintyre et al., 2002).

Spatial processes and patterns are as important in health as are political economic features. The geography of health approach builds on traditional approaches to health inequalities research by specifically analyzing features of the physical and social environment. A growing number of health studies that emphasize spatial processes and that take into account the characteristics of where health outcomes occur provide evidence that these geographical features affect health outcomes separately from individual-based attributes (Bernard et al., 2007; Cummins et al., 2007; Macintyre, 2007; Macintyre et al., 2002). Studies that use geographical approaches to study the effects of social and economic deprivation report that deprived regions significantly raise the risk of mortality and the probability of individual engagement in health damaging behaviors, such as smoking (Diez-Roux et al., 1997; Ross & Mirowsky, 2001; Waitzman & Smith, 1998). Studies of geographical processes in health outcomes do not always produce consistent, or linear, findings (Braveman et al., 2005; Krieger et al., 1999). Conflicting research findings, however, may not negate the importance of these different study outcomes, but instead could reflect the unique circumstances of the local or geographical areas under study. Likewise, a geographical analysis of health outcomes identifies different associations depending on the scale or unit of analysis. For example, when differentiated by race and ethnicity and gender, rate differences in CRC at the national
scale are quite distinct from patterns at the regional scale, as described earlier in this chapter (Lai et al., 2006). Therefore, the choice of spatial scale and the unit of analysis (individuals versus groups) can produce significantly different results (and as noted earlier, are linked to problem of the ecological fallacy).

Geographical features describe both the physical and social attributes of a location. Macintyre and colleagues (2003) describe geographical features as the material infrastructure of a location, such as: built environments; its physical features, e.g. air quality; the availability of healthy environments, e.g. safe neighborhoods; the provisions of public and private services, e.g. education and transportation; as well as social and cultural relationships. Krieger (2005) and colleagues (2005) further define geography as being comprised of the compositional characteristics of individuals in addition to the contextual features of places. Research goals and aims will determine the specific place-based social and economic characteristics of the analysis. Traditionally, geographical analysis of health employs spatial analysis of health events, measured by geographical distribution, spatial concentration, and distance from medical services (Meade, 2000).

A dominant criticism of this approach (Conner & Norman, 1996; Kaufman, 1999; Krieger & Aierler, 1996; Link et al., 1998), however, is that while geographical methodologies capture important features of health, the approach often does not explain pathways to health outcomes and does not sufficiently develop or employ theories. Most recently, a growing number of researchers in the fields of medical geography, health geography, and social epidemiology have addressed this gap and provided more substantive theoretical discussions of the processes of and pathways to health (Bernard et al., 2007; Curtis & Jones, 1998; Davidson et al., 2008; Kaufman & Cooper, 1999;
Krieger, 2001, 2005; Macintyre et al., 2002). As an example, Krieger’s (2001; 2005) ecosocial theory of ‘embodiment’ posits that the material and social world are incorporated biologically and at multidimensional organizational scales (e.g. the biological, individual, family, community, and population levels). While Krieger draws from the traditional ‘Web of Causation’ theory in epidemiology, which describes a triangle of agent, disease, and environmental interactions, she goes further and makes connections between biology and social relationships. This type of chain-of-causation explanation, goes beyond exclusively traditional approaches and incorporates a multi-scalar and multi-theoretical approach to health.

Health inequalities research incorporates multiple and often overlapping approaches, each capturing significant aspects of health inequalities. What one approach may lack methodologically, another could supply through connections to theory, and vice versa. The potential strength of health inequalities research, therefore, lies not in the application of a singular traditional disciplinary approach, but in the integration of them, thus providing a more comprehensive approach to studying the deep complexities of health today.
2.3 Miami, CRC Patterns, and Health Insurance

2.3.1 Miami-Dade County

Miami-Dade is a county of approximately 2000 square miles, 32 municipalities, and over 2.4 million people (Census, 2000a). Its residents come from more than 121 countries and speak 68 different languages (MDC, 2008). The City of Miami and Hialeah are the largest municipalities, with 362,470 and 226,411 people, respectively. Since the 1960s, Miami-Dade County has transformed itself from a relatively small resort beach town to an international and globalized metropolitan area. Known as the ‘Gateway to Latin America,’ over 50 percent of its population consists of foreign-born residents (Census, 2000a; Miami, 2008), largely made up of Latinos. Among the Latino population, Cubans represent the majority ethnic population and comprise 46 percent of all foreign-born residents.

Miami-Dade County is characterized by extreme inequalities among racial and ethnic groups, particularly in relation to their social and economic characteristics. Black Caribbean immigrants, mostly Haitians, immigrated to Miami in large numbers due to severe economic deprivation and political violence in their home countries. The U.S. significantly marginalized immigrant Haitians, causing extreme social and economic hardship by deeming them economic refugees (even while many feared persecution for political reasons), a designation that denied them the level of U.S. protection and access to government provided services that they would have received had they been deemed political refugees (Portes & Stepick, 1993). Based on the 1951 Geneva Convention Article I, a political refugee is someone who is outside their country due to fear of persecution because of his or her race, religion, nationality, or membership of a particular
group and not merely due to economic hardship (UNHCR, 2001). The U.S. classified Cuban immigrants as political refugees, but not Haitians, which heavily influenced their respective social and economic achievements in the U.S. In 2000, Haitians had the lowest median income of any racial and ethnic group in Miami-Dade, including non-Haitian blacks, and were overrepresented in low-wage jobs (Sohmer, 2005).

Over the past four decades, the expanding globalized market and burgeoning migration flows have transformed the county economically and demographically. Tourism, international trade, and real estate are the leading economic sectors, with more than 100 financial institutions and over 300 “Fortune 500” multinational corporation headquarters located in the county (MDC, 2008). The City of Miami also is ranked third in number of foreign banking offices, after New York and Los Angeles (Nijman, 2000). While major growth and economic development afforded residents with opportunities for upward mobility, this transformation also led to political disenfranchisement and ethnically divided occupational sectors (Nijman, 2000; Portes & Stepick, 1993). Blacks have suffered more than any other racial group. Currently, native and foreign-born black residents in Miami-Dade County have the lowest levels of representation in local government, the highest rates of unemployment at 14 percent, and the greatest rates of poverty at 29 percent (Census, 2000a). Comparatively, non-Hispanic whites have unemployment and poverty rates of 5 and 9 percent, while Latinos have rates of 9 and 18 percent. In 2002, blacks owned 10 percent of all county firms, compared to 55 percent that were owned by Latinos (MDC, 2008).

Additionally, residential patterns in Miami-Dade County show racial and ethnic divisions that are characterized by highly segregated communities.
within the County, such as Liberty City, Overtown, Opa-Locka, Hialeah, Homestead and Little Havana, have high concentrations of single racial and ethnic groups. Boswell (1990) attributes the high levels of segregation partly to “white-flight,” racial discrimination, suburban development, the influx of immigrant Latinos, and the concentration of public housing projects in low-income black neighborhoods. Racial segregation is not only significant in itself, but also because of its relationship to other predictors of health outcomes. Massey and Denton (1993) argue that residential segregation is the fundamental cause of poverty. In their view, residential segregation promotes the creation of ghettos and isolates communities both socially and economically from educational and employment opportunities.

The Lewis Mumford Center of the State University of New York at Albany analyzed racial and ethnic neighborhood segregation using the 2000 census (Mumford Center, 2001). The Mumford Center calculated exposure and isolation indices that describe levels of racial and ethnic residential segregation and found that U.S. metropolitan areas are defined by a higher degree of residential separation than in 1990, with an average white person living in a neighborhood that is 80 percent white, and a only 7 percent black. Conversely, an average black individual lives in a neighborhood that is only 33 percent white and 51 percent black. The Center also calculated the segregation index or index of dissimilarity that captures the degree to which groups are evenly spread across census tracts. The index represents the percentage of one group who would have to move to attain an equally distributed residential pattern. According to the Center’s analysis of fifty metropolitan areas in 2000, Miami ranked 10th in black to
white segregation in the entire country with an index of 74, and was additionally found to be most resistant to racial integration.

Among all racial and ethnic groups, blacks had the highest index of residential segregation in the U.S in 2001 (Mumford Center, 2001). Residential patterns analyzed by the Center seem to have undergone only minor changes since the 1980s. Using the 1980 census, Boswell (1990) calculated segregation patterns using an index of dissimilarity that provides a detailed description of neighborhood residential patterns among racial and ethnic groups in Miami-Dade County. Blacks were concentrated mostly in the areas of Liberty City and Overtown. In 1980, seven out of every ten blacks lived in communities that were 70 percent black. Residential segregation in Miami-Dade County also occurred within various Latino groups and among broader racial and ethnic populations. Within the Latino demographic, residents of Mexican descent were, on average, the most segregated from other Latino groups, showing distinct residential separation from Cubans, and residing predominantly in the agricultural regions of Florida City and Homestead. When comparing Latinos to other racial groups, Cubans had the highest level of residential segregation, with primary concentrations in Little Havana and Hialeah.

Another critical issue facing Miami-Dade County is the severe lack of affordable housing, which is only complicated by the documented mismanagement of the Miami-Dade Housing Agency, the agency charged with (re)building affordable housing projects. In 2006, the Miami Herald exposed how the agency had misspent millions of dollars that were earmarked for housing developments in Liberty City, Little Havana, and other blighted areas (Cenziper, 2006). Forty percent of housing projects, worth a total of over
$80 million, have been cancelled, and others are delayed. These delays and
cancellations, combined with current conditions in the housing market that are putting
financial constraints on Miami-Dade residents, have created considerable obstacles for
the poor and disadvantaged in finding affordable housing.

Such problems increase social and economic vulnerabilities of Miami-Dade
County residents and affect, directly and indirectly, the overall health of the population.
As mentioned in the discussion on social and economic determinants of health, both
structural level variables and place-based socio-political characteristics highly influence
the ability to maintain good health. The situation in Miami-Dade County is such that
significant levels of social polarization, residential segregation, and unstable economic
conditions make health outcomes highly variable. As is the case with national and
regional patterns of CRC epidemiology, Miami-Dade County exhibits similar trends in
CRC outcomes by race and ethnicity, gender, and socioeconomic status. The following
sections discuss these patterns, as well as the powerful effect of health coverage.

2.3.2 CRC Patterns in Florida and Miami-Dade County

The descriptive epidemiology of CRC in Florida and Miami-Dade County is
similar to national rates and trends, with subtle differences. Age-adjusted mortality rates
and the average percent change (APC) are two important measures for CRC status, as
they indicate public health improvements in CRC outcomes. Since the 1970s, national
rates of CRC-specific mortality have dropped steadily, with the greatest reductions
occurring since the mid-1980s. As discussed above however, once rates are
disaggregated by race, ethnicity, and gender, this pattern varies significantly. National
CRC rates over time have shown greater declines in CRC-specific mortality in white men and women (at APCs of negative 2.2 and negative 1.8, respectively), than in black men and women (at APCs of negative 0.8 and negative 0.7 respectively) (Jemal et al., 2004). In the state of Florida, the overall APCs for CRC are broadly similar but show greater reductions in overall mortality, with an APC of 3.0 annually between the years of 1994 and 2003 (FCDS, 2005). Similar to national figures by race and gender, blacks in Florida have a smaller reduction in mortality than whites, the greatest difference being between black and white men (negative 3.0 for whites, and negative 0.9 for blacks). The Florida Cancer Data System (FCDS), the state of Florida’s cancer incidence registry, reports that since 1982, CRC-specific mortality trends show a 39 and 46 percent decline for males and females, respectively; however, black males had the only reported increase in CRC mortality since that time (FCDS, 2005). In 2003, white males and females had mortality rates of 18.6 and 12.1 per 100,000 men and women respectively, compared to rates in black men and women of 28.1 and 17.8 (FCDS 2005). In 2005, Miami-Dade County was one of five counties to lead all other counties in CRC-mortality, and surpassed the statewide rate by a small margin (16.33 vs. 15.61). The average percentage decline in mortality rates between the years 1994 and 2003 depicts a smaller decline for the county, at a rate of 2.1, than for the state.

The descriptive epidemiology of CRC often lacks rate reports by Latino/Hispanic subgroups and comparisons with other non-Hispanic racial categories. In one of the few studies of this kind, Howe and colleagues (2006) reviewed cancer data from 38 cancer registries across the country, distinguishing between Hispanic and non-Hispanic groups in their reporting of CRC rates. Historically, the Hispanic population has had lower
incidence and mortality rates than blacks and whites (CDC, 2007c); however, between the years 1999-2003, the Hispanic population in the U.S. had an APC similar to blacks at negative 1.0 and negative 1.1 respectively, while non-Hispanic whites had an APC of negative 2.2. This trend suggests that Hispanic CRC-specific mortality rates could exceed rates for non-Hispanic whites in the future.

### 2.3.3 Health Insurance, Cancer Screening and Miami-Dade County

Health insurance coverage and regularity of cancer screenings highly correlate with earlier diagnosis, lower risk for advanced carcinomas, and greater therapeutic care (Bond, 2002). Diagnostic screenings significantly reduce the risk for late stage cancer incidence and CRC-specific mortality (Fazio, 2005). In 1999, the Behavioral Risk Factor Surveillance System (BRFSS) showed that only 44 percent of the target screening population had ever received a Fecal Occult Blood Test (FOBT), sigmoidoscopy or colonoscopy within the recommended time period (CDC, 2001). In a study using the 1998 National Health Interview Survey (NHIS), Breen and colleagues (2001) linked screening compliance to higher levels of education, income, and insurance coverage. Results from analysis using individual and neighborhood-level indicators consistently found relationships among SES, health coverage, and screenings (Ionescu et al., 1998; Mandelblatt, 1996; Matthews et al., 2005; Neilson & Whynes, 1995; Parikh-Patel et al., 2006; Schwartz et al., 2003). Link and colleagues (1998) explain the association by arguing that possession of material resources, social networks, and knowledge enable people to employ protective strategies against illness.
In 2005, the rate of uninsured individuals had increased by 7 million people since the year 2000, totaling approximately 47 million, or 16 percent, of the U.S. population (DeNavas-Walt et al., 2006). Working age adults make up the majority of those uninsured, largely due to reductions in the percentage covered by employment-based health insurance, which dropped from 70 percent in 1987 to 59.5 percent in 2005. In the same year, Florida ranked close to last in health insurance coverage, with a total of 3 million uninsured residents (Nissen et al., 2006). Among racial and ethnic and foreign-born groups in Florida, Hispanics had the highest percent of uninsured individuals at 31.3 percent, followed by blacks at 24.5 percent. Only 18 percent of naturalized immigrants were uninsured versus 48 percent of non-citizen legal immigrants. These figures demonstrate the significant health coverage disparities in Florida and Miami-Dade County, indicating that millions of people are potentially without critical preventive care.

The connection between health coverage and cancer screenings cannot be overemphasized. Vogelaar and colleagues (2006) estimate that if screening usage were to increase to 70 percent of the target population, accompanied by greater usage of chemotherapy, by the year 2020 CRC mortality could decline by 49 percent. This study underscores the significant impact of health coverage and CRC screenings on CRC outcomes. Given the rising trends in the percentage of people uninsured, the overall decline in CRC mortality could soon be reversed. In addition to the reduction of employment-based health care, new restrictions on government assistance for medical care further limits access to preventive services. New eligibility requirements enacted by federal and state legislation place particular restrictions on legal immigrants (The Personal Responsibility and Work Opportunity Act, 1996). The population of Miami-
Dade County faces serious health concerns in light of the fact that it is home to a large and highly vulnerable group of people.

2.4 Spatial Analysis of Health

The spatial analysis of public health data increasingly is used to evaluate area differences in disease rates, identify spatial clusters of disease, and measure associations between environmental exposures and health outcomes. Spatial methods provide a useful approach to public health surveillance, as well as for more rigorous studies that explore pathways to disease causation (Pickle et al., 2006).

2.4.1 Spatial Representation and Analysis of Public Health Data

Public health data can be represented in different ways, as well as at many spatial scales (Waller & Gotway 2004). Public health studies represent spatial data in one of two ways: the first is by points on a map, such as patient address information; the second is by two-dimensional polygonal areas, using aggregated health information. In addition to different types of spatial representations, health data can be represented at multiple spatial scales. A research question, and to a great extent the accessibility of data, largely determines the scale of analysis.

Spatial analysis of public health data encompasses a range of spatial methods and techniques. Basic descriptive mapping of public health data is the first step to analyzing spatial trends, and most often consists of exploratory spatial data analysis (ESDA) (Lawson, 2007). This method describes the process pursued by John Snow in his investigation of the cholera outbreak in central London, whereby he simply mapped cases
to neighborhood blocks and explored potential disease pathways in the community (Cameron & Jones, 1983). To the public’s benefit, he identified the spatial patterning of cholera cases near water wells, which lead to the discovery that the cholera outbreak was spread through contaminated water wells. This is the most widely cited example of ESDA in public health, as it exemplifies its utility for pattern detection and hypothesis building in potential disease pathways.

While ESDA may have been the most powerful method of spatial analysis available at the time, advances in spatial statistical methods and tools since then have had a particularly large impact on the fields of public health and medical geography. Now, spatial clusters of disease occurrence can be evaluated using powerful statistical analysis through both global and local spatial cluster detection methods that scan rates and produce test statistics of significant disease rate variations from expected outcomes (Waller & Gotway, 2004). Examples of these methods include Moran’s I, Geary’s C, and the spatial scan statistic. A discussion of the particular techniques applied in each of these methods is beyond the scope of this chapter, but their application in the analysis of health inequalities is extremely relevant, especially considering the contribution of analyzing the spatial dimension of health outcomes.

In addition to spatial analysis of disease clustering, spatial methodologies and techniques include powerful methods for spatial modeling, whereby the relationships between disease patterns over a geographical area and potential pathways can be statistically evaluated (Lawson, 2007). Many types of spatial models can be applied to public health data, each having its own strengths and limitations. However, appropriate spatial models often are determined by the scale of analysis, the type of public health data
– rare cases versus common cases – and the theoretical approach of a researcher (Gardner et al., 1995). Despite such variation in spatial techniques and methodologies, the statistical analysis of public health data now can be evaluated spatially, contributing a deeper geographic understanding of public health patterns and relationships.

### 2.4.2 Area-Based Social and Economic Measures

Ecological studies typically incorporate area-based health data at the neighborhood, regional or national scales, because socioeconomic data are more widely collected at these scales (i.e., by the U.S. Census) and often are used by researchers to supplement patient-level data (Kreiger et al. 2005). In assessing neighborhood effects on health status, researchers affiliated with the Public Health Disparities Geocoding Project (Krieger et al. 2005) use census tracts to represent neighborhoods and report independent and unique effects at this scale in comparison to smaller or larger geographical scales.

Studies that employ spatial analysis of area-based health outcomes by economic and social characteristics often disagree regarding which variables are most appropriate to include. Partly, studies delimit variables by what is available and accessible to them; researchers also select variables based on a priori assumptions and theoretical approaches. Most commonly, spatial analysis of health draws social and economic indicators from the U.S. Census Bureau, as it provides extensive area-based information at many spatial scales. Health research commonly utilizes race, ethnicity, and gender to evaluate social effects on health, given that health disparities are distinctly characterized by these variables. Studies that analyze socioeconomic impacts on health employ a range of indicators to capture SES, such as educational attainment, employment status, income,
and poverty (Braveman et al., 2005; Krieger et al., 2005; Lannin et al., 1998; Lantz et al.,
population health surveys by Braveman et al. (2005) reports that variables such as income
and education capture somewhat different aspects of material conditions, and that neither
should be used as a proxy for the other. Therefore, social and economic variables may
not always be entirely independent of each other, but may still represent somewhat
different covariates.

2.4.3 Limitations and Considerations

Spatial analytical methods of area-based public health data are challenged by
issues of ecological fallacy, the modifiable areal unit problem (MAUP), limitations of
cross-sectional data, and spatial autocorrelation. Ecological fallacy occurs when
conclusions drawn from group-level analysis are applied to the individual level, or when
statistics for one scale are compared to statistics across scales (Waller & Gotway, 2004).
When using area-based variables as measures of exposure, it is erroneous to conclude that
the relationship estimated at one scale actually applies at more disaggregated scales. In
addition, ecological studies cannot determine direct causal relationships between disease
outcomes and exposures at the level of the individual, because they are observational and
not experimental studies.

There are two main features of the MAUP: the scale effect and the grouping or
zoning effect. The “scale effect” refers to the application of inferences from one areal
unit to an areal unit of a different spatial extent, while “zoning” effects refer to changes in
the geographic divisions of data, resulting in different variable associations (Waller &
Gotway, 2004). Findings at one scale, such as the community level, may not be consistent with findings at a different scale, such as the county level.

While cross-sectional studies are widely utilized in ecological studies, they often are limited by their reference to a single snapshot in time and space. Schwartz and colleagues (1999:29) note that “an individual’s risk of disease is an ever changing probability based on interactions among the individual’s personal history, biology, and physical and social environments.” They recommend utilizing longitudinal data or cross-sectional datasets at multiple time periods. However, this approach has its own set of biases and assumptions, mainly that the same population is evaluated over time.

A separate issue in the application of spatial analysis of public health data is the concern over spatial autocorrelation, defined as:

the correlation among values of a single variable strictly attributable to the proximity of those values in geographic space, introducing a deviation from the independent observations assumption of classical statistics (Griffith, 2003 p. 3)

Spatial autocorrelation presents unique issues of spatial dependency when measuring relationships between disease and associative affects that are distributed in space, as it violates the assumption that observations are independent from one another. The presence of positive spatial autocorrelation means that contiguous, geographical areal units are similar because of their proximity. These effects can inflate the statistical coefficients of independent variables and must be accounted for in spatial models in order to reflect accurate measures of association.
2.5 Summary

Colorectal cancer is a serious national concern, and presents highly differentiated outcomes by race and ethnicity, gender, and area-level poverty. Miami-Dade County faces a distinct set of CRC vulnerabilities as it has significantly high proportions of people who lack health insurance and one of the highest county concentrations of people with late stage CRC incidence and CRC mortality in Florida. Public health surveillance of current CRC trends in Miami-Dade County can benefit particularly from the analysis of the spatial distribution of CRC outcomes at the neighborhood level to detect areas with higher than expected cases. Because CRC is largely preventable, the identification of these clusters can help to target prevention efforts. As discussed, utilization of interdisciplinary approaches more effectively captures patterns and processes of health outcomes that are affected by contextual features of areas, which include social and economic characteristics. Given these factors, this dissertation spatially explores the distribution of CRC in Miami-Dade County and particular social and economic features of neighborhoods in relation to CRC outcomes.
CHAPTER III:

WELFARE REFORM EFFECTS ON ELDERLY IMMIGRANTS

3.0 Overview

This dissertation assesses the effects of race, ethnicity, and social status on health outcomes of colorectal cancer. The immigrant population, and particularly elderly immigrants, are one population that are particularly likely to have lower incomes and to be members of racial and ethnic minority groups, which are often associated with lower levels of health insurance and decreased access to health care. As previously discussed, these rates have been associated in literature with worse CRC outcomes. Although this study was unable to assess differences in health insurance status between 1990 and 2000, some evidence suggests that changes to national welfare laws in 1996 may have influenced health outcomes among elderly, non-citizen legal immigrants in Miami-Dade County by reducing the number of these individuals that qualify for government-sponsored health care assistance programs.

In 1996, the United States Congress enacted the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), which cut federal spending on means-tested benefits – those social assistance programs that are based on an individual or family’s income and assets – by $54 billion dollars (Weil & Finegold, 2002). A total of $23.8 billion, or 44% of all estimated savings from PRWORA’s changes, resulted from cuts to federal and state benefits for low-income legal non-citizens (Binstock & Jean-Baptiste, 1999; Weil & Finegold, 2002). The Act mandated immediate denial of Supplemental Security Income (SSI) and Medicaid for immigrants arriving after 1996, requiring five years of legal residency before those immigrants would be eligible for such
benefits. It also changed eligibility criteria for legal immigrants who had arrived prior to 1996 (Binstock & Jean-Baptiste, 1999). In addition to these new federal requirements, the Act delegated greater authority to the states, allowing them to enact more restrictive qualification requirements; it also shifted greater financial responsibility for funding the programs to the states (Zimmerman & Tumlin, 1999). These changes had a significant impact on those legal immigrants who are most vulnerable to chronic disease: the elderly. This chapter begins by describing the new eligibility criteria for legal immigrants and how welfare policy changes significantly impacted elderly immigrants’ health care coverage and access to medical care. Based on an analysis of the literature and demographic statistics in Miami-Dade County, the likely effects of PRWORA on health insurance status of the large elderly legal immigrant population in Miami-Dade County are analyzed. It is theorized that reductions in access to care due to elderly legal immigrant’s loss of health insurance have had an impact on CRC outcomes in the county.

3.1 New Public Assistance Eligibility Criteria Under PRWORA

PRWORA enacted sweeping changes to all federal and state means-tested benefit programs (The Personal Responsibility and Work Opportunity Act, 1996). One of the major changes replaced Aid to Families with Dependent Children (AFDC) with a time-limited assistance program called Temporary Aid to Needy Families (TANF). Under the Act, the federal government would also no longer fund food stamps for any legal immigrant, deferring to the states to continue these benefits (Weil & Finegold, 2002). A few states enacted legislation to temporarily continue provision of food stamps to
immigrants. In addition to these structural programmatic changes, the Act revamped eligibility requirements for SSI and Medicaid for immigrants living in the U.S.

Under PRWORA, eligibility criteria for federal and state benefits created two classes of immigrants, “qualified” and “unqualified.” Unqualified status applies to those foreign-born residents who live in the United States without proper documentation. This group is barred from all federal and state benefits, with the primary exception of emergency health care (Weil & Finegold, 2002). Qualified immigrants are made up of legal residents, refugees, assylees, and foreign-born residents whose deportation is being withheld (Friedland & Pankaj, 1997). All immigrants except political asylees must have a U.S. sponsor to apply for legal residency status.

Among qualified immigrants, the Act imposed a set of additional restrictions and exemptions, treating pre-enactment residents differently from those arriving in the U.S. after 1996, the year the law was passed. For pre-enactment legal immigrants receiving SSI in 1996, the law extends benefits until cases are reviewed by the state, at which point the beneficiary must demonstrate 40 quarters (10 years) of U.S. employment to continue to receive benefits (Weil & Finegold, 2002). The law further requires pre-enactment immigrants who were not receiving benefits in 1996, but who applied thereafter for SSI, Medicaid, or TANF, to demonstrate 40 quarters of U.S. employment. Additionally, before PRWORA, recipients of SSI were automatically eligible for Medicaid; PRWORA eliminated this automatic eligibility and required immigrants to qualify for Medicaid through other criteria (Binstock & Jean-Baptiste, 1999).

PRWORA also eliminated SSI benefits entirely and imposed a requirement of five years of legal residency before a post-enactment, qualified legal immigrant is even
eligible for Medicaid or TANF. Once the five-year threshold is met, the state must assess the financial need of the immigrant. However, PRWORA mandates that the sponsor’s income and assets must also be reviewed, a process called “deeming,” before an immigrant can be found to be eligible for benefits (Angel, 2003; Friedland & Pankaj, 1997). Although deeming existed prior to the Act, it was only required for the first five years of residency; the new law extends deeming until an immigrant is either naturalized or becomes eligible for social security benefits through 40 quarters of U.S. employment.

3.2 PRWORA and Elderly Immigrants

Elderly non-citizens are particularly vulnerable to the changes effected by PRWORA. This is also a group that, based on literature, is particularly susceptible to negative CRC outcomes. In the United States, legal immigrants over the age of 65 – a group consisting of 1.2 million people – have limited options for earning income and have great need for financial and medical support services (Friedland & Pankaj, 1997). Elderly legal immigrants have greater long-term healthcare needs than elderly naturalized or native-born citizens (36% vs. 17%) (Binstock & Jean-Baptiste, 1999). Yet, compared to elderly naturalized immigrants, elderly non-citizens are twice as likely to be uninsured, and they are more likely to experience higher rates of poverty and to receive public assistance (Friedland & Pankaj, 1997; Fronstin, 2005). In fact, as of December 1996, almost half of elderly legal immigrants were poor enough to be receiving SSI. Prior to PRWORA, 40 percent of immigrant welfare users were either elderly legal immigrants or refugees, but those two classes constituted only 21 percent of the total immigrant population (Fix & Zimmerman, 1995).
PRWORA restricted eligibility to the public aid programs most utilized by elderly legal immigrants. In particular, the Act limited access to SSI for pre-enactment legal immigrants and denied access altogether to post-enactment immigrants. Before PRWORA, legal elderly immigrants’ greatest level of welfare use occurred through the SSI cash assistance program, largely because this program did not require 40 quarters of employment to be eligible for benefits, as opposed to social security benefits which did require that duration of employment (Fix & Zimmerman, 1995). In the year prior to enactment of PRWORA, while about three quarters of pre-enactment elderly non-citizens receiving SSI benefits also received social security income, 24 percent were eligible only for SSI assistance (Friedland & Pankaj, 1997).

Additionally, PRWORA’s severance of the linkage between Medicaid and SSI eligibility, along with more restrictive eligibility requirements for Medicaid, poses the greatest threat to elderly immigrant’s access to care. Medicaid is the primary source of nursing home care and long term care for the elderly legal immigrant (Binstock & Jean-Baptiste, 1999). Medicaid is a means-tested entitlement program, jointly funded by the federal and state governments, that provides medical care for low-income pregnant women and children, as well as the elderly, blind, or disabled (NCSL, 1997). Even before PRWORA, legal elderly immigrants predominantly used Medicaid rather than Medicare, the federal health care program established primarily for those 65 years of age or older, because these immigrants had not accumulated the number of working years required to qualify for Medicare. For example, in 1995, about a quarter of elderly legal immigrants had not worked long enough to be eligible for Medicare (Friedland & Pankaj, 1997). In the same year, Medicaid covered 24 percent of elderly non-citizens as opposed
to only 12 percent of elderly citizens. In contrast, Medicare covered only 77 percent of elderly non-citizens, compared to 98 percent of elderly citizens. While most elderly legal immigrants utilize a combination of Medicaid and Medicare, about 16 percent of these non-citizens, in contrast to only 1% of elderly citizens, relied on Medicaid alone (Friedland & Pankaj, 1997). These statistics demonstrate that elderly legal immigrants who rely on Medicaid alone are the group most likely to lose health care coverage altogether due to PRWORA, since any non-citizen who qualified for Medicare before 1996 would also qualify for that program after the passage of PRWORA.

The rapid growth of the population of elderly immigrants has amplified the importance of the problems related to access to health care coverage for that population. Between 1990 and 2000, roughly 440,000 individuals 65 years of age or older immigrated to the U.S. (Friedland & Pankaj, 1997). While a majority of elderly legal immigrants arrived to the U.S. before the age of 65, roughly 22 percent of them arrived after the age of 65 (Friedland & Pankaj, 1997). These newly arrived elderly immigrants made up roughly half of the total elderly legal immigrant population in 2000 (Census, 2003). In 1996 alone, 67,000 immigrants over 60 years and older were legally admitted to the U.S. (Angel, 2003). Wilmoth and colleagues (1997) have attributed this trend largely to family reunification efforts. This recent influx of elderly legal immigrants has expanded the class of post-enactment immigrants who are not eligible for public assistance because they have not acquired the 40 quarters of employment necessary for Medicare, are not eligible for SSI or food stamps due to PRWORA’s elimination of such benefits for post-enactment immigrants, and have to wait five years before even
becoming eligible for Medicaid or TANF under the strict new requirements that require consideration of the income of both the immigrant and the immigrant’s sponsor.

Reports on the impact of PRWORA on Medicaid participation do not provide data or results detailing specific changes in participation among elderly legal immigrants. Most studies focus on immigrants of working age and do not always distinguish between non-citizen and naturalized immigrants. Moreover, due to wide variability in state-level welfare changes, researchers typically delimit their studies to state-level units of analysis. To date, a number of studies look at Medicaid usage of those under the age of 65, but the effects of welfare reform on elderly non-citizens in particular remain under-studied. Fix and Passel (1999) did publish a rare study that attempted to measure the impact of PRWORA on the use of Medicaid by elderly legal immigrants. However, while that analysis reported no significant effects, it was conducted only two years after the laws’ enactment, which may not have allowed adequate time for an accurate assessment of the Act’s longer term effects.

Despite the research gap in PRWORA’s impact on elderly legal immigrants, many studies have convincingly connected health care coverage to health status, as reported in this chapter. Moreover, no conclusive evidence of a change effected by PRWORA or of an increase in uninsured rates among elderly legal immigrants exists, other studies suggest that the number of uninsured foreign-born individuals is growing. Historically, immigrant populations in the United States have been more likely to be uninsured than the native born population, but temporal trends show that immigrants increasingly make up a greater proportion of the uninsured. Between the years 1998 and 2003, immigrants accounted for 86 percent of the growth in the uninsured (Fronstin,
Additionally, non-citizen immigrants are more likely to be uninsured than naturalized immigrants. In 2003, nearly half of the foreign-born non-citizens were uninsured. Recent immigrants also typically have higher proportions of the uninsured. Among immigrants who arrived between 2000 and 2004, 54 percent were uninsured compared to 21 percent of immigrants who arrived before 1970 (Fronstin, 2005).

3.3 PRWORA and Elderly Legal Immigrants in Florida and Miami

Because of the uneven distribution of elderly non-citizens across the country and differences in state welfare laws, the effects of PRWORA on elderly non-citizens are not evenly distributed across the country. The Act likely impacted Florida disproportionately due to the large immigrant population of Florida. As noted above, PRWORA effectively shifted a large portion of responsibility for welfare assistance programs to the states. Zimmermann and Tumlin (1999) referred to the uneven, fragmented, and exclusionary state assistance programs enacted by the states after PRWORA as “patchwork policies,” and they claim that immigrant exceptionalism, which singles out immigrants for differential treatment, characterizes most state policies. While states including California and Massachusetts enacted policies that restored many benefits to legal immigrants, states such as Texas and Florida passed state aid packages that did not fully restore the federal aid cuts (Quint et al., 1999). These changes resulted from the “New Federalism” movement that is characterized by the shift of power in policy making decisions from the federal to the state level (Golden, 2005). This shift led to a general devolution of federally and state provided public benefits. The fact that each state determined its own
welfare policy also led to greater inequality in immigrants’ access to public assistance across states.

The elderly immigrant population of Florida, and of Miami-Dade County in particular, is among the largest in the country and, as a result, PRWORA threatens access to public assistance to a disproportionately large number of elderly immigrants in this State and County. Roughly 60 percent of elderly legal immigrants are concentrated in California, Florida, and New York. Florida has approximately 160,000 elderly legal immigrants making up 6 percent of the state’s elderly population, of which a third reside in Miami-Dade County (Census, 2006a). Around the time of the enactment of PRWORA, the Florida Department of Children and Families estimated that more than 70,000 legal immigrants would be affected by PRWORA’s changes, many of them elderly immigrants receiving SSI assistance (Quint et al., 1999). Additionally, in July and August of 1997, a total of 55,000 non-citizens were terminated from the food stamps program, with benefits being temporarily reinstated for only 2,100 people under a one-year state-funded program.

Some evidence suggests that Florida’s welfare policy programs added additional obstacles that may have prevented legal immigrants from obtaining state-funded aid or having public assistance restored. In response to PRWORA’s delegation of funding and policy implementation responsibilities, Florida passed four different versions of the policy in the following year (Quint et al., 1999). The state policies were intended to encourage naturalization among legal immigrants, thereby giving them greater access to state-funded public assistance. This created a backlog of cases at the Immigration and Naturalization Service, delaying the process of naturalization for many immigrants for up
to a year. However, the additional state restrictions and the multiple changes to the policies also added to the confusion among immigrants regarding eligibility requirements and exemptions. The Kaiser Family Foundation (Feld & Power, 2000) conducted a study of immigrants’ access to health care after welfare reform, using focus groups in four major cities, including Miami. The Miami focus group primarily consisted of Haitian and Cuban immigrants who expressed difficulty in obtaining Medicaid coverage and who believed that applying to the program would threaten their ability to sponsor family members for immigration in the future. Others complained that fewer doctors were willing to see Medicaid patients and that the quality of care was substandard in comparison to care received in Cuba.

These concerns that perceptions of the changes resulting from the PRWORA decreased access to care in Florida are supported by studies at the national level. Studies produced by Fix and Passel (1999; 2002) confirm that non-citizens’ use of public benefits declined after PRWORA was enacted. Community surveys suggest that this decline was not only related to strict new eligibility requirements, but that the welfare changes created a “chilling” effect among legal immigrants who were confused by policy changes and fearful that applying for public benefits would affect their immigration status (Capps et al., 2002). In addition to welfare policy changes, the Deficit Reduction Act passed by Congress in 2006 placed new requirements for Medicaid eligibility that mandated proof of citizenship status or immigrant qualifications of all new applicants (Pear, 2007). Furthermore, the Bush administration added an additional hurdle to this policy by requiring submission of original documents or of copies certified by the issuing agency. These policies were intended to reduce fraudulent applications for Medicaid by illegal
immigrants. However, since the policy was enacted, the number of children and parents on Medicaid declined by 39,000, believed to primarily consist of American citizens who would qualify (Pear, 2007).

3.4 Conclusion

This chapter describes the policy changes of PRWORA and how they may impact elderly legal immigrants. The creation of new state powers in funding and eligibility requirements significantly affect elderly legal immigrants who have limited earning alternatives and significant health care needs. Historically, uninsured rates are higher among elderly legal immigrants than their citizen counterparts. Establishment of distinct criteria within the immigrant population as it applies to qualified immigrants and to pre- and post-enactment immigrants in effect creates tiers of immigrant classes for which public assistance is either offered, limited, or denied.

This dissertation examined the incidence and geographic distribution of colorectal cancer (CRC) in Miami-Dade County. Particularly relevant to the impact of PRWORA on elderly immigrants, the study findings, reported in more detail below in Chapter 5, show a statistically significant relationship between age and CRC health outcomes, such that older individuals are more likely to have late stage CRC incidence and a higher incidence of CRC-specific mortality than younger individuals. Although this dissertation does not analyze the extent to which health insurance status influences these outcomes due to limitations in the FCDS’s reporting of insurance status, as discussed in Chapter 2, population-level public health studies consistently report significant associations between lack of health insurance and the prevention of CRC through screenings (Potosky, 1998;
Powell-Griner, 1999; Roetzheim et al., 2000). Based on current population statistics, while Miami-Dade county, with a population of 2.3 million, constitutes approximately 0.8 percent of the nation’s total population, Miami-Dade’s elderly legal immigrant population of 50,601 is approximately 4.2 percent of the nations’ total population in that category (Census, 2000b, 2006a). These figures suggest that, to the extent PRWORA has adversely affected elderly legal immigrants’ access to health care, Miami-Dade County’s large elderly immigrant population is likely to experience significant health effects due to CRC.

Despite this dissertation’s inability to specifically measure the relationship among health insurance status, elderly legal immigrants, and CRC outcomes, health officials in Miami-Dade County should focus public health surveillance resources on this group with the intention of targeting this population for intervention and early treatment programs. Because such a large percentage of the country’s elderly legal immigrant population resides in Miami-Dade County, health officials should be particularly cognizant of the risks to this group, and endeavor to institute local and state-wide programs that would guarantee appropriate CRC screenings, as well as other health care prevention and treatment services, regardless of the individual’s health insurance status. Moreover, additional studies regarding PRWORA’s effect on Medicaid health coverage of this group are warranted, and Miami-Dade County would be a particularly suitable site for research in light of its demographics.
CHAPTER IV:
RESEARCH DESIGN AND METHODS

4.0 Overview of Study Design

This study examines spatial patterns of disease clusters to identify neighborhoods in Miami-Dade County with statistically significantly higher than expected numbers of cases of late-stage CRC incidence and CRC-specific mortality, and measures how neighborhood SES and racial and ethnic composition affect CRC outcomes. Based on these analyses, the researcher considers how social and economic factors influence CRC outcomes. Such an analysis is important not only to understand better how these underlying conditions can result in negative health outcomes, but also to provide health organizations and public health agencies with data that will allow them to better allocate resources to diagnose and treat this preventable form of cancer among those who are most at risk of late diagnosis and poor outcomes.

The study drew its data on diagnosed CRC cases in Miami-Dade County from FCDS and then analyzes these cases at the census tract level using census population data and geographic files for census years 1990 and 2000. In order to compare trends across time, the study examines diagnosed cases of CRC late-stage incidence and CRC-related mortality between the years 1988-1992, referred to as time 1 (T1), and between 1998-2002, or time 2 (T2). This study employs an observational design using cross-sectional population data from two points in time. It applies the spatial scan statistic to test for significant clusters of CRC and the negative binomial regression analysis to measure neighborhood-level social and economic covariate associations with particular CRC outcomes. After controlling for differences in population size and age distribution for all
Miami-Dade tracts, the researcher expected to find significant clusters of higher than expected late-stage CRC incidence and CRC mortality in both time periods. Additionally, the researcher expected to find significant associations of CRC outcomes with tract-level economic and social variables, such as percent black, percent Hispanic, and percent below poverty, for both time periods, after controlling for population size, age, and spatial autocorrelation.

4.1 Data Sources

4.1.1 The Florida Cancer Data System (FCDS)

This study’s primary data source was FCDS, which provided incidence data for registered cases of diagnosed late stage CRC incidence and CRC-specific mortality data for the years 1998-1992 and 1998-2002. FCDS is a statewide, population-based cancer incidence registry created by the State of Florida Department of Health in 1978. Once implemented, the State contracted the University of Miami’s Miller School of Medicine to maintain the Registry. The FCDS is wholly supported by the State of Florida Department of Health (DOH), and the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC). It is based in the Sylvester Comprehensive Cancer Center at the University of Miami’s Miller School of Medicine. The Registry obtains its data on cancer cases from patient records as reported by certified cancer registrars from hospitals as well as from radiation therapy facilities, private physicians, free-standing ambulatory surgical facilities, and death certificates (FCDS, 2008).
The Florida Department of Health’s Vital Statistics Office provided cancer-specific mortality data. On July 6, 2007, the office approved the mortality data for re-release for use in this study. The Florida Department of Health Institutional Review Board (IRB) approved this study under “Expedited” review (IRB Protocol # H06105), and the University of Miami IRB granted the study “Exempt” review status (IRB Protocol # 20060518). The views expressed herein are solely those of the researcher, and do not necessarily reflect those of the contractor or the State of Florida DOH.

This dissertation analyzed a total of 13,599 de-identified CRC diagnosed cases (cases not identifying personal patient information such as name or social-security number) that were extracted for T1 (n = 6,129) and T2 (n = 7,470). Cancer primary site data (typology and morphology) were coded by the FCDS using the third *International Classification of Diseases: Oncology* (ICD-0-3) for cancer of the colon and rectum. Mortality cases were coded using the ICD-9 codes for all deaths that occurred before 1999, and ICD-10 codes for all deaths that occurred after 1999 (CDC, 2007b). Table 4.1 lists the edition-specific codes for CRC incidence and mortality used in this study.

**Table 4.1: International Classification of Diseases for Colorectal Cancer**

<table>
<thead>
<tr>
<th>Malignant neoplasm of colon</th>
<th>ICD-9</th>
<th>ICD-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatic flexure</td>
<td>153.0</td>
<td>C18.3</td>
</tr>
<tr>
<td>Transverse colon</td>
<td>153.1</td>
<td>C18.4</td>
</tr>
<tr>
<td>Descending colon</td>
<td>153.2</td>
<td>C18.6</td>
</tr>
<tr>
<td>Sigmoid colon</td>
<td>153.3</td>
<td>C18.7</td>
</tr>
<tr>
<td>Cecum</td>
<td>153.4</td>
<td>C18.0</td>
</tr>
<tr>
<td>Appendix</td>
<td>153.5</td>
<td>C18.1</td>
</tr>
<tr>
<td>Ascending colon</td>
<td>153.6</td>
<td>C18.2</td>
</tr>
<tr>
<td>Splenic flexure</td>
<td>153.7</td>
<td>C18.5</td>
</tr>
<tr>
<td>Other</td>
<td>153.8</td>
<td>C18.8</td>
</tr>
<tr>
<td>Colon, unspecified</td>
<td>153.9</td>
<td>C18.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Malignant neoplasms of rectum, rectosigmoid junction</th>
<th>ICD-9</th>
<th>ICD-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectosigmoid junction</td>
<td>154.0</td>
<td>C19</td>
</tr>
<tr>
<td>Rectum</td>
<td>154.1</td>
<td>C20</td>
</tr>
</tbody>
</table>

*Source: Centers for Disease Control, Classification of Diseases (2007b)*
This study particularly focused on late stage CRC and CRC-specific mortality cases. These two variables were selected because they represent the more serious outcomes of CRC and thus provide good data for analyzing how geographical conditions may relate to negative CRC outcomes, and may permit further insights into how those conditions and outcomes relate to the availability of preventive care, screenings and effective treatment. For tumors diagnosed between 1988 and 2000, the FCDS reported cancer stage based on the Surveillance Epidemiology and End Results (SEER) General Summary Stage 1977, a coding system for staging of cancer. For tumors that occurred after 2000, the FCDS identified cancer stage using the SEER General Summary Stage 2000. CRC late stage cases include both regional and distant stage cancers. CRC-specific mortality cases represent patients diagnosed with CRC during one of the two study periods who subsequently died from the disease. In T1, 3,311 cases (or 54 percent of all cases) were registered as late stage CRC, and 2,281 cases (37 percent) resulted in death. A total of 1,036 cases (17 percent) were registered in T1 as of unknown CRC stage. In T2, 3,782 cases (or 51 percent) were registered as late stage CRC and 1,961 cases (26 percent) resulted in death due to CRC. 1,245 cases (16 percent) were registered in T2 as of unknown CRC stage.

The FCDS also contained patient demographic and residential information. The database recorded patients’ date of birth, gender, race, Hispanic origin, marital status, primary site, age at diagnosis, vital status, cause of death, and date of death. Residential location at the time of diagnosis was initially designated by census year 2000 tracts for the two study periods. The FCDS provided geocoding and census tract assignments.
However, preliminary review of the geocoded data by the researcher identified inconsistencies in the tract assignments, prompting an investigation into their quality and accuracy. The State of Florida DOH, as well as FCDS, agreed to re-release the dataset with patient address information, as well as the 1990 census tract assignments, pending amendment to the original IRB protocol. This study received an amended approval for release of patient address information on January 23, 2008 (amendment #1, H06105).

Table 4.2 lists the file layout provided by FCDS under this amended approval.

**Table 4.2: File Layout of FCDS Data Provided Under January 23, 2008 Release H06105, Amendment #1**

<table>
<thead>
<tr>
<th>*NAACCR Item #</th>
<th>NAACCR Item Name</th>
<th>Start</th>
<th>Length</th>
<th>Year Start-End</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>Primary Site</td>
<td>1</td>
<td>4</td>
<td>1981</td>
</tr>
<tr>
<td>a:759</td>
<td>SEER Summary Stage 1977</td>
<td>10</td>
<td>1</td>
<td>1995</td>
</tr>
<tr>
<td>a:760</td>
<td>SEER Summary Stage 2000</td>
<td>20</td>
<td>1</td>
<td>2001</td>
</tr>
<tr>
<td>230</td>
<td>Age at Diagnosis</td>
<td>30</td>
<td>3</td>
<td>1981</td>
</tr>
<tr>
<td>l:360</td>
<td>2000 Census Tract</td>
<td>40</td>
<td>6</td>
<td>1981</td>
</tr>
<tr>
<td>i:365</td>
<td>2000 Census Tract Certainty</td>
<td>50</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>250</td>
<td>Birthplace Geocode</td>
<td>60</td>
<td>3</td>
<td>1981</td>
</tr>
<tr>
<td>160</td>
<td>Race</td>
<td>70</td>
<td>2</td>
<td>1981</td>
</tr>
<tr>
<td>190</td>
<td>Ethnicity</td>
<td>80</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>150</td>
<td>Marital Status</td>
<td>90</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>310</td>
<td>Usual Occupation</td>
<td>100</td>
<td>100</td>
<td>1995</td>
</tr>
<tr>
<td>320</td>
<td>Usual Industry</td>
<td>210</td>
<td>40</td>
<td>2001</td>
</tr>
<tr>
<td>2220</td>
<td>FCDS Tobacco Use</td>
<td>300</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>z:1910</td>
<td>Cause of Death</td>
<td>310</td>
<td>4</td>
<td>1981</td>
</tr>
<tr>
<td>x:2220</td>
<td>Date of Death</td>
<td>320</td>
<td>8</td>
<td>1981</td>
</tr>
<tr>
<td>390</td>
<td>Diagnosis Year</td>
<td>330</td>
<td>4</td>
<td>1981</td>
</tr>
<tr>
<td>h:2220</td>
<td>Vital Status</td>
<td>340</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>220</td>
<td>Gender</td>
<td>345</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>110</td>
<td>1990 Census Tract</td>
<td>350</td>
<td>6</td>
<td>1981</td>
</tr>
<tr>
<td>364</td>
<td>1990 Census Tract Certainty</td>
<td>360</td>
<td>1</td>
<td>1981</td>
</tr>
<tr>
<td>2330</td>
<td>Address at DX – Street</td>
<td>370</td>
<td>40</td>
<td>2001</td>
</tr>
<tr>
<td>1810</td>
<td>Address at DX – city</td>
<td>410</td>
<td>20</td>
<td>2001</td>
</tr>
<tr>
<td>m:2220</td>
<td>Address at DX- state</td>
<td>430</td>
<td>3</td>
<td>2001</td>
</tr>
<tr>
<td>1830</td>
<td>Address at DX – zip code</td>
<td>440</td>
<td>5</td>
<td>2001</td>
</tr>
</tbody>
</table>

* North American Association of Central Cancer Registry
All cases that occurred in T1 were assigned to the 1990 census tracts provided by FCDS, as the researcher identified no inconsistencies in geocoding for that census year. Due to the inconsistencies in the 2000 census tract assignments, the researcher geocoded all patient address information for cases occurring in T2 to Miami-Dade street and parcel layer files obtained from the City of Miami and Florida International University using ArcGIS 9.2 (ESRI, 2007). A total of 6,814 patient addresses were geocoded from the T2 cancer dataset with a match rate of 92 percent in comparison to a 35 percent match rate from the 2000 census tract geocodes provided by FCDS. The remaining cases were assigned using FCDS 2000 census tract assignments.

4.1.2 US Census Bureau Variables

Census tract population variables for Miami-Dade County were extracted from the U.S. Census Bureau website (1990; 2000b) for the years 1990 and 2000, and were attached to cancer data files for T1 and T2, respectively. Census tract shapefiles and geographic coordinates were downloaded from the Census Bureau’s Topologically Integrated Geographic Encoding and Referencing (TIGER) system (Census, 2006b). This process identified a total of 264 census tracts for the census year 1990 and 346 for the year 2000. Social and economic variables were extracted from the census and calculated as percentages of the total tract population. Census population estimates for T1 and T2 were used as denominators for each census tract, and were multiplied by five because the CRC data represent five years of diagnosed cases. Rather than taking the average of CRC cases for all five years (the sum of these numbers divided by 5), the population for each census year was multiplied by five to coincide with the five years of
cancer data (i.e., simplification of a complex fraction). Table 4.3 lists the variables taken from the U.S. Census Bureau website.

<table>
<thead>
<tr>
<th>Census Variables</th>
<th>2000</th>
<th>1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Hispanic or Latino; white alone</td>
<td>P007003</td>
<td>P0120001</td>
</tr>
<tr>
<td>Not Hispanic or Latino; black or African American alone</td>
<td>P007004</td>
<td>P0120002</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>P007010</td>
<td>P0100001</td>
</tr>
<tr>
<td>Total population</td>
<td>P006001</td>
<td>P0010001</td>
</tr>
<tr>
<td>Income below poverty</td>
<td>P087002</td>
<td>P1170001-24</td>
</tr>
<tr>
<td>Education; male and female; up to 12th grade education no diploma</td>
<td>P037004-10 and P037021-27</td>
<td>P0570001 and P0570002</td>
</tr>
<tr>
<td>Population 16 and over; female and male; unemployed</td>
<td>P043014 and P043007</td>
<td>P0700007 and P0700003</td>
</tr>
<tr>
<td>Foreign-born</td>
<td>P021013</td>
<td>P0370004</td>
</tr>
<tr>
<td>Foreign-born; naturalized citizen</td>
<td>P021014</td>
<td>P0370005</td>
</tr>
<tr>
<td>Foreign-born; not a citizen</td>
<td>P021015</td>
<td>P0370006</td>
</tr>
</tbody>
</table>

### 4.2 Cluster Analysis

The first analytical goal of this study was to identify spatial clusters of higher than expected late stage CRC incidence and CRC-specific mortality cases in Miami-Dade County by census tract for T1 and T2 of those aged 50 and above. This analysis was performed using the SaTScan cluster detection software 7.0.3 (Kulldorff, 2007). Dr. Martin Kulldorff developed this method, which scans circular ‘windows’ of varying radii across a study area (Goodchild & Janelle, 2004; Kuldovert, 1997). For each ‘window,’ SaTScan estimates the probability that the frequency of observed events is higher or lower than that expected by chance, a process producing a Poisson distribution. The method estimates probabilities based on the likelihood ratio test statistic, which calculates the expected number of cases within and outside each window, and compares the null
hypothesis of constant risk to the alternative hypothesis of higher than expected risk (Waller & Gotway, 2004). The likelihood function, under the Poisson assumption, is equal to:

$$T_{\text{scan}} = \left( \frac{e^{c}}{E[c]} \right)^{c} \left( \frac{e^{c-E[c]}}{c-E[c]} \right)^{c-E[c]} I(c > E[c]),$$

Equation 4.1

where \( C \) is the total number of cases, \( c \) is the observed number of cases within the window, and \( E[c] \) is the covariate adjusted for expected number of cases within the window under the null hypothesis. The term \( C-E[c] \) is the expected number of cases outside the window. \( I(\cdot) \) denotes the indicator function. SaTScan was set to scan for geographic areas with high rates. Therefore, \( I(\cdot) \) was equal to 1 when the window had more cases than expected under the null-hypothesis, and \( I(\cdot) \) otherwise is set to zero (Kulldorff, 2006).

To determine the significance of the spatial scan probabilities, SaTScan generates a large number of replications of the data set through Monte Carlo simulations under the null hypothesis of a random distribution. The program then computes likelihood ratios for each simulation, and if the real data set is among the top 0.05 percent of all the values, it is statistically significant. The Monte Carlo technique also corrects for multiple comparisons.

This study employs a spatial analysis of CRC count data using the Poisson probability model to identify areas with high rates. Only late stage CRC incidence and CRC-specific mortality cases for those aged 50 years and over were included in this analysis. This age represents the typical cut-off for CRC studies because it is the point at
which CRC incidence begins to increase significantly with age (NCI, 2007a). This is exemplified in Figure 4.1. The age covariate was divided into eight separate age categories to adjust for heterogeneous age distributions across census tracts (see Table 4.4). A total of 9,999 Monte Carlo replications were generated to test for significance, and all window sizes were defined by the default in SaTScan.

**Figure 4.1: CRC Incidence by Age Category**

![CRC Incidence by Age Category](image)

**Table 4.4: Number of CRC Late Stage and Mortality Cases by Study Period and Age Category**

<table>
<thead>
<tr>
<th>Age Category</th>
<th>1988-1992 (T1)</th>
<th>1998-2002 (T2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Late Stage</td>
<td>CRC Death</td>
</tr>
<tr>
<td>50-54</td>
<td>166</td>
<td>122</td>
</tr>
<tr>
<td>55-59</td>
<td>250</td>
<td>168</td>
</tr>
<tr>
<td>60-64</td>
<td>383</td>
<td>262</td>
</tr>
<tr>
<td>65-69</td>
<td>467</td>
<td>309</td>
</tr>
<tr>
<td>70-74</td>
<td>552</td>
<td>370</td>
</tr>
<tr>
<td>75-79</td>
<td>535</td>
<td>372</td>
</tr>
<tr>
<td>80-84</td>
<td>406</td>
<td>266</td>
</tr>
<tr>
<td>85 and older</td>
<td>362</td>
<td>300</td>
</tr>
<tr>
<td>Total</td>
<td>3121</td>
<td>2169</td>
</tr>
</tbody>
</table>
The researcher then performed a spatial cluster analysis using the Poisson probability model, controlling for age and scanning for high rates of late stage CRC incidence and CRC-specific mortality for the two time periods. As mentioned previously, the Monte Carlo replications were set to 9,999. Each cluster analysis required three data files, all of which were defined at the tract level: late stage CRC incidence and/or CRC-specific mortality cases of those aged 50 and older; total population aged 50 and older; and census tract latitude and longitude coordinates (geographic centroids). The index key in each file consisted of the census tract number. Fields in the case and population files consisted of the census tract index key, age group, and number of cases or population for each census tract.

The study used the commercial software Statistical Package for the Social Sciences (SPSS) version 15.0 to create case and population files for both the 1988 to 1992 and 1998 to 2002 data files (SPSS, 2007). After aggregating cases by census tract index key for each time period, all CRC late stage and CRC-specific mortality cases of those diagnosed at age 50 or above were extracted from the FDCS data and saved to a “.dbf” file. The researcher applied the same method to the population files – the 1990 and 2000 files downloaded from the U.S. census – and saved these data sets to a separate “.dbf” file. To correspond to the five-year intervals of cancer diagnoses (1988-1992 and 1998-2002), the underlying 1990 and 2000 tract populations were multiplied by five to estimate the total number of people at risk for developing CRC over the five-year periods. The researcher identified tracts with no CRC cases or census populations. These census tracts were removed from the analyses as the log-population for these tracts is undefined.
The null hypotheses for the cluster analysis were that in T1 and T2, there were no spatial clusters of higher than expected late stage CRC incidence or CRC mortality cases. The alternate hypotheses were that after controlling for age and population size in T1, there were clusters of higher than expected cases of late stage CRC incidence (hypothesis A1) and CRC mortality (hypothesis A2), and that spatial clusters also existed in T2 for late stage CRC incidence (hypothesis B1) and CRC mortality cases (hypothesis B2).

4.3 Spatial Regression Analysis

The study applied a negative binomial regression analysis to test the relationship between CRC count data and social and economic variables at the tract level in Miami-Dade County for T1 and T2. Specifically, multiple regression analysis was employed to measure neighborhood effects on late stage cancer (an advanced stage cancer with poor survival prognosis) and on CRC mortality (the most severe and preventable cancer outcome). The social and economic variables included in the model were percent black, percent Hispanic, and percent below poverty. The variable percent black was included in the model due to national and regional trends that show higher CRC late stage incidence and mortality among blacks than any other racial or ethnic group (FCDS, 2005; NCI, 2008). As noted in the previous chapter, blacks also face particularly severe economic and social conditions in Miami (Nijman, 2000; Portes & Stepick, 1993). Percent Hispanic and percent below poverty were included to reflect those groups with higher proportions of people without health insurance (Duncan et al., 2005; Garvan et al., 2005); percent Hispanic also captures a significant proportion of the foreign-born population (Census, 2000b). Moreover, poverty variables are commonly used in area-based health
studies (Braveman et al., 2005; Krieger et al., 2005) as proxies for income, employment, and education. This dissertation uses these particular variables to describe conditions for which barriers to CRC prevention and treatment are most often found.

4.3.1 Dependent Variables

The regression analysis included two models for T1 and T2, with a dependent variable for each model. CRC cases diagnosed as regional or distant stage were combined to make up the first dependent variable, “late stage” CRC, and are comprised of counts. The second dependent variable is also comprised of counts and includes all diagnosed CRC cases in T1 and T2 that resulted in death due to CRC by 2005, as reported by the Florida DOH Office of Vital Statistics. The total number of late stage CRC incidence and CRC-specific mortality cases in T1 and T2 were calculated for all Miami-Dade census tracts using SPSS version 15.0 (SPSS, 2007). The study included in the analysis all diagnosed CRC cases, regardless of age, due to the unavailability of independent variables by age categories, such as non-Hispanic white and non-Hispanic black from the 1990 census.

4.3.2 Independent Variables

All independent variables originate from the 1990 and 2000 U.S. Census Bureau. The researcher downloaded tract-level measures of race, ethnicity, socioeconomic status, and age distribution from the U.S. Census Bureau website, and imported that data into a single relational database. To control for age, the study included three population age groups as separate covariates in the model: total persons aged 30 to 49 years, 50 to 79,
and 80 and older. The researcher calculated the percentages of persons in each age group for each tract to permit comparisons across tracts of different population sizes. The study evaluated multicollinearity using Pearson’s coefficients to assess the strength and direction of the linear relationship between each variable. The researcher did not include highly correlated variables with correlations greater than 0.8 in the dataset. Specifically, the variables identifying individuals with below a 12th grade education and percentage of foreign-born individuals were excluded due to strong correlations with percent poverty and percent Hispanic, respectively. The final set of population variables included percent black, percent Hispanic, percent below poverty, percent aged 30 to 49, percent aged 50 to 79, and percent aged 80 and older. The variance inflation factors (VIF) were generated for these variables using the commercial software Stata/IC (StataCorp, 2007) to further evaluate the impact of multicollinearity on the numerical estimates. The study design set a mean VIF of 10 as the cutoff point beyond which presence of multicollinearity was considered too strong. Correlation and VIF tables are presented in the next chapter.

4.3.3 Negative Binomial Regression: Models and Hypotheses

Because cancer data comprise counts, classical linear regression models are not appropriate, as count variables are discrete with a minimum value of zero (Beck, 1995; Land et al., 1996; Osgood, 2000). When using discrete data, such as counts, Poisson regression models often are applied. A Poisson random variable does not mimic the bell-shaped curve well where mu, or the mean, is less than several thousands, although it is asymptotically normal. The Poisson regression model uses a discrete probability distribution to analyze the likelihood that an event will occur based on a random and
independent distribution in time or space. However, simple Poisson regression has limitations because it assumes that the mean rate of occurrence is invariant in time or space, where the variance and the mean are equal to each other; a more sophisticated Poisson regression assumes the mean and variance, which are still equal, vary across time or space. In many cases, count variables display “overdispersion” where the mean and variance fail to be equal across a map, representing spatial or temporal “clumping” of events. In these instances, a negative binomial regression may be a better model, as it relaxes the mean-equal-to-variance constraint. The conditional likelihood of the negative binomial model is:

\[
    f(y_i | \mu_i) = \frac{(\nu_i \mu_i)^{y_i} e^{-\nu_i \mu_i}}{\Gamma(y_i + 1)},
\]

Equation 4.2

where \( y_i \) is the number of times something occurs, \( \nu_i \) is an observed parameter and \( (\nu_i \mu_i) \) is the Poisson distribution defined as:

\[
    \mu_i = \exp(x_i \beta + \text{offset}_i),
\]

Equation 4.3

where \( x_i \) is a covariate for the \( i \)-th observation. This implies that the dispersion for the \( i \)-th observation is:

\[
    \text{Var}(y_i) = \mu_i (1 + \alpha \mu_i),
\]

Equation 4.4
The Gamma \((1/\alpha, \alpha)\) density for \(v_x\) is:

\[
g(v) = \frac{v^{1/\alpha} e^{v/\alpha} e^{-v/\alpha}}{\alpha^1 \Gamma(1/\alpha)}, \tag{4.5}
\]

The gamma distribution has a mean of 1 and variance \(\alpha\), where \(\alpha\) is an ancillary parameter (Stata: Base Reference Manual, 2007).

This study employs spatial regression analysis using a negative binomial model. The study used the software package Stata/IC version 10.0 to conduct the regression analysis and diagnostics. GLM regression diagnostics were performed to evaluate statistical significance of overdispersion using the Likelihood-Ratio test with a probability cut-off value of 0.05. To control for spatial autocorrelation, or the phenomenon that a value observed at one location is not independent of the value of its neighboring locations, a spatially lagged variable was generated for all models using the commercial software SpaceStat version 1.93 (TerraSeer, 2001-2004). Essentially, this process spatially lags the dependent count variable by weighting the average of neighboring counts. If an observation for variable \(x\) at location \(i\) is represented by \(X_i\), then its spatial lag is:

\[
\text{Spatial Lag} = \sum_j W_{ij} X_j \tag{4.6}
\]

where the data set is a weighted average of the values in the locations neighboring each observation, defined as the sum of the product of each observation in the data set with a corresponding weight from the i-th row of the spatial weights matrix (Anselin, 1992). Spatial lags were created by the researcher by applying an inverse distance weight to the power of one to reflect declining influence of values with increasing distance. These weights were row standardized to facilitate interpretation and comparison across models.
(Anselin et al., 2000). The researcher included a spatial lag term as an autoregressive term in each regression model, represented as the spatially lagged dependent variable, to capture “neighbor effects” or the average CRC values for neighboring tracts.

The study then tested the relationship between CRC outcomes and neighborhood-level socioeconomic variables to assess the study hypotheses C1 to D2 posited below. The CRC variables late stage and mortality were entered separately as dependent variables in two multivariate negative binomial regressions using the independent variables percent black, percent Hispanic, percent below poverty, percent aged 30 to 49, percent aged 50 to 79, and percent aged 80 and older to test the study hypotheses presented in Chapter 1, and further detailed below:

**C1:** In T1, prevalence of late stage CRC incidence is positively associated with percent black, percent Hispanic, percent poverty, each age category, and the spatially lagged covariate.

**C2:** In T1, prevalence of CRC-specific mortality is positively associated with percent black, percent Hispanic, percent poverty, each age category, and the spatially lagged covariate.

**D1:** In T2, prevalence of late stage CRC incidence is positively associated with percent black, percent Hispanic, percent poverty, each age category, and the spatially lagged covariate.

**D2:** In T2, prevalence of CRC-specific mortality is positively associated with percent black, percent Hispanic, percent poverty, each age category, and the spatially lagged covariate.
4.3.4 Limitations of the Negative Binomial Model with the Spatial Lag Term

The introduction of a spatial lag term as a covariate, intended to account for spatial autocorrelation, places limitations on model inferences. The inclusion of a spatial lag as an independent variable creates a pseudo-likelihood estimate whereby standard errors may not be correct. Corrections to standard error estimates require calculation of a normalizing factor. In addition, a proper auto-negative binomial model specification can capture only negative spatial autocorrelation (Besag, 1974), making its applicability here of little use. Currently, there is no available software that applies this technique. Therefore, interpretations of standard errors should be made with caution.
CHAPTER V: RESULTS

5.0 Overview of Results

CRC incidence and health effects in Miami-Dade County show patterns of outcomes, some expected and others unexpected. Prior studies at the national level or in other communities have found that race, ethnicity, and SES are associated with CRC incidence, stage, and mortality. This study was designed to determine if these patterns exist in the racially and culturally diverse Miami-Dade County. The first part of this chapter summarizes the county-wide descriptive statistics of CRC patterns during T1 and T2 by age, race, and ethnicity. The next two sections present the results of the spatial cluster analysis and negative binomial regression analyses. The regression analysis found significant associations at the community level only between late stage CRC incidence and mortality for percent poverty during the 1988-1992 time period. Although the age-adjusted descriptive statistics suggest a relationship between rates of late stage CRC incidence and CRC-mortality with race and ethnicity, the regression analysis did not find significant associations between race or ethnicity and these CRC outcomes in Miami-Dade County during either time period. However, the cluster analysis did identify significant clusters in each time period of higher than expected late stage CRC incidence and CRC mortality, suggesting the need for public health intervention and outreach in these communities. The chapter concludes with an interpretation of the results.
5.1 Descriptive Statistics of CRC Patient Data

As the first step in the study, the researcher examined the FDCS data, looking at race, ethnicity, and gender differences in CRC incidence and outcomes to determine whether patterns described in the literature appear to exist in Miami-Dade County. Between 1990 and 2000, significant public health and political changes occurred that may have influenced CRC outcomes. In particular, during those years, medical organizations began recommending more aggressive screening guidelines for CRC. At the same time, Congress passed national welfare reform that altered eligibility for public health insurance or assistance programs, likely decreasing access to health care in general for low-income individuals and, in particular, cancer screenings for groups of people at higher risk. This study compared cancer statistics across two time periods, from 1988 to 1992 (T1) and from 1998 to 2002 (T2), to explore whether and why differences in CRC incidence or outcomes changed over time.

5.1.1 1988-1992 (T1) Cancer Cases By Race/Ethnicity and Gender in Miami-Dade County

CRC cancer cases were drawn from the FDCS database for Miami-Dade County for T1: 1988 to 1992. For those years, Miami-Dade County reported 6,129 diagnosed CRC cases, 48 percent of which occurred among men and 52 percent among women. The racial and ethnic classification of patients was 49 percent non-Hispanic white, 41 percent Hispanic, and 9 percent black. At the time of tumor diagnosis, the data identified 30 percent of CRCs as in situ or local stage, 39 percent as regional, and 15 percent as distant stage. The FDCS registry did not list the stage for approximately 16 percent of all cases. As of 2007, colorectal cancer-specific mortality occurred among 37 percent of all
CRC cases diagnosed in T1. Compared to blacks, Hispanics and non-Hispanic whites have higher proportions of local and regional stages of CRC, but lower proportions of distant stage cancer and CRC-specific mortality. Table 5.1 lists the CRC distributions by cancer stage, CRC-specific mortality, and race and ethnicity.

Table 5.1: CRC Stage Distribution by Race/Ethnicity and Gender, 1988-1992

<table>
<thead>
<tr>
<th></th>
<th>Total Cases</th>
<th>Local %</th>
<th>Regional %</th>
<th>Distant %</th>
<th>Mortality %</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Cases</td>
<td>6,129</td>
<td>29</td>
<td>39</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>Male</td>
<td>2,933</td>
<td>29</td>
<td>39</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>Female</td>
<td>3,196</td>
<td>29</td>
<td>40</td>
<td>15</td>
<td>36</td>
</tr>
<tr>
<td>NH White</td>
<td>3,013</td>
<td>29</td>
<td>37</td>
<td>14</td>
<td>34</td>
</tr>
<tr>
<td>NH Black</td>
<td>549</td>
<td>24</td>
<td>36</td>
<td>21</td>
<td>46</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2,523</td>
<td>30</td>
<td>41</td>
<td>14</td>
<td>39</td>
</tr>
</tbody>
</table>

Like most cancers, CRC rates increase with age. The rate of increase and mean age at onset, however, varies by race and ethnicity. The study restricted the categories analyzed by age for T1 to white, black, and Hispanic because the 1990 census did not provide age distributions for non-Hispanic white or non-Hispanic black racial categories that would permit calculation of rates by these classifications. Although the cancer statistics show little overlap between the Hispanic and black classifications among CRC cases (5%), 95 percent of Hispanic patients also are classified as white. Therefore, the CRC incidence rates presented in Figures 5.1 through 5.3 should be interpreted with caution because the cases classified as Hispanics include a high percentage of whites and a small percentage of blacks, and 55% of the population of white CRC patients are ethnically identified as Hispanic. For T1, Figure 5.1 shows age-specific CRC incidence
per 100,000 people by age group and race and ethnicity, Figure 5.2 shows age-specific late stage CRC incidence by age group, and Figure 5.3 shows age-specific rates or CRC-specific mortality rates by stage and race. These statistics are discussed in more detail below.

**Figure 5.1: CRC Incidence Rates by Age Group and Race/Ethnicity, 1988-1992**

**Figure 5.2: CRC Late Stage Rates by Age Group and Race/Ethnicity, 1988-1992**
In T1, age-specific incidence rates are similar across groups until age 55. At age 55, white incidence rates begin to exceed black rates, although black rates rise dramatically at age 80. Hispanic incidence rates are lower than those for blacks across most age categories. Late stage incidence rates by race and ethnicity are similar across age categories up to age 55, but for age groups 55 and above, black rates exceed both white and Hispanic rates, showing the largest margin after age 75. Mortality rates are higher generally among whites after age 65, but converge with black rates at age 85. Hispanic mortality rates are lower than black rates across all age categories except ages 70 to 74 and 80 to 84.

Age-adjusted rates were standardized for blacks, whites, and Hispanics to the 2000 U.S. Standard Million population, a weighted scale. Table 5.2 lists age-adjusted rates by stage and race and ethnicity for T1. Age-adjusted rates reported below are the number of cases per 100,000 people. The age-adjusted incidence for all racial and ethnic groups combined was 71.4. Whites had the highest incidence at 74, while blacks and Hispanics had rates of 66 and 59, respectively. Late stage incidence across groups was 38.5. The rate of late stage incidence among whites was 39, whereas black and Hispanic
rates were lower at 37 and 32, respectively. The CRC-specific mortality rate for all groups was 26.6. Blacks had a higher age-adjusted mortality rate at 31.6 than did whites at a rate of 26.9. Hispanics had the lowest mortality rate at 22.6.

Table 5.2: Age-Adjusted CRC Incidence, Late Stage, and Mortality Rates per 100,000 People by Race/Ethnicity, 1988-1992

<table>
<thead>
<tr>
<th></th>
<th>Incidence</th>
<th>Late Stage</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Cases</td>
<td>71.4</td>
<td>38.5</td>
<td>26.6</td>
</tr>
<tr>
<td>CI</td>
<td>69.6 - 73.2</td>
<td>37.2 - 39.8</td>
<td>25.5 - 27.6</td>
</tr>
<tr>
<td>White</td>
<td>74</td>
<td>39.0</td>
<td>26.9</td>
</tr>
<tr>
<td>CI</td>
<td>64 - 67.7</td>
<td>38.8 - 39.1</td>
<td>25.7 – 28</td>
</tr>
<tr>
<td>Black</td>
<td>66.0</td>
<td>36.9</td>
<td>31.6</td>
</tr>
<tr>
<td>CI</td>
<td>60.6 - 71.4</td>
<td>33 - 40.9</td>
<td>27.9 - 35.3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>58.8</td>
<td>32.0</td>
<td>22.6</td>
</tr>
<tr>
<td>CI</td>
<td>56.5 - 61</td>
<td>30.3 - 33.7</td>
<td>21.2 - 24</td>
</tr>
</tbody>
</table>

CI = 95% Confidence Interval

5.1.2 1998-2002 (T2) Cancer Cases By Race/Ethnicity and Gender in Miami-Dade County

The researcher determined CRC distributions across the same variables for T2: 1998-2002. In that time period, the cancer data identifies a total of 7,470 diagnosed CRC cases, 49 percent of which occurred in men and 51 percent in women. In terms of race/ethnicity, of the patients suffering these cases, thirty one percent were non-Hispanic white, 56 percent Hispanic, and 11 percent black. Table 5.3 lists the CRC distributions by cancer stage (though the FCDS did not report a stage for 17 percent of cases), CRC-specific mortality, race, and ethnicity.
As with T1, the researcher identified age-specific rates by race, ethnicity, incidence, late stage rate, and mortality. As noted above, T1 data did not allow the separation of white and black racial classifications from Hispanic ethnicity. In T2, due to reporting differences between the 1990 and 2000 census, the researcher was able to separate non-Hispanic whites from the Hispanic group. Specifically, the 2000 census data includes age distributions by Hispanic and non-Hispanic white, although not by non-Hispanic black. Thus, this section reports only non-Hispanic white rates. The researcher calculated the age-specific rates for blacks using the 2000 Miami-Dade county black population, but these rates include a small percentage of Hispanic black cases. The 2000 census data shows little overlap between black and Hispanic population groups in Miami-Dade County, with only 7 percent of the black population ethnically categorized as Hispanic. Thus, while the Hispanic rates reported cannot be interpreted as entirely distinct from black rates, there is little overlap. Figures 5.4 through 5.6 display age-specific rates per 100,000 people by race and ethnicity, CRC incidence, late stage CRC and CRC mortality for T2.
Figure 5.4: CRC Incidence Rates by Age Group and Race/Ethnicity, 1998-2002

Figure 5.5: CRC Late Stage Rates by Age Group and Race/Ethnicity, 1998-2002

Figure 5.6: CRC Mortality Rates by Age Group and Race/Ethnicity, 1998-2002
The graphs for T2 show several patterns. As in T1, incidence rates by age category begin to vary across race and ethnic group at around age 55, with non-Hispanic whites having the highest rates beyond that age. Incidence rates for blacks are lower than for whites across all age categories, with the exception of a spike at age 75 when the two converge, followed by a drop in rates for blacks. Hispanic rates fall between black and white rates until age 75, when their rates fall below black rates. A similar pattern is observed for late stage incidence rates. After age 50, CRC mortality rates fluctuate across race and ethnicity. Whites have higher mortality rates at age 60 to 64, 70 to 74, and over 80, while blacks have higher rates at ages 65 to 69 and 80 to 84. Hispanic mortality rates fall between black and white rates between the ages of 75 and 84 and are similar to black rates thereafter.

Age-adjusted rates again were standardized for blacks, non-Hispanic whites, and Hispanics to the U.S. Standard Million population in the same manner described in Section 5.1.1. Table 5.4 lists age-adjusted rates by stage and race/ethnicity. Again, all age-adjusted rates are cases per 100,000 people. The age-adjusted incidence for all groups combined was 75.6. Non-Hispanic whites had the highest incidence at 87.9, blacks had the second highest at 70.7, and Hispanics were lowest at 69.8. Late stage CRC rates were more similar across groups than incidence rates, with identical rates among blacks and Hispanics (35.8 and 35.7). The white late stage incidence rate of 46.4, however, lead black and Hispanic rates by nearly 11 cases per 100,000. CRC mortality rates for non-Hispanic whites and blacks are the same at rates of 23, but lower among Hispanics at a rate of 18.4. Compared to T1, Hispanics had the lowest drop in age-standardized, CRC-specific mortality rates of any group. However, the stability of
subpopulation sample sizes and their respective age-adjusted rates makes it difficult to interpret changes over time.

### Table 5.4: Age-Adjusted CRC Incidence, Late Stage, and Mortality Rates per 100,000 People by Race/Ethnicity, 1998-2002

<table>
<thead>
<tr>
<th></th>
<th>Incidence</th>
<th>Late Stage</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Cases</strong></td>
<td>75.6</td>
<td>38.7</td>
<td>20.4</td>
</tr>
<tr>
<td>CI</td>
<td>73.8 - 77.3</td>
<td>37.5 - 40</td>
<td>19.6 - 21.4</td>
</tr>
<tr>
<td><strong>NH-White</strong></td>
<td>87.9</td>
<td>46.4</td>
<td>23.0</td>
</tr>
<tr>
<td>CI</td>
<td>85.7 - 90.1</td>
<td>44.8 - 48</td>
<td>22.2 - 24.4</td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td>70.7</td>
<td>35.8</td>
<td>23.0</td>
</tr>
<tr>
<td>CI</td>
<td>66.1 - 75.4</td>
<td>29.1 - 32.6</td>
<td>20.3 - 25.7</td>
</tr>
<tr>
<td><strong>Hispanic</strong></td>
<td>69.8</td>
<td>35.7</td>
<td>18.4</td>
</tr>
<tr>
<td>CI</td>
<td>67.7 - 72</td>
<td>34.1 - 37.2</td>
<td>17.3 - 19.5</td>
</tr>
</tbody>
</table>

CI = 95% Confidence Interval

### 5.1.3 Comparison of Incidence, Late Stage, and Mortality from T1 to T2

The statistics show both similarities between and differences in the proportion of cases and age-adjusted rates by race and ethnicity between T1 and T2. The percentage of the total CRC cases occurring among Hispanics increased by 15 percent. The stage distribution at T2 differs only slightly from T1. The proportion of cases that resulted in CRC-specific mortality as of 2007 was much smaller at T2 than T1 (29% vs. 37%), an expected result given that the time lapse from T2 is much shorter than from T1. Site distributions also varied over time. By T2, Hispanic and black patients had higher proportions of local stage CRC than non-Hispanic whites, but had lower proportions of regional stage CRC. However, non-Hispanic whites had higher age-adjusted rates for
CRC incidence and late stage CRC incidence in comparison with black and Hispanic rates. Late stage CRC incidence and CRC-specific mortality exhibited a higher proportion of cases in blacks than in either non-Hispanic whites or Hispanics, but age-adjusted mortality rates were roughly equal across groups by T2, though Hispanic mortality rates were still lower.

5.2 Results of the Spatial Clustering Analysis for T1 and T2

To determine any significant geographical variations in CRC outcomes, a spatial analysis was conducted to detect higher than expected clusters of late-stage CRC incidence and CRC-specific mortality in Miami-Dade County for the years 1988 to 1992 (T1) and 1998 to 2002 (T2) using the software program SaTScan version 7.0.3 (Kuldorff, 2007). Relative risk probabilities for each census tract were generated using the Monte Carlo technique. The study included all cases diagnosed at age 50 or older in this analysis. Eight age groups – each consisting of five-year intervals starting at age 50, with the exception of the eighth age group which includes ages 85 and older – were used as covariates to control for differences in age distributions across tracts. Pearson’s bivariate correlations of age categories show high levels of multicollinearity, between 0.92 and 0.97, for consecutive age categories. The cluster analysis reported statistically significant clusters of higher than expected outcomes (“significant clusters”) for late stage CRC incidence and CRC mortality cases in each time period. The analysis identified other clusters of higher than expected cases that, while not statistically significant, may nonetheless be relevant for public health surveillance efforts—they are therefore included in the maps and tables in this section.
5.2.1 Study Hypothesis A1: *In T1, after controlling for age and population size, there are clusters of higher than expected cases of late stage CRC incidence.*

The cluster analysis of T1 late stage CRC incidence using the Poisson model identified 48 tracts where the incidence of late stage CRC cases was higher than expected. The results showed a total of four clusters that had statistically significantly higher than expected numbers of cases (Clusters 1-4) (P < 0.05), showing positive spatial autocorrelation, in addition to eleven clusters that displayed higher than expected late stage cases that were not statistically significant (Clusters 5-15). Table 5.5 lists the cluster analysis output of expected and observed cases, relative risk probabilities, tract composition, and significance values by cluster. The statistically significant clusters are all located in the outer edges of Miami-Dade county in: Doral and an unincorporated tract in western Miami-Dade (Cluster 1); Miami Beach and Fisher Island in eastern Miami-Dade (Cluster 2); Andover, Ives Estates, and Norland in northern Miami-Dade (Cluster 3); and in the Hammocks located in southwest Miami-Dade county (Cluster 4). Figure 5.7 maps the spatial distribution of these tracts. The results support the hypothesis that there are clusters of higher than expected late-stage CRC in Miami-Dade County during T1, identifying Clusters 1, 2, 3, and 4.
Table 5.5: Cluster Analysis Results for Late Stage CRC Incidence Cases, Age 50 and Older, 1988-1992

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Tract Number</th>
<th>Significance value</th>
<th>Cluster Observed Cases</th>
<th>Cluster Expected Cases</th>
<th>Cluster Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>9003, 10103</td>
<td>.00001</td>
<td>40</td>
<td>11.27</td>
<td>3.584</td>
</tr>
<tr>
<td>2*</td>
<td>4500</td>
<td>.00004</td>
<td>34</td>
<td>10.67</td>
<td>3.210</td>
</tr>
<tr>
<td>3*</td>
<td>9600, 9800, 9901</td>
<td>.00378</td>
<td>84</td>
<td>48.11</td>
<td>1.767</td>
</tr>
<tr>
<td>4*</td>
<td>10125</td>
<td>.03580</td>
<td>30</td>
<td>12.66</td>
<td>2.383</td>
</tr>
<tr>
<td>5</td>
<td>205</td>
<td>.06499</td>
<td>21</td>
<td>7.65</td>
<td>2.758</td>
</tr>
<tr>
<td>6</td>
<td>1501</td>
<td>.39162</td>
<td>14</td>
<td>4.91</td>
<td>2.861</td>
</tr>
<tr>
<td>7</td>
<td>103, 108</td>
<td>.55043</td>
<td>55</td>
<td>34.78</td>
<td>1.592</td>
</tr>
<tr>
<td>8</td>
<td>8201, 8203</td>
<td>.75769</td>
<td>29</td>
<td>15.85</td>
<td>1.838</td>
</tr>
<tr>
<td>9</td>
<td>7604</td>
<td>.79873</td>
<td>15</td>
<td>6.35</td>
<td>2.370</td>
</tr>
<tr>
<td>10</td>
<td>10201</td>
<td>.98847</td>
<td>6</td>
<td>1.76</td>
<td>3.414</td>
</tr>
<tr>
<td>11</td>
<td>6102, 6200, 6302, 6800, 6900, 7001, 7002, 7100, 7200, 7300, 7400, 7501, 7502, 7503, 7901, 1803</td>
<td>.99154</td>
<td>175</td>
<td>145.08</td>
<td>1.218</td>
</tr>
<tr>
<td>12</td>
<td>408, 1001, 1101, 1102, 1103, 1104, 1402</td>
<td>.99935</td>
<td>12</td>
<td>5.75</td>
<td>2.092</td>
</tr>
<tr>
<td>13</td>
<td>408, 1001, 1101, 1102, 1103, 1104, 1402</td>
<td>.99981</td>
<td>57</td>
<td>42.10</td>
<td>1.360</td>
</tr>
<tr>
<td>14</td>
<td>10701</td>
<td>.99982</td>
<td>14</td>
<td>7.33</td>
<td>1.914</td>
</tr>
<tr>
<td>15</td>
<td>4901, 4902, 5001, 5002, 5100, 5401, 5501, 5502</td>
<td>.99997</td>
<td>139</td>
<td>115.99</td>
<td>1.208</td>
</tr>
</tbody>
</table>

* Statistically significant clusters at the 95% Confidence Interval
Figure 5.7: Map of Clusters of Higher than Expected Late-Stage CRC Cases in Miami-Dade, 1988-1992
5.2.2 Study Hypothesis A2: In T1, after controlling for age and population size, there are clusters of higher than expected cases of CRC-specific mortality.

SaTScan also was used to analyze clusters of higher than expected CRC deaths. In T1, the analysis identified a total of 37 tracts that had higher than expected CRC deaths compared to the rest of Miami-Dade county tracts showing positive spatial autocorrelation. Table 5.6 lists the cluster analysis output of expected and observed cases, relative risk probabilities, tract composition, and significance values by cluster. A total of four out of fifteen clusters with higher than expected cases, illustrated in Figure 5.8, showed significantly higher than expected CRC cases in comparison with all other Miami-Dade County tracts (P < 0.05). The first statistically significant cluster includes the communities of West Little River, Gladeview, Brownsville, and Liberty City; Cluster 2 is comprised of one unincorporated tract in western Miami-Dade County adjacent to Doral; the third and fourth clusters are located in part of the Golden Glades (Cluster 3) and in North Miami (Cluster 4). The results support the hypothesis that there are clusters of higher than expected cases of CRC-specific mortality in Miami-Dade County during T1.
Table 5.6: Cluster Analysis Results for CRC-Specific Mortality, Age 50 and Older, 1988-1992

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Tract Number</th>
<th>Significance Value</th>
<th>Cluster Observed Cases</th>
<th>Cluster Expected Cases</th>
<th>Cluster Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>1502, 1801, 1501, 1004, 1903, 903, 1003, 1901, 1002, 1803, 1802, 1904, 1702, 1701, 902</td>
<td>0.00138</td>
<td>99</td>
<td>58.25</td>
<td>1.733</td>
</tr>
<tr>
<td>2*</td>
<td>10103</td>
<td>0.02856</td>
<td>16</td>
<td>4.5</td>
<td>3.574</td>
</tr>
<tr>
<td>3*</td>
<td>103</td>
<td>0.03572</td>
<td>33</td>
<td>14.61</td>
<td>2.279</td>
</tr>
<tr>
<td>4*</td>
<td>205</td>
<td>0.04392</td>
<td>17</td>
<td>5.22</td>
<td>3.275</td>
</tr>
<tr>
<td>5</td>
<td>2202</td>
<td>0.08382</td>
<td>16</td>
<td>5.03</td>
<td>3.195</td>
</tr>
<tr>
<td>6</td>
<td>10201</td>
<td>0.17895</td>
<td>7</td>
<td>1.18</td>
<td>5.939</td>
</tr>
<tr>
<td>7</td>
<td>9600</td>
<td>0.21728</td>
<td>20</td>
<td>7.97</td>
<td>2.522</td>
</tr>
<tr>
<td>8</td>
<td>3702</td>
<td>0.57984</td>
<td>4</td>
<td>0.48</td>
<td>8.279</td>
</tr>
<tr>
<td>9</td>
<td>5802</td>
<td>0.70507</td>
<td>23</td>
<td>11.43</td>
<td>2.023</td>
</tr>
<tr>
<td>10</td>
<td>8203, 8201, 8204</td>
<td>0.72452</td>
<td>30</td>
<td>16.5</td>
<td>1.83</td>
</tr>
<tr>
<td>11</td>
<td>7604</td>
<td>0.73439</td>
<td>12</td>
<td>4.4</td>
<td>2.735</td>
</tr>
<tr>
<td>12</td>
<td>10125</td>
<td>0.90007</td>
<td>18</td>
<td>8.68</td>
<td>2.082</td>
</tr>
<tr>
<td>13</td>
<td>9003, 9100</td>
<td>0.98871</td>
<td>20</td>
<td>10.86</td>
<td>1.849</td>
</tr>
<tr>
<td>14</td>
<td>5002, 5501, 5001, 4901, 5502, 4902</td>
<td>0.99114</td>
<td>72</td>
<td>53.33</td>
<td>1.362</td>
</tr>
<tr>
<td>15</td>
<td>4500</td>
<td>0.99685</td>
<td>15</td>
<td>7.59</td>
<td>1.983</td>
</tr>
</tbody>
</table>

* Statistically significant clusters at the 95% Confidence Interval
Figure 5.8: Map of Clusters of Higher than Expected CRC Mortality Cases in Miami-Dade, 1988-1992
5.2.3 Study Hypothesis B1: In T2, after controlling for age and population size, there are clusters of higher than expected cases of late stage CRC incidence.

In T2, the cluster analysis of late stage CRC incidence cases produced 11 clusters of a total of 108 tracts that had higher than expected cases showing positive spatial autocorrelation. Table 5.7 lists the SaTScan output of expected and observed cases, relative risk probabilities, tract composition, and significance values by cluster. One out of the eleven clusters with higher than expected rates of late-stage CRC was statistically significantly different from expected cases in the remaining Miami-Dade county tracts. This cluster was south of downtown Miami in the area designated by the census as “Coral Way.” These spatial patterns are illustrated in Figure 5.9, with the statistically significant cluster displayed in the inset map. The results support the hypothesis because there is a statistically significant cluster of late-stage CRC in Miami-Dade County during T2.
Table 5.7: Cluster Analysis Results for Late Stage CRC Incidence Cases, Age 50 and Older, 1998-2002

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Tract Number</th>
<th>Significance Value</th>
<th>Cluster Observed Cases</th>
<th>Cluster Expected Cases</th>
<th>Cluster Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>6702</td>
<td>0.01700</td>
<td>19</td>
<td>5.65</td>
<td>3.376</td>
</tr>
<tr>
<td>2</td>
<td>401, 302, 403, 405, 402, 301, 9502, 304, 305, 404, 9501, 406, 204, 306, 205, 9400, 210, 503, 9902, 1101, 407, 9903, 1102, 9600, 208, 203, 501, 408, 10006, 502, 10002, 209, 206, 1103, 1203, 202, 9901, 1005, 10010, 9802, 901, 1104, 9904, 1202, 9702, 1002, 110, 601, 1204, 1003, 1006, 10005, 10009, 201, 902, 604, 9801, 602, 109, 1004, 10001, 1402, 10001, 10007, 1301, 603, 1401, 106, 605, 9307, 9701, 9308, 9200</td>
<td>0.52853</td>
<td>697</td>
<td>621.43</td>
<td>1.153</td>
</tr>
<tr>
<td>3</td>
<td>1903, 1904, 1501, 1901, 1801, 1802, 2300</td>
<td>0.74992</td>
<td>56</td>
<td>36.11</td>
<td>1.560</td>
</tr>
<tr>
<td>4</td>
<td>8204, 10604, 8203, 8306, 8303, 10606, 10605, 8305, 8307, 8201, 10206, 10607, 8304, 8409, 10205, 8100, 8415, 10114,</td>
<td>0.90622</td>
<td>200</td>
<td>162.96</td>
<td>1.241</td>
</tr>
<tr>
<td>5</td>
<td>3702, 3701, 3400</td>
<td>0.96872</td>
<td>13</td>
<td>5.49</td>
<td>2.373</td>
</tr>
<tr>
<td>6</td>
<td>10124</td>
<td>0.97461</td>
<td>9</td>
<td>3.13</td>
<td>2.882</td>
</tr>
<tr>
<td>7</td>
<td>8412</td>
<td>0.99080</td>
<td>17</td>
<td>8.41</td>
<td>2.025</td>
</tr>
<tr>
<td>8</td>
<td>10201, 10172</td>
<td>0.99974</td>
<td>13</td>
<td>6.27</td>
<td>2.077</td>
</tr>
<tr>
<td>9</td>
<td>10141</td>
<td>0.99978</td>
<td>14</td>
<td>6.98</td>
<td>2.010</td>
</tr>
<tr>
<td>10</td>
<td>4602</td>
<td>0.99999</td>
<td>10</td>
<td>4.64</td>
<td>2.158</td>
</tr>
<tr>
<td>11</td>
<td>9100</td>
<td>0.99999</td>
<td>17</td>
<td>9.97</td>
<td>1.709</td>
</tr>
</tbody>
</table>

* Statistically significant clusters at the 95% Confidence Interval
Figure 5.9: Map of Clusters of Higher than Expected Late-Stage CRC Cases in Miami-Dade, 1998-2002
5.2.4 Study Hypothesis B2: In T2, after controlling for age and population size, there are clusters of higher than expected cases of CRC-specific mortality.

The analysis of higher than expected CRC-specific mortality cases in T2 generated a total of eight clusters consisting of 101 tracts showing positive spatial autocorrelation. Table 5.8 lists the cluster output of expected and observed cases, relative risk probabilities, tract composition, and significance values by cluster. Of the eight clusters with higher than expected cases of CRC-specific mortality, one cluster, consisting of 21 tracts, was statistically significantly different from expected cases in the remaining tracts. These tracts are spread among the census designated places of Liberty City, Little Haiti, Allapattah, Brownsville, Gladeview, and West Little River. Figure 5.10 displays the spatial distribution of clusters with higher than expected rates, with the statistically significant cluster displayed in the inset map. The results support the hypothesis because there is a statistically significant cluster of CRC-specific mortality in Miami-Dade County during T2.
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Tract Number</th>
<th>Significance Value</th>
<th>Cluster Observed Cases</th>
<th>Cluster Expected Cases</th>
<th>Cluster Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10006, 10010, 501, 9400, 10005, 10009, 10002, 10007, 10145, 402, 10143, 502, 10001, 9903, 9502, 403, 9304, 10144, 9904, 9305, 10142, 9307, 401, 503, 9902, 9501, 404, 9308, 601, 9901, 10132, 602, 9306, 405, 10010, 9200, 302, 301, 204, 9311, 9600, 10134, 9309, 406, 603, 604, 407, 901, 10136, 304, 305</td>
<td>0.13072</td>
<td>279</td>
<td>223.60</td>
<td>1.293</td>
</tr>
<tr>
<td>3</td>
<td>10155</td>
<td>0.25092</td>
<td>14</td>
<td>4.46</td>
<td>3.157</td>
</tr>
<tr>
<td>4</td>
<td>8306, 8307, 8303, 8305, 8304, 8204, 8203, 10604, 8415, 10206, 10606, 10178, 8409, 10114, 8414, 10205, 8201, 10607, 10177, 10204, 10605, 8100, 8405</td>
<td>0.58217</td>
<td>135</td>
<td>101.79</td>
<td>1.352</td>
</tr>
<tr>
<td>5</td>
<td>6800</td>
<td>0.70016</td>
<td>17</td>
<td>7.17</td>
<td>2.383</td>
</tr>
<tr>
<td>6</td>
<td>803, 804, 801, 70</td>
<td>0.88321</td>
<td>53</td>
<td>34.76</td>
<td>1.540</td>
</tr>
<tr>
<td>7</td>
<td>8412</td>
<td>0.91252</td>
<td>12</td>
<td>4.64</td>
<td>2.598</td>
</tr>
<tr>
<td>8</td>
<td>201</td>
<td>0.97336</td>
<td>10</td>
<td>3.72</td>
<td>2.694</td>
</tr>
</tbody>
</table>

* Statistically significant clusters at the 95% Confidence Interval
Figure 5.10: Map of Clusters of Higher than Expected CRC Mortality Cases in Miami-Dade, 1998-2002
5.2.5 Descriptive Statistics for Significant CRC Clusters

The neighborhood social and economic characteristics were reviewed for each statistically significant CRC cluster in T1 and T2 to identify persistent patterns across tracts with higher than expected late stage CRC incidence and CRC mortality. Table 5.9 summarizes community-level racial, ethnic, citizenship, and nativity characteristics for the T1 and T2 clusters.

Table 5.9: Selected Population Characteristics of Statistically Significant Clusters

<table>
<thead>
<tr>
<th></th>
<th>% Black</th>
<th>% Hispanic</th>
<th>% Foreign-born</th>
<th>% Non-Citizen</th>
<th>% Below Poverty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T1 Late Stage CRC Incidence Clusters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 1 (Doral and adjacent unincorporated tract)</td>
<td>11</td>
<td>61</td>
<td>49</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>Cluster 2 (Southern Point of South Beach)</td>
<td>3</td>
<td>52</td>
<td>52</td>
<td>30</td>
<td>41</td>
</tr>
<tr>
<td>Cluster 3 (Andover, Ives Estates, Norland)</td>
<td>23</td>
<td>16</td>
<td>27</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Cluster 4 (The Hammocks)</td>
<td>6</td>
<td>50</td>
<td>40</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td><strong>T1 CRC Mortality Clusters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 1 (Liberty City, Brownsville, Gladeview, West Little River)</td>
<td>86</td>
<td>11</td>
<td>14</td>
<td>8</td>
<td>41</td>
</tr>
<tr>
<td>Cluster 2 (unincorporated tract near Doral)</td>
<td>14</td>
<td>71</td>
<td>52</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>Cluster 3 (Golden Glades)</td>
<td>9</td>
<td>17</td>
<td>30</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Cluster 4 (North Miami)</td>
<td>37</td>
<td>23</td>
<td>40</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td><strong>T2 Late Stage CRC Incidence Cluster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 1 (Coral Way)</td>
<td>1</td>
<td>64</td>
<td>54</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td><strong>T2 CRC-Mortality Cluster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 1 (Liberty City, Little Haiti, Allapattah, Brownsville, Gladeview, West Little River)</td>
<td>71</td>
<td>21</td>
<td>27</td>
<td>16</td>
<td>40</td>
</tr>
</tbody>
</table>
In T1, the late stage Clusters 1, 2 and 4 include the communities Doral and the Hammocks in western Dade-County, and the most southern tip of South Beach. At that time, these communities had Hispanic majorities, 50 percent foreign-born residents, and roughly a quarter foreign-born non-citizens. The percentage of people below poverty was low in the most western communities (6% to 7%), but high in South Beach (40%). Cluster 3, consisting of the North Miami-Dade County communities of Andover, Ives Estates and Norland, exhibited a moderate percentage of black residents (23%), a relatively small percentage of Hispanic residents (16%) a moderate foreign-born population (27%), and a relatively low below-poverty rate (9%). Mortality clusters in T1 are located in the Liberty City area (Cluster 1), the community of Doral (Cluster 2), the Golden Glades in northeast Miami-Dade (Cluster 3), and North Miami (Cluster 4). At that time, the population of Cluster 1, which included Liberty City, was 86 percent blacks, 41 percent below poverty, 11 percent Hispanic, and 14 percent foreign-born. The North Miami cluster demographics was a mix of Hispanic (23%), black (37%), and foreign-born residents (40%), with a relatively low proportion of people below poverty (13%). The Golden Glades cluster had very few Hispanic residents (17%), black residents (9%), and residents below poverty (13%), but a moderate percentage of foreign-born residents (30%).

Although the cluster analysis for T2 generated just one late stage cluster and one for mortality, these clusters also were characterized by high percentages of minority populations. The late stage cluster – the “Coral Way” tract in Downtown Miami – consisted of 64 percent Hispanic residents, 54 percent foreign-born residents, 27 percent foreign-born non-citizen residents, and 16 percent of residents below poverty. The CRC
mortality cluster straddles the predominantly non-Hispanic black communities of Liberty City, Little Haiti, Allapattah, Brownsville, Gladeview, and West Little River. This cluster is characterized by a high proportion of black residents (71%), fewer Hispanics (21%), a modest foreign-born population (27%), and a high percentage of individuals living below poverty (40%). Thus, the late stage cluster is a Hispanic, foreign-born, and low poverty population, while the mortality cluster is the opposite, with a large black and below poverty population.

5.3 Negative Binomial Spatial Regression Analysis

In the next stage of the study, the researcher estimated a negative binomial regression model to measure late stage CRC incidence and CRC-specific mortality count data in Miami-Dade county in T1 and T2. This dissertation seeks to measure the social and economic relationships at the neighborhood level to CRC outcomes while controlling for differences in age, population size, and the spatial distribution of CRC. While an increasing number of public health studies evaluate the association of area-based characteristics with cancer outcomes, many such studies are limited in their analysis of neighborhood factors because they employ rates that are only stable at larger geographical units, such as the county level (Hsu et al., 2006; Lai et al., 2006; Vinnakota & Lam, 2006). As discussed in Chapter 4, the negative binomial regression model permits the analysis of count level data displaying overdispersion.

For T1 and T2, the descriptive statistics from the regression model show similar mean percentages for most model variables: 19 to 20 percent below poverty; 22 to 23 percent non-Hispanic black; 28 to 30 percent age 30 to 49; 25 to 26 percent age 50 to 79; and 4 percent age 80 and older. Average percentage of Hispanics was the only variable
that showed a major change, increasing from 45 percent in T1 to 55 percent in T2. The average number of late stage cases per tract was 12.73 in T1 and 10.89 in T2, and the average number of mortality cases was 8.64 in T1 and 5.73 in T2. Tables 5.10 and 5.12 summarize the ranges, means, and standard deviations of the dependent and independent variables. The bivariate results indicate significant correlations between percent Hispanic and percent non-Hispanic black, as well as between ages 50 to 79 and ages 80 and older. The population variable used in the model (“population log”) was transformed using the natural logarithm of the total population for each tract. Table 5.11 and 5.13 present Pearson’s correlations used in the analyses for T1 and T2. While researchers use Pearson’s correlation values to indicate presence of multicollinearity among model variables, it is typically used in combination with other methods.

In addition to generating a Pearson’s correlation matrix to evaluate multicollinearity, the researcher produced VIFs for each model to evaluate the impact of multicollinearity on the quality of the numerical estimates. Independent variables with VIFs greater to or equal to 10 were considered to have standard error estimates that are strongly impacted by multicollinearity. For this dissertation, factors less than four were considered to have a weak impact on model estimates due to multicollinearity, and VIFs of four to ten were considered to represent moderate impacts on estimates. As discussed in the next two sections detailing the results of the regression analysis, no variables showed a sufficiently high impact, or VIFs over 10, to be excluded from the regression analyses.
Table 5.10: Descriptive Statistics of T1 Model Variables (n=264 tracts)

<table>
<thead>
<tr>
<th>Model Variables</th>
<th>Low</th>
<th>High</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRC mortality cases</td>
<td>.00</td>
<td>46.00</td>
<td>8.64</td>
<td>7.23</td>
</tr>
<tr>
<td>Late stage CRC incidence cases</td>
<td>.00</td>
<td>70.00</td>
<td>12.54</td>
<td>11.00</td>
</tr>
<tr>
<td>Percent non-Hispanic black</td>
<td>.00</td>
<td>0.99</td>
<td>0.23</td>
<td>0.32</td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>.01</td>
<td>0.96</td>
<td>0.45</td>
<td>0.30</td>
</tr>
<tr>
<td>Percent below poverty</td>
<td>.01</td>
<td>0.68</td>
<td>0.20</td>
<td>0.14</td>
</tr>
<tr>
<td>Percent age 30-49</td>
<td>.05</td>
<td>0.51</td>
<td>0.28</td>
<td>0.06</td>
</tr>
<tr>
<td>Percent age 50-79</td>
<td>.00</td>
<td>0.56</td>
<td>0.26</td>
<td>0.09</td>
</tr>
<tr>
<td>Percent age 80 and older</td>
<td>.00</td>
<td>0.36</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Population x 5</td>
<td>560</td>
<td>148,840</td>
<td>36,684</td>
<td>23,811</td>
</tr>
</tbody>
</table>

Table 5.11: Zero Order Correlations, T1 CRC Cases with 1990 Census Variables

<table>
<thead>
<tr>
<th></th>
<th>CRC Mortality</th>
<th>Late Stage CRC Incidence</th>
<th>Non-Hisp. Black %</th>
<th>Hisp. %</th>
<th>Below Poverty %</th>
<th>Pop Log</th>
<th>Age 30-49 %</th>
<th>Age 50-79 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRC Mortality</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRC Incidence</td>
<td>0.911**</td>
<td>-0.284**</td>
<td>-0.325**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hisp. Black %</td>
<td>-0.284**</td>
<td>-0.325**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hisp. %</td>
<td>0.211**</td>
<td>0.199**</td>
<td>-0.644**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below Poverty %</td>
<td>-0.117</td>
<td>-0.112</td>
<td>0.572**</td>
<td>-0.088</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PopLog</td>
<td>0.494**</td>
<td>0.470**</td>
<td>-0.145*</td>
<td>0.171**</td>
<td>-0.0281*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 30-49 %</td>
<td>-0.188**</td>
<td>-0.196**</td>
<td>-0.079</td>
<td>-0.180**</td>
<td>-0.328**</td>
<td>0.077</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Age 50-79 %</td>
<td>0.383**</td>
<td>0.413**</td>
<td>-0.485**</td>
<td>0.430**</td>
<td>-0.190**</td>
<td>-0.079</td>
<td>-0.0469**</td>
<td>1</td>
</tr>
<tr>
<td>Age 80+ %</td>
<td>0.390**</td>
<td>0.461**</td>
<td>-0.319**</td>
<td>0.076</td>
<td>-0.029</td>
<td>-0.092</td>
<td>-0.460**</td>
<td>0.660**</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Table 5.12: Descriptive Statistics of T2 Model Variables (n=345 tracts)

<table>
<thead>
<tr>
<th>Model Variables</th>
<th>Low</th>
<th>High</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRC mortality cases</td>
<td>.00</td>
<td>21.00</td>
<td>5.73</td>
<td>3.92</td>
</tr>
<tr>
<td>Late stage CRC incidence cases</td>
<td>.00</td>
<td>36.00</td>
<td>10.89</td>
<td>6.35</td>
</tr>
<tr>
<td>Percent non-Hispanic black</td>
<td>.00</td>
<td>0.97</td>
<td>0.21</td>
<td>0.30</td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>.01</td>
<td>0.95</td>
<td>0.55</td>
<td>0.13</td>
</tr>
<tr>
<td>Percent below poverty</td>
<td>.00</td>
<td>0.65</td>
<td>0.19</td>
<td>0.13</td>
</tr>
<tr>
<td>Percent age 30-49</td>
<td>.01</td>
<td>0.52</td>
<td>0.30</td>
<td>0.05</td>
</tr>
<tr>
<td>Percent age 50-79</td>
<td>.00</td>
<td>0.45</td>
<td>0.25</td>
<td>0.07</td>
</tr>
<tr>
<td>Percent age 80 and older</td>
<td>.00</td>
<td>0.19</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Population x 5</td>
<td>750</td>
<td>92,735</td>
<td>32,657</td>
<td>13,897</td>
</tr>
</tbody>
</table>
Table 5.13: Zero Order Correlations, T2 CRC Cases with 2000 Census Variables

<table>
<thead>
<tr>
<th></th>
<th>Late Stage CRC Mortality</th>
<th>Late Stage CRC Incidence</th>
<th>Non-Hisp. Black %</th>
<th>Hisp. %</th>
<th>Below Poverty %</th>
<th>PopLog</th>
<th>Age 30-49 %</th>
<th>Age 50-79 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRC Mortality</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRC Incidence</td>
<td>0.803**</td>
<td>-0.299**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hisp. Black %</td>
<td>-0.122**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hisp. %</td>
<td>0.165**</td>
<td>0.263**</td>
<td>-0.771**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below Poverty %</td>
<td>-0.078</td>
<td>-0.197**</td>
<td>0.574**</td>
<td>-0.244**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PopLog</td>
<td>0.446**</td>
<td>0.525**</td>
<td>-0.173**</td>
<td>0.226**</td>
<td>-0.266**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 30-49 %</td>
<td>-0.119**</td>
<td>-0.052</td>
<td>-0.309**</td>
<td>0.127**</td>
<td>-0.309**</td>
<td>0.145**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Age 50-79 %</td>
<td>0.350**</td>
<td>0.469**</td>
<td>-0.437**</td>
<td>0.360**</td>
<td>-0.251**</td>
<td>-0.025**</td>
<td>-0.180**</td>
<td>1</td>
</tr>
<tr>
<td>Age 80+ %</td>
<td>0.348**</td>
<td>0.451**</td>
<td>-0.326**</td>
<td>0.180**</td>
<td>-0.074</td>
<td>-0.054</td>
<td>-0.228**</td>
<td>0.719**</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

5.3.1 Study Hypotheses C1 and C2: In T1, incidence of late stage CRC incidence (C1) and CRC-specific mortality (C2) are positively associated with percent black, percent Hispanic, percent poverty, age, and spatial autocorrelation (lag).

For late stage CRC incidence and CRC-specific mortality in T1, Tables 5.14 and 5.15 present regression coefficient estimates, standard errors, t statistics, and statistics that evaluate the extent to which the negative binomial regression model fits the distribution of the data ("summary fit statistics"). With the introduction of a spatial lag term, standardized estimates become statistically inefficient estimates (Griffith, 2002, 2004). Therefore, all models report standardized pseudo-estimates that should be interpreted with caution. In both models, the Likelihood-Ratio (LR) test for the dispersion parameters (alpha) (see Tables 5.14 and 5.15) was statistically significantly different from zero, indicating that overdispersion exists. These results indicate that the negative binomial regression model is a better fit than the Poisson regression model. VIFs for both models are presented in Tables 5.16 and 5.17. Most of the VIFs for both
models were below 4, with only percent black showing a moderate indication of multicollinearity impacts with a factor of just over 5. Thus, the independent variables did not produce high enough VIFs of 10 or greater, the cut-off point, to be excluded from the regression analysis.

The results from the first model show that late stage CRC incidence cases, which was the dependent variable, was statistically significantly associated at the 5 percent level with the variables of percent below poverty, age 30 to 49, age 50 to 79, age 80 and older, and the log-population. Percent black, percent Hispanic, and the spatial lag term were not significant. The second model showed that CRC-specific mortality, the dependent variable, was significantly associated with percent below poverty, percent age 30 to 49, percent age 50 to 79, age 80 and older, and the log-population. Percent black and percent Hispanic were not statistically significant. The data support the hypothesis that late stage CRC incidence and CRC-specific mortality are positively associated with age and percent poverty at T1, but do not support the hypothesis of an association with percent black or percent Hispanic. The spatial lag term was not statistically significant, meaning there is no significant spatial autocorrelation with the current model specification, although this outcome is unreliable because standard error estimates are inefficient due to the use of pseudo-likelihood estimation techniques. Visual inspection of the maps in Figures 5.7, 5.8, 5.9 and 5.10 reveals that positive spatial autocorrelation indeed is present.
### Table 5.14: Negative Binomial Regression, Late Stage CRC Incidence, 1988-1992

<table>
<thead>
<tr>
<th>Late Stage CRC Incidence</th>
<th>Coefficient (SE)</th>
<th>Pseudo Standard Error (SE)</th>
<th>Z score</th>
<th>Confidence Interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black %</td>
<td>-0.33 (0.28)</td>
<td>0.78</td>
<td>-0.47</td>
<td>-.63</td>
</tr>
<tr>
<td>Hispanic %</td>
<td>-0.13 (0.23)</td>
<td>-0.57</td>
<td>-0.58</td>
<td>0.31</td>
</tr>
<tr>
<td>Below Poverty %</td>
<td>0.84* (0.40)</td>
<td>2.08</td>
<td>0.50</td>
<td>1.63</td>
</tr>
<tr>
<td>Age 30 to 49 %</td>
<td>2.94** (1.17)</td>
<td>2.50</td>
<td>0.63</td>
<td>5.24</td>
</tr>
<tr>
<td>Age 50 to 79 %</td>
<td>4.47** (0.82)</td>
<td>5.45</td>
<td>2.86</td>
<td>6.08</td>
</tr>
<tr>
<td>Age 80 and older %</td>
<td>3.29** (1.38)</td>
<td>2.38</td>
<td>0.58</td>
<td>6.00</td>
</tr>
<tr>
<td>Population Log</td>
<td>0.95** (0.68)</td>
<td>13.87</td>
<td>0.82</td>
<td>1.08</td>
</tr>
<tr>
<td>CRC Spatial Lag</td>
<td>0.001 (0.002)</td>
<td>0.90</td>
<td>-0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>Intercept</td>
<td>-10.22** (0.98)</td>
<td>-10.45</td>
<td>-12.14</td>
<td>-8.30</td>
</tr>
</tbody>
</table>

| Dispersion Parameter     | 0.13             | 0.02                       | 0.09    | 0.18                   |

Test of Fit Statistics
- Likelihood-ratio test of alpha = 0  
- Goodness-of-fit X square (1)a = 88.27  
- Prob. Value= 0.000

* Significant at the p < 0.05  
** Significant at the p< 0.01  
a = Degrees of freedom

### Table 5.15: Negative Binomial Regression, CRC Mortality, 1988-1992

<table>
<thead>
<tr>
<th>CRC Mortality</th>
<th>Coefficient (SE)</th>
<th>Pseudo Standard Error (SE)</th>
<th>Z score</th>
<th>Confidence Interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black %</td>
<td>0.08 (0.28)</td>
<td>0.29</td>
<td>-0.47</td>
<td>0.63</td>
</tr>
<tr>
<td>Hispanic %</td>
<td>-0.13 (0.23)</td>
<td>-0.57</td>
<td>-0.58</td>
<td>0.32</td>
</tr>
<tr>
<td>Below Poverty %</td>
<td>0.84* (0.40)</td>
<td>2.08</td>
<td>0.05</td>
<td>1.63</td>
</tr>
<tr>
<td>Age 30 to 49 %</td>
<td>2.94* (1.76)</td>
<td>2.50</td>
<td>0.63</td>
<td>5.24</td>
</tr>
<tr>
<td>Age 50 to 79 %</td>
<td>4.47** (0.82)</td>
<td>5.45</td>
<td>2.87</td>
<td>6.08</td>
</tr>
<tr>
<td>Age 80 and older %</td>
<td>3.29* (1.38)</td>
<td>2.38</td>
<td>0.58</td>
<td>6.00</td>
</tr>
<tr>
<td>Population Log</td>
<td>0.95** (0.07)</td>
<td>13.87</td>
<td>0.81</td>
<td>1.08</td>
</tr>
<tr>
<td>CRC Spatial Lag</td>
<td>0.001 (0.002)</td>
<td>0.90</td>
<td>-0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>Intercept</td>
<td>-10.22** (0.98)</td>
<td>-10.45</td>
<td>-12.14</td>
<td>-8.30</td>
</tr>
</tbody>
</table>

| Dispersion Parameter    | 0.13             | 0.02                       | 0.09    | 0.18                   |

Test of Fit Statistics
- Likelihood-ratio test of alpha = 0  
- Goodness-of-fit X square (1)a = 88.27  
- Prob. Value= 0.000

* Significant at the p < 0.05  
** Significant at the p< 0.01  
a = Degrees of freedom

Model Statistics
- Model Pseudo R square = 0.14  
- Model LR(7)a X square = 230.61
### Table 5.16: VIFs, Late Stage CRC Incidence, 1988-1992

<table>
<thead>
<tr>
<th></th>
<th>VIF</th>
<th>VIF/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent black</td>
<td>5.09</td>
<td>0.1966</td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>3.92</td>
<td>0.2552</td>
</tr>
<tr>
<td>Percent age 50 – 79</td>
<td>3.39</td>
<td>0.2948</td>
</tr>
<tr>
<td>Percent below poverty</td>
<td>2.74</td>
<td>0.3651</td>
</tr>
<tr>
<td>Percent 80 plus</td>
<td>2.68</td>
<td>0.3733</td>
</tr>
<tr>
<td>Spatial Lag</td>
<td>2.12</td>
<td>0.4722</td>
</tr>
<tr>
<td>Percent age 30 – 49</td>
<td>2.00</td>
<td>0.5001</td>
</tr>
<tr>
<td>Population Log</td>
<td>1.28</td>
<td>0.7803</td>
</tr>
<tr>
<td><strong>Mean VIF</strong></td>
<td><strong>2.90</strong></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5.17: VIFs, CRC Mortality, 1988-1992

<table>
<thead>
<tr>
<th></th>
<th>VIF</th>
<th>VIF/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent black</td>
<td>5.19</td>
<td>0.1928</td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>3.99</td>
<td>0.2505</td>
</tr>
<tr>
<td>Percent age 50 – 79</td>
<td>3.39</td>
<td>0.2951</td>
</tr>
<tr>
<td>Percent below poverty</td>
<td>2.71</td>
<td>0.3692</td>
</tr>
<tr>
<td>Percent 80 plus</td>
<td>2.67</td>
<td>0.3740</td>
</tr>
<tr>
<td>Spatial Lag</td>
<td>2.14</td>
<td>0.4680</td>
</tr>
<tr>
<td>Percent age 30 – 49</td>
<td>1.99</td>
<td>0.5020</td>
</tr>
<tr>
<td>Population Log</td>
<td>1.29</td>
<td>0.7770</td>
</tr>
<tr>
<td><strong>Mean VIF</strong></td>
<td><strong>2.92</strong></td>
<td></td>
</tr>
</tbody>
</table>

5.3.2 **Study Hypotheses D1 and D2:** In T2, incidence of late stage CRC incidence (D1) and CRC-specific mortality (D2) are positively associated with percent black, percent Hispanic, percent poverty, age, and spatial autocorrelation (lag).

Tables 5.18 and 5.19 present regression coefficient estimates, standard errors, t statistics, and summary fit statistics for late stage CRC incidence and CRC-specific mortality regression models in T2. In both models, the LR test for the dispersion parameter (alpha) was statistically significantly different from zero, indicating the negative binomial regression specifications had a good model fit. Tables 5.20 and 5.21 summarize the VIFs for each model. The bivariate correlation table identified strong correlations among percent black, percent Hispanic, percent age 50 to 79, and percent age 80 and older. However, because the VIFs for each of these variables produced mean factors below 10, none of the variables were excluded from the models. In model 1, late stage CRC incidence was significantly associated at the 5 percent level with the age group 30 to 49, ages 50 to 79, ages 80 and older, and log-population. Percent black, percent Hispanic, percent below poverty, and the spatial lag term were not statistically significant. In model 2, CRC-specific mortality was significantly associated with percent
age 50 to 79, age 80 and older, log-population, and the spatial lag covariate. Percent
black, percent Hispanic, percent below poverty, and percent age 30 to 49 were not
significant. The data thus support the hypothesis that late stage CRC incidence and CRC-
specific mortality are positively associated with age at T2, but does not support the
hypothesis of an association with percent black, percent Hispanic, or percent poverty for
that period. The results demonstrate significant spatial autocorrelation for CRC-
mortality, but spatial autocorrelation was not found with respect to late stage CRC
incidence.

Table 5.18: Negative Binomial Regression, Late Stage CRC Incidence, 1998-2002

<table>
<thead>
<tr>
<th>Late Stage CRC Incidence</th>
<th>Coefficient</th>
<th>Pseudo Standard Error</th>
<th>Z score</th>
<th>Confidence Interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black %</td>
<td>0.12</td>
<td>(0.19)</td>
<td>0.64</td>
<td>-0.25 - 0.50</td>
</tr>
<tr>
<td>Hispanic %</td>
<td>-0.10</td>
<td>(0.15)</td>
<td>-0.65</td>
<td>-0.34 - 0.20</td>
</tr>
<tr>
<td>Below Poverty %</td>
<td>0.11</td>
<td>(0.26)</td>
<td>0.42</td>
<td>-0.40 - 0.61</td>
</tr>
<tr>
<td>Age 30 to 49 %</td>
<td>1.61*</td>
<td>(0.81)</td>
<td>1.99</td>
<td>0.03 - 3.19</td>
</tr>
<tr>
<td>Age 50 to 79 %</td>
<td>3.82**</td>
<td>(0.59)</td>
<td>6.43</td>
<td>2.65 - 4.98</td>
</tr>
<tr>
<td>Age 80 plus %</td>
<td>3.92**</td>
<td>(1.13)</td>
<td>3.46</td>
<td>1.70 - 6.13</td>
</tr>
<tr>
<td>Population Log</td>
<td>0.92**</td>
<td>(0.53)</td>
<td>17.17</td>
<td>0.81 - 1.02</td>
</tr>
<tr>
<td>CRC Spatial Lag</td>
<td>0.002</td>
<td>(0.001)</td>
<td>1.74</td>
<td>-0.0002 - 0.003</td>
</tr>
<tr>
<td>Intercept</td>
<td>-9.97**</td>
<td>(0.69)</td>
<td>-12.95</td>
<td>-10.32 - 7.61</td>
</tr>
</tbody>
</table>

| Dispersion Parameter     | 0.34        | 0.01                 | 0.02    | 0.06                    |

Test of Fit Statistics
Likelihood-ratio test of alpha = 0
Goodness-of-fit X square (1)² = 19.74
Prob. Value = 0.000

Model Statistics
Model Pseudo R square = 0.16
Model LR(8)² X square = 349.93

* Significant at the p < 0.05
** Significant at the p< 0.01
a. Degrees of freedom
Table 5.19: Negative Binomial Regression, CRC Mortality, 1998-2002

<table>
<thead>
<tr>
<th>CRC Mortality</th>
<th>Coefficient</th>
<th>Pseudo Standard Error</th>
<th>Z score</th>
<th>Confidence Interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black %</td>
<td>0.42</td>
<td>(0.27)</td>
<td>1.56</td>
<td>-0.11 0.95</td>
</tr>
<tr>
<td>Hispanic %</td>
<td>-0.90</td>
<td>(0.22)</td>
<td>-0.41</td>
<td>-0.52 0.34</td>
</tr>
<tr>
<td>Below Poverty %</td>
<td>-0.08</td>
<td>(0.36)</td>
<td>-0.23</td>
<td>-0.80 0.63</td>
</tr>
<tr>
<td>Age 30 to 49 %</td>
<td>0.64</td>
<td>(1.11)</td>
<td>0.57</td>
<td>-1.54 2.81</td>
</tr>
<tr>
<td>Age 50 to 79 %</td>
<td>3.06**</td>
<td>(0.82)</td>
<td>3.72</td>
<td>1.45 4.67</td>
</tr>
<tr>
<td>Age 80 plus %</td>
<td>4.42**</td>
<td>(1.62)</td>
<td>2.73</td>
<td>1.25 7.6</td>
</tr>
<tr>
<td>Population Log</td>
<td>1.00**</td>
<td>(0.08)</td>
<td>12.85</td>
<td>0.84 1.15</td>
</tr>
<tr>
<td>CRC Spatial Lag</td>
<td>0.008**</td>
<td>(0.002)</td>
<td>3.17</td>
<td>0.003 0.013</td>
</tr>
<tr>
<td>Intercept</td>
<td>-10.32**</td>
<td>(0.98)</td>
<td>-10.52</td>
<td>-12.25 -8.40</td>
</tr>
</tbody>
</table>

**Dispersion Parameter**

<table>
<thead>
<tr>
<th>Test of Fit Statistics</th>
<th>Model Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood-ratio test of alpha = 0</td>
<td>Model Pseudo R square = 0.16</td>
</tr>
<tr>
<td>Goodness-of-fit X square (1)* = 27.55</td>
<td>Model LR(8)² X square = 213.87</td>
</tr>
</tbody>
</table>

Prob. Value= 0.000

* Significant at the p < 0.05
** Significant at the p< 0.01

a. Degrees of freedom

Table 5.20: VIFs, Late Stage CRC Incidence, 1998-2002

<table>
<thead>
<tr>
<th>VIF</th>
<th>VIF/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black %</td>
<td>5.94</td>
</tr>
<tr>
<td>Hispanic %</td>
<td>4.30</td>
</tr>
<tr>
<td>Age 50 – 79 %</td>
<td>2.82</td>
</tr>
<tr>
<td>Age 80 plus %</td>
<td>2.49</td>
</tr>
<tr>
<td>Below poverty %</td>
<td>2.17</td>
</tr>
<tr>
<td>Spatial Lag</td>
<td>1.8</td>
</tr>
<tr>
<td>Age 30 – 49 %</td>
<td>1.4</td>
</tr>
<tr>
<td>Population Log</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Mean VIF 2.77

Table 5.21: VIFs, CRC Mortality, 1998-2002

<table>
<thead>
<tr>
<th>VIF</th>
<th>VIF/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black %</td>
<td>6.07</td>
</tr>
<tr>
<td>Hispanic %</td>
<td>4.28</td>
</tr>
<tr>
<td>Age 50-79 %</td>
<td>2.82</td>
</tr>
<tr>
<td>Age 80 plus %</td>
<td>2.49</td>
</tr>
<tr>
<td>Below poverty %</td>
<td>2.17</td>
</tr>
<tr>
<td>Spatial Lag</td>
<td>1.8</td>
</tr>
<tr>
<td>Age 30-49 %</td>
<td>1.4</td>
</tr>
<tr>
<td>Population Log</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Mean VIF 2.78
5.4 Discussion of Results

5.4.1 Patterns of Late Stage CRC Incidence and CRC-Specific Mortality in Miami-Dade County

Age-adjusted CRC rates in Miami-Dade County vary by race and ethnicity and between the two time periods studied. In T1, age-adjusted rates were calculated for the categories Hispanic, white, and black. This study does not list non-Hispanic white age-adjusted rates in this time period due to a lack of an appropriate population file from the census (i.e., non-Hispanic white by age). The lowest rates of CRC occurred among Hispanics for CRC incidence, late stage, and mortality. Age-adjusted rates, which include roughly 95 percent of Hispanic cases, depict the highest figures for late stage CRC incidence. These age-adjusted rates would likely be higher if the Hispanic cases were calculated separately from non-Hispanic white rates. CRC-specific mortality rates were highest among blacks during this time period. In T2, age-adjusted rates were calculated for non-Hispanic whites, blacks, and Hispanics, given the more detailed variables available in the 2000 census data. This time period shows similar distributions to T1 by race and ethnicity for all age groups. However, age-adjusted rate trends between T1 and T2 depict an increase in incidence for all racial and ethnic groups studied and an increase in late stage incidence among whites and Hispanics. In contrast, mortality rates declined for all groups from T1 to T2, most substantially among blacks. Comparisons between white rates from T1 and non-Hispanic white rates from T2 also show a decrease. The decrease likely would have been larger for non-Hispanic white rates if they were calculated separately from Hispanic rates in T1, as noted above. The reduction in Hispanic mortality rates was the smallest overall. Non-Hispanic whites had higher rates than any other group of late stage CRC incidence in T2.
This overall decline in CRC-specific mortality from T1 (1988-1992) to T2 (1998-2002) is the most notable descriptive trend, and may suggest that CRC cases are increasingly detected through screenings, thus reducing the number of deaths due to CRC. However, since vital statistics are reported up to the year 2005, mortality rates are likely lower due to the fact that less time has transpired for diagnoses that occurred in T2. Nevertheless, in the mid-1990s, strong evidence of the effect of screenings on CRC-mortality prompted a nation-wide adoption of more aggressive screening guidelines that could partly account for this decline (Mandel, 1993). These new guidelines recommend annual FOBT screening, sigmoidoscopy and colonoscopy every five to ten years for those with average risk, starting at the age of 50 (Rex et al., 2000). In a summary of research results about the effect of screenings on CRC-specific mortality, the National Cancer Institute (NCI 2008) provide a summary of studies that reported a 15-30 percent decrease in mortality through the use of FOBTs, and a 50 percent decrease in mortality through use of sigmoidoscopy (Cotterchio et al., 2005; Gondal et al., 2003; Schoenfeld et al., 2005; Segnan et al., 2005). Since the 1990s, rates in CRC-related mortality have decreased substantially across the nation (NCI 2008). The changes in screening guidelines, which permit earlier detection and treatment, may potentially explain the reduction in CRC-specific mortality in Miami-Dade County between T1 and T2.
5.4.2 Discussion of Cluster Analysis Results

As predicted by the study hypotheses, the cluster analysis identified statistically significant spatial clusters for CRC outcomes in each time period. Controlling for population size and age distribution, the analysis of CRC occurrence in T1 shows four statistically significant clusters of late stage CRC incidence and four significant clusters of CRC-specific mortality, supporting study hypotheses A1 and A2. The cluster analysis also supported hypotheses B1 and B2, identifying in T2 one significant cluster of late stage CRC incidence and a significant cluster of CRC mortality. These results demonstrate that these communities have a higher than expected rate of late stage CRC incidence or CRC-specific mortality than the expected rate for the County.

The identification of neighborhoods with higher than expected rates is important for public health surveillance and interventions that target high risk populations. Additionally it provides a way to identify potential patterns and explore the social and economic characteristics of the communities most at risk. While the cluster analysis in this dissertation did not control for race and ethnicity, poverty, or foreign-born populations, it is at least important to know the population characteristics of those tracts that have a statistically significantly higher number of expected CRC cases. The racial, ethnic, and SES levels of the communities composing those clusters also may provide some insight into underlying conditions that lead to higher than expected CRC cases.

The characteristics of these communities do not follow a single template for racial composition or SES for the CRC clusters across the time periods and outcomes. However, the demographics of the eight significant tract clusters in T1 for late stage CRC incidence and mortality do show patterns of tracts with high proportions of Hispanic,
black, and foreign-born populations. There is a range of mean poverty rates across tracts, suggesting that the average SES of communities may not have as much of an impact on CRC outcomes as does racial and ethnic composition, or that median level income might be a better indicator of SES. One exception is the T1 Mortality Cluster 3 in Golden Glades, which is only a quarter black and Hispanic and has a relatively low poverty rate at 13 percent. Neither SES nor racial and ethnic composition – at least not of those categories studied here – appears to explain the clusters of CRC-related deaths in that tract.

Fewer clusters were found in T2 than in T1, suggesting a significant decline in clustering. Whether or not the changes to screening guidelines caused the decline in CRC late stage incidence and CRC-related mortality cannot be concluded by this study. Imputing the risk of CRC outcomes faced by an individual living in one of the neighborhoods based on the results of this group-level analysis may commit the ecological fallacy. However, while admittedly speculative, in T2, the Hispanic population in the area of “Coral Way” could have had lower screening compliance rates, resulting in higher occurrences of late stage CRC, but greater overall treatment once diagnosed. The T2 CRC mortality cluster may be caused by high rates of poverty in that area that resulted in the lack of adequate preventive and tertiary care.

5.4.3 Discussion of Regression Analyses

The spatial regression analysis measured the extent to which social and economic factors – percent black, percent Hispanic, percent poverty – were associated with late stage CRC incidence and CRC mortality in Miami-Dade County in the two time periods,
T1 (1988-1992) and T2 (1998-2002). These variables were included in the analysis largely on account of other research that correlated the effects of neighborhood-level socioeconomic status with cancer outcomes (e.g., Hsu et al., 2006; Krieger et al., 1999; Krieger et al., 2003; Singh et al., 2003). A number of health inequalities studies more generally support the theory that racial and ethnic characteristics of neighborhoods influence health outcomes (Bernard et al., 2007; Krieger et al., 1999; Lillie-Blanton & Laveist, 1996; Strait, 2006; Williams, 1999). These researchers theorize that area-level context plays a significant role in determining health outcomes, independent from individual-level factors. Miami-Dade County is a place with a highly diverse racial and ethnic population, with wide socioeconomic disparities, and that ranks first in the country in percentage of foreign-born residents. To determine whether these social and economic variables influenced CRC incidence at the neighborhood-level, population characteristics and CRC incidence were evaluated at the tract level using a negative binomial regression analysis, controlling for spatial autocorrelation, population size, and age.

The regression analysis identified an association with area-level poverty for T1, but the social variables were not statistically significant, thus only partially supporting the study hypotheses C1 and C2. Among the additional variables – age, population size, and the spatial lag – all but the spatial lag showed a statistical association with late stage CRC incidence and CRC mortality. This analysis did not show a correlation between race and ethnicity and CRC outcomes. The most significant result was the positive association between CRC outcomes and percent poverty during this time period, and the spatial clusters of higher than expected rates. In T2, the study hypotheses D1 and D2 posited the same variable associations with CRC outcomes as did those stated in hypotheses C1 and
C2. However, while results from these analyses show significant associations with age and population size, they do not indicate significant associations between CRC outcomes and percent black, percent Hispanic, or percent poverty. Positive spatial autocorrelation, measured by the inclusion of the spatial lag term, was significantly associated with CRC mortality but not late stage CRC incidence. All age categories, 30 and older, were positively associated with CRC late stage incidence, but only the age categories 50 to 79, and 80 and above were associated with CRC mortality. The lifetime probability of dying from CRC between age 30 to 49 is very low at 0.07 percent, which might explain the lack of statistical significance for the variable age 30 to 49. As expected, population size was significantly and positively correlated with CRC outcomes in both T1 and T2.

The results from this study show that neighborhood-level poverty is an important factor influencing CRC outcomes, a finding that is supported by health inequalities literature that associates poverty with lower health care access and screening compliance (DeNavas-Walt et al., 2006; Rosenberg & Wilson, 2000; Strait, 2006; Wagstaff, 2002; Ward et al., 2008). However, the poverty variable was significant only for T1 and not for T2, suggesting that barriers to health care may have been reduced for the T2 population. According to study findings, the poor also are likely to be uninsured, and those who are uninsured are at a higher risk for late stage CRC incidence and death from CRC as a result of lower access to screenings (Roetzheim et al., 2000; Roetzheim et al., 1999; Ward et al., 2008). Among the poor, there is a greater reliance on emergency care as individuals typically lack the financial resources for private or outpatient care services (Begley et al., 2006). Although speculative, more aggressive CRC screening guidelines could account for the disappearance of an association to the poverty variable in T2.
While screening compliance is positively associated with health insurance and SES (Matthews et al., 2005; Neilson & Whynes, 1995), perhaps lower income groups were particularly targeted for intervention and treatment by public health officials.

Alternatively, unobserved variables could be changing the effect of poverty. For example, recent immigrants are more likely to live in poverty and as a result of PRWORA, also have less access to public aid services. The distinction between those who are impoverished but have greater access to public aid versus those who live in poverty and have fewer public resources available to them is one that might account for the insignificant effect of the poverty variable in T2.

The lack of an association between racial and ethnic categories and late stage CRC incidence and CRC mortality was an unexpected finding given prior studies that had found correlations with the percentage of ethnic minorities (Nelson et al., 1997; Piffath et al., 2001; Roetzheim et al., 2000; Wu, 2001). However, there are studies that report no significant associations between these CRC outcomes and race or ethnicity after incorporating socioeconomic variables into statistical models (Ionescu et al., 1998; Lannin et al., 1998; Singh et al., 2003). Nevertheless, the finding that race, ethnicity, or poverty variables did not significantly explain CRC patterns in T2 was unexpected in light of numerous reports (Hsu et al., 2006; Krieger et al., 1999; Neilson & Whynes, 1995; Roetzheim et al., 2000) that discuss the relationship between SES, insurance status, and cancer outcomes. It is possible that more complex social and economic processes explain the CRC rate differences in T2, such as type of medical insurance, cultural beliefs, screening behaviors, and federal and state-level social policies. Conclusions regarding these factors are presented in this dissertation as possible explanations of CRC
differences in Miami-Dade County, but do not attempt to provide an exhaustive list of potential intervening variables.

5.5 Limitations

This dissertation confronted a series of limitations in its analysis of CRC in Miami-Dade County. Regarding the study variables, the cancer database files accessed through FCDS were limited in the availability and reliability of economic and social variables. While insurance status and occupation sector are provided, some studies report the inadequacies in depicting socioeconomic status, given high rates of insurance enrollment at time of diagnosis (Morgan et al., 1997; Roetzheim et al., 1999), and the imperfect correlation between occupation and socioeconomic status (Braveman et al., 2005). Additionally, FCDS patient data on country of birth are unreliable because many cases are not reported directly by patients but through physicians or clinicians. The data quality, therefore, is insufficient for direct analysis of CRC by foreign-born status. For these reasons, cluster analyses of CRC by FCDS-provided economic and foreign-born variables were not conducted.

This dissertation utilized cross-sectional cancer and population data that are static representations of Miami-Dade demographic characteristics and CRC outcomes at two distinct time periods. While only descriptive comparisons were made between T1 and T2 CRC clusters or neighborhood associations, this research does not take into account the spatial mobility of residential populations between time periods. Furthermore, the latency period for which CRC develops can be between 15 to 20 years. This dissertation attempts to identify significant social and economic characteristics at the neighborhood-
level that impact the risk for late stage CRC incidence or CRC mortality, but it does not control for the length of residency. Perhaps the greatest limitation regarding latency is the shorter time period between T2 CRC diagnosis and the time at which vital status was assessed. Vital statistics were reported as all deaths that occurred in 2005 as a result of CRC. CRC survival odds are typically evaluated over a five year period (NCI, 2008). CRC cases that were diagnosed between 2001 and 2002 fall short of this time span, and could account for the fewer mortality cases observed in T2 than were observed in T1.

With respect to the identification of significant clusters, the cluster analysis did not control for race or ethnicity. Rather, this dissertation evaluates the effects of race and ethnicity through spatial regressions. Therefore, clusters of higher than expected CRC rates may be explained by higher numbers of black CRC patients in particular census tracts, given their overall higher mortality rates. Additionally, the geocodes provided by the FCDS showed significant errors in the accuracy of the 2000 census tract assignments, with a match rate of only 35 percent compared to the researcher’s match rate of 92 percent. For this reason, the analysis used the geocodes assigned by the researcher using parcel and street databases as geo-referencing files for a majority of cases diagnosed during 1998 to 2002. While match rates were generally high, many of these geocodes could, in fact, be assigned incorrectly, potentially affecting results.

Limitations of the negative binomial regression analysis include the unreliability of the standardized error estimates resulting from the inclusion of the spatial lag term (Griffith, 2002, 2004). For models in T1, the inclusion of this term reduced the dispersion parameter (from 0.128 to 0.022 for late stage CRC incidence and from 0.143 to 0.034 for CRC mortality), creating a better overall model fit. However, while
variables such as age and population were statistically significant in all models as expected, the lack of significant spatial autocorrelation in three of the models raises some doubt as to the adequacy of the model specifications employed. It is uncommon that events are distributed randomly across space, even when taking into account other variables (Gatrell & Rigby, 2004). Regardless, standardized error estimates should be interpreted with caution.

Finally, while this study examines CRC outcomes in Miami-Dade County in two separate periods in time, it is limited by associating disease occurrence to population characteristics of census tracts at time of diagnosis, and does not consider other variables throughout the life-course of each individual or at different locations of residence that may influence patient-level CRC outcomes. Part of the difficulty in assessing these characteristics is the availability or accessibility of such individualized mobility data. Likewise, census estimates at the tract level are produced every ten years, with limited variable reports during intercensal years. This situation should change as the American Community Survey, which conducts population surveys in intercensal years, becomes fully implemented. Due to these limitations, only cross-sectional census data were used for the years 1990 and 2000. In addition to these limitations, the small sample size of CRC count data at the level of the census tract could affect the dissertation results.

5.6 Areas for Future Research

Findings from this dissertation point to the persistent effects over time of CRC on morbidity and mortality. While trends indicate a decline in late stage CRC incidence and CRC-specific mortality rates in Miami-Dade County between 1988-1992 and 1998-2002,
more significant improvements in CRC outcomes still can be made. Results from this study provide vital information on the status of CRC in Miami-Dade County, but efforts to further understand the complexities are necessary. The first of these efforts should extend spatial cluster analysis to consider CRC outcome distributions by race and ethnicity in addition to age. Given availability of race and ethnicity data from FCDS, comparison of late-stage CRC and CRC-specific mortality can be evaluated using the same methods applied here.

Second, due to the limitations of the negative binomial regression analysis with the inclusion of the spatial lag term to control for spatial autocorrelation, use of additional spatial methods could produce more reliable standard error estimates. One approach that is particularly worth exploring is the spatial filtering method. Comparing results from the negative binomial regression analysis including a spatial lag term with results produced by a spatial filtering approach might indicate significantly different outcomes, and may prove useful in identifying the most appropriate methodological applications and techniques to use in area-based analyses of CRC outcomes.

Finally, primary data collection involving neighborhoods in Miami-Dade County, where higher than expected rates of late stage CRC incidence and CRC mortality were identified, could further facilitate analysis of the mechanisms that drive spatial patterns and influence negative outcomes of CRC. Community surveys, focus groups, or one-on-one interviews would provide greater detailed information on additional contextual features of neighborhoods that were not available in the FCDS or census datasets. Primary data collection methods would also make possible the evaluation of health
seeking behaviors and attitudes as they relate to access to healthcare and screening compliance.
CHAPTER VI

CONCLUSIONS

This dissertation examines the spatial distribution of late stage colorectal cancer (CRC) and CRC-specific mortality in Miami-Dade County in relation to neighborhood-level social and economic characteristics in two distinct time periods that straddle major public policy and health guideline changes. Based on the premise that health outcomes are a product of a complex set of individual, structural, and geographical processes, the analysis of CRC outcomes integrates cross-disciplinary theoretical and methodological approaches from the fields of public health and epidemiology, medical sociology, and medical geography. In particular, this dissertation considers the effects of individual-level characteristics such as age and race, group-level socioeconomic inequalities, and the impact of these factors on CRC status at the scale of the neighborhood. This research employs two distinct analyses of CRC outcomes and social and economic effects. Quantitative spatial statistics are used to directly measure neighborhood associations with CRC outcomes, while a policy-based analysis considers the indirect impacts on CRC risk among vulnerable populations. For the latter, this dissertation examines how the enactment of the Personal Responsibility and Worker Opportunity Act of 1996 affected immigrants’ access to health care, and focuses on barriers to healthcare among elderly non-citizens. Policy instruments such as the PRWORA can exacerbate structural inequities in health patterns, including the risk for delay in diagnosis until the more advanced stages of CRC develop, thus increasing mortality from CRC, a chronic disease condition that is the second leading cause of all cancer related deaths for men and women combined.
CRC is a significant health problem in the United States, but one that is more preventable than any other type of cancer. Over a million men and women alive today have been affected by colorectal cancer. Based on rates from 2002 to 2004, approximately one of every eighteen men and women will be diagnosed with CRC, and one in forty five people will die from CRC (NCI, 2008). Although these figures are disconcerting, even more problematic is the wide disparity in CRC outcomes across racial, ethnic, and socioeconomic groups. Historically, white men and women have had higher CRC incidence and mortality rates than black men and women, but, during the 1980s, national CRC mortality trends began to show significant declines among whites but increases among blacks. Results from randomized clinical trials and case-control studies of CRC screening modalities in the 1990s associated FOBT and sigmoidoscopy with significant reductions in CRC mortality (Hardcastle, 1996; Mandel, 1993). As the beneficial effects of screening for mortality reductions became evident, national screening guidelines adopted more aggressive practices in CRC prevention and early detection. Yet, despite these efforts to promulgate new screening guidelines in primary care practices, screening compliance among the U.S. population has remained very low, and strongly determined by access to and quality of health insurance (Nelson et al., 2003; Roetzheim et al., 2000; Ward et al., 2008). Similar to associations with CRC outcome disparities, access to health insurance is strongly correlated with race, ethnicity, socioeconomic status, and immigration status (Breen et al., 2001; Neilson & Whynes, 1995; Vogelaar et al., 2006). With the enactment of PRWORA in 1996, new public assistance eligibility requirements restricted access to Medicaid for thousands of legal immigrants, creating differences in access to government-provided health insurance by
immigration status. These barriers to healthcare maintenance and disease prevention services can produce widely disparate health outcomes across social and economic groups.

Miami-Dade County provides a unique opportunity for a case study of health inequalities in this nation. The county is a demographically diverse region consisting of a predominantly Hispanic and foreign-born population and exhibits significant patterns of racial and ethnic residential segregation, along with concentrations of neighborhood-level poverty (Boswell, 1990; Census, 1990, 2000b). While Miami-Dade’s participation in global economic trade and finance has produced rapidly rising levels of wealth overall, this transformation has also led to widening income disparities, political disenfranchisement, and the creation of ethnically segregated neighborhoods (Nijman, 1997, 2000; Portes & Stepick, 1993). Access to affordable housing, especially among those who are poor and rely on public housing, is an additional problem facing Miami-Dade County and compounds the financial difficulties of many who reside there. These conditions define the social and economic context in the County and influence, directly and indirectly, the County’s overall public health. In 2005, cancer deaths ranked second among all causes of death in Miami-Dade County, with CRC constituting the second leading cause of all cancer-related deaths (FCDS, 2005). In 2005, the Florida Health Insurance Survey reported that Miami-Dade had the highest level (29%) of uninsured county in the state (Garvan et al., 2005). Given the prevalence of such socio-economic inequalities, as well as high proportions of uninsured individuals in the County, particular attention to population health differences and CRC risks is warranted herein.
The results of this study show that Miami-Dade County exhibits spatial clusters with higher than expected late stage and mortality outcomes for CRC cases diagnosed between 1988 to 1992 (T1) and between 1998 to 2002 (T2), after controlling for age and population size. Many of these neighborhoods are characterized by high percentages of Hispanic, foreign-born, and black residents that have moderate to high levels of poverty, although not all of the identified clusters show these demographics. Spatial regression analyses demonstrate that poverty, age, and population size account for the observed differences in CRC incidence and spatial clusters in T1. While the social variables – percent black and percent Hispanic – did not show the expected associations with CRC outcomes, the significant association of poverty with late stage CRC incidence and CRC mortality in T1 indicates the influence of socioeconomic status on CRC outcomes at the neighborhood level.

The empirical relationship between poverty and the two CRC model outcomes is consistent with literature that links neighborhood-level socioeconomic disadvantage, often measured by poverty, to higher rates of late stage CRC incidence and mortality (Ionescu et al., 1998; Krieger et al., 2005; Link et al., 1998; Mandelblatt, 1996; Parikh-Patel et al., 2006). The poor are disadvantaged by lower incomes, fewer assets, lower levels of education and social cohesion, and greater barriers to health insurance and health services (Coburn, 2000; Duncan et al., 2005; Powell-Griner, 1999; Wagstaff, 2002; Winkleby et al., 1992). These challenges have real health consequences, as is demonstrated in the relationship between poverty and CRC outcomes in Miami-Dade County in T1.
In T2, late stage CRC incidence and CRC mortality in Miami-Dade County displayed fewer spatial clusters and no significant associations with neighborhood compositions of black or Hispanic residents, or with neighborhoods characterized by poverty. Although general improvements in CRC outcomes for T2 were consistent with increasingly aggressive screening guidelines since T1, the lack of association between CRC outcomes and socioeconomic variables were unexpected (Rex et al., 2000; Zoorob et al., 2001). Surprisingly, the BRFSS reported that FOBT and sigmoidoscopy screening compliance in Miami-Dade County remained below 50 percent in the early 2000s (Nelson et al. 2003), which contradicts the expected improvements due to guideline changes. While fewer CRC mortality “hotspots” occurred in T2 than in T1, this may be explained by the shorter time increment between the year of diagnosis and the year that vital status was assessed (2005). However, late stage CRC clusters also were fewer in T2, which is not related to the year of vital status, and suggests that CRC outcomes did improve since T1. While neighborhood social and economic variables did not explain CRC outcomes in T2, clusters displayed similar racial, ethnic, and socioeconomic neighborhood characteristics as clusters in T1; late-stage CRC was a predominantly Hispanic immigrant community with only moderate poverty, and the CRC mortality cluster was made up of overwhelmingly black and impoverished neighborhoods.

A significant limitation in this dissertation was the inability to directly measure individual health insurance status, screening compliance, and foreign-born status, or to control for latency effects. As discussed in this dissertation, public health guidelines in the early 1990s published more aggressive CRC screening guidelines due to evidence supporting that screening can reduce mortality rates, while federal and state changes in
welfare policy put elderly immigrants at increased risk for advanced CRC by limiting access to healthcare services. These events present two opposite effects on CRC risk, which make expectations about CRC in T2 difficult to formulate. The latency periods for which screening guidelines and welfare changes influence CRC outcomes are difficult to account for due to the lack of FCDS data on screening compliance or the lack of efficient foreign-born information, preventing direct measurements. Due to the inadequacy of available data, this dissertation provided a policy-based analysis of national and regional policy changes occurring between T1 and T2 that attempts to indirectly capture the effects of screenings and access to healthcare on CRC risk. While recommendations for CRC screenings changed roughly ten years prior to the second study period, welfare reform was enacted in 1996 and took some years for full implementation, occurring right up to T2. Therefore, the results from the cluster and regression analysis in T2 may not adequately captures the effect of policy changes due to a latency effect.

Although medical guidelines recommend more frequent and aggressive CRC screenings, screening compliance is associated with more complex factors, such as type of health insurance, and cultural beliefs and attitudes. Ward and colleagues (2008) compared cancer screening compliance between the uninsured, privately insured, and those with Medicaid. They found that type of insurance significantly influences screening rates: higher screening rates were found among those privately insured when compared with either Medicaid beneficiaries or the uninsured. Florida residents who were either uninsured or Medicaid beneficiaries also were less likely to receive surgery to treat cancer (Roetzheim et al., 2000). Screening rates among immigrants are also lower than the native or naturalized population due to more fatalistic attitudes toward cancer
and general misconceptions regarding the origin of cancer (Gany et al., 2006). Therefore, while major nationwide studies show a general improvement in CRC outcomes, and while significant associations with social and economic variables are lacking in the analysis at T2, more complex conditions such as health insurance type, attitudes and beliefs, and specific racial and ethnic immigrant characteristics could more adequately explain differences in compliance with and access to early detection and treatment of CRC.

General improvements in the Miami-Dade County public health system resulting in increased access to care for the uninsured between T1 and T2 could also explain the fact that percent poverty was significant in T1 but not in T2. The Urban Institute distributed a household survey to low income residents in Miami-Dade County from 1990 to 1995 (Lipson et al., 1997). Low income respondents reported considerable improvements in access to health care for low income individuals. Other data show growth in the Florida public health budgets and in the number of facilities offering services to the indigent. Between 1990 and 1995, state spending on Medicaid grew by 12.7 percent per year. Beginning in 1992, Miami-Dade also began to collect a half-cent sales tax that was earmarked for the county's public health system and providing insurance for low-income residents, and which, as of 1995, was raising approximately $96 million per year (Lipson et al., 1997). The Urban Institute study also showed that throughout the 1990s, an increasing number of medical care facilities began accepting Medicaid patients, leading to a 20% decrease in the amount of Medicaid-funded care provided by Miami Dade's Jackson-Memorial Hospital, the primary source of care for low-income and uninsured individuals in the County. In 2004, Miami-Dade County's
health department had a budget of nearly $62 million and 1,023 employees (Napier et al., 2004).

Additionally, changes in Florida's administration of its public health system may have led to improved treatment and access to care among the lowest income groups in Miami-Dade County. In 1997, the state created the Florida Department of Health, separating public health administration from the state's social services agency. According to Napier and colleagues (2004), this new agency collaborated with local public health departments in Florida to establish and implement better practice policies, which has likely led to greater effectiveness and efficiency at the local level. Health delivery improvements, such as these, may influence earlier CRC diagnosis and treatment among low-income residents in Miami-Dade County, accounting for the insignificant association of neighborhood-level poverty with late stage CRC incidence or CRC mortality in T2.

While improvements in healthcare delivery were made between T1 and T2, as explained above, this dissertation expected to see persistent effects of neighborhood-level racial and ethnic compositions on CRC outcomes given trends reported in the literature. A limitation of this study, potentially affecting results, is the level of social and economic aggregations, which broadly categorized populations based on race and Hispanic ethnicity, rather than more disaggregated variables that define groups by race, ethnicity, ancestry, and foreign-born status. More specific individual characteristics of FCDS data were either not available or were inadequate measures, and census tract foreign-born status was highly correlated with percent Hispanic.
Miami-Dade County is racially and ethnically diverse, with immigrant residents originating from many different countries. Miami-Dade County’s black population, for example, includes a large number of Haitians. In fact, the cluster in T2 that showed higher than expected cases of CRC-mortality includes the community of Little Haiti, which is home to a concentration of Miami’s 18,000 Haitians, many of whom immigrated to the U.S. between 1995 and 2000 (Sohmer, 2005). Studies focusing particularly on Haitian communities have identified some characteristics that make this group more susceptible to negative CRC outcomes. In 2003, Haitian immigrants in New York had low cancer screening rates of only 30 percent (Gany et al., 2008). Additionally, household surveys of the Haitian community in Miami-Dade County showed low usage rates of public assistance services despite high needs for basic services, including health care (Kretsedemas, 2003). Even though 80 percent of the Haitian survey respondents qualified for public assistance, only seven percent had applied. A majority of the respondents expressed concerns that utilization of public assistance services would lead them to be labeled a public charge, although the authors stated that these concerns were based on misconceptions regarding public charge guidelines. Studies such as these suggest that aggregate racial and ethnic variable at the neighborhood-level do not always capture significant group differences that influence health outcomes.

Legal elderly immigrants represent another group that is not fully accounted for in the data and model analyzed here. Such individuals are particularly affected by the loss of public assistance, which may make them increasingly vulnerable to CRC. The significant cuts in funding and legal changes in qualification requirements enacted by PRWORA have reduced access to public assistance among this group. Before
PRWORA, elderly immigrants (age 65 and older) had the highest usage rate of public aid, as well as the greatest need for long term health care (Binstock & Jean-Baptiste, 1999).

In the state of Florida, Miami-Dade County has the highest concentration of legal elderly immigrants, accounting for 4.2 percent of the nation’s total in 2006 (ACS, 2008). Elderly legal immigrants have higher rates of poverty, lower rates of health insurance coverage, and greater reliance on supplemental security income and Medicaid than their naturalized counterparts (Friedland & Pankaj, 1997). PRWORA’s restrictions on eligibility increases health risks for populations already limited in access to health services, such as disease prevention and treatments.

To date, very few studies have evaluated the impact of PRWORA on elderly non-citizen’s enrollment in public aid programs. Analyses of PRWORA tend to focus more on the law’s effect on working age immigrants. However, ethnographic research shows that legal immigrants across the country, and in Miami-Dade County in particular, express deep fears about seeking public assistance in the aftermath of PRWORA, especially citing its potential impact on their ability to sponsor immigration of family members in the future (Feld & Power, 2000). This reluctance to seek public assistance is amplified by the growing immigrant elderly population in the United States, which is at least partly due to family reunification. Given the greater economic disadvantage of elderly non-citizens, as well as the limitations on legal immigrants’ access to public aid, this group bears increasing risks for negative health outcomes in the future.

Given the recent declines in CRC mortality, as well as the narrowing gap in social and economic differences in CRC outcomes, some CRC prevention efforts in Miami-Dade County appear to be effective. However, while this dissertation reports fewer CRC
“hotspots” and insignificant associations with neighborhood-level social or economic factors, Miami-Dade continues to face pressing public health challenges in the areas of access to insurance and quality care. Access to adequate health care is most consistently associated with racial and ethnic, as well as socioeconomic, inequalities, but is more increasingly determined by recent public assistance restrictions that subject even legal immigrants to status differences. In light of these challenges, the greatest task facing Miami-Dade County in the achievement of greater health equality lies in actively working to mitigate the root causes of present health inequalities and counteract policy effects that could widen disparities in the future.
WORKS CITED


StataCorp. (2007). Stata Statistical Software: Release 10. College, Station, TX: StataCorp LP.


