Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

[Nancy Nguyen] Date

Description of the Spatial Heterogeneity in Colorectal Cancer (CRC) Mortality by Race across Georgia and an Investigation of the Association between Area-Level Factors and the Disparity

By

Nancy B. Nguyen
Master of Public Health

Epidemiology

_____ [Chair's signature]
Lauren E. McCullough, PhD, MSPH
Committee Chair

Description of the Spatial Heterogeneity in Colorectal Cancer (CRC) Mortality by Race across Georgia and an Investigation of the Association between Area-Level Factors and the Disparity

By

Nancy B. Nguyen

B.S., University of Nebraska – Lincoln, 2020

Faculty Thesis Advisor: Lauren E. McCullough, PhD, MSPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
In partial fulfillment of the requirements for the degree of
Master of Public Health
In Epidemiology
2022

Abstract

Description of the Spatial Heterogeneity in Colorectal Cancer (CRC) Mortality by Race across Georgia and an Investigation of the Association between Area-Level Factors and the Disparity

By Nancy B Nguyen

Background
In the US and Georgia, colorectal cancer (CRC) is the 3rd leading cause of cancer death in both men and women (1, 2). Even as mortality rates decline each year, race disparities in CRC mortality persist. Many studies have investigated individual and area-level SES on colorectal cancer mortality and the racial disparity however few have investigated the relationship in a spatial lens. This analysis examines the variation in CRC mortality across counties in Georgia by race and area-level factors that may be drivers in the heterogeneity.

Methods
2,622 colorectal cancer deaths obtained from the 2005-2011 Georgia SEER cancer registries were aggregated at the county level and by race. A descriptive borrowing approach was implemented to describe the heterogeneity of CRC mortality. A conditional auto-regressive (CAR) Bayesian model – a disease mapping method – was used to examine spatial patterns of mortality and whether area-level factors were drivers in the variation seen. Area-level data were obtained from the American Community Survey 5-year estimates, for the period 2006-2010. To determine whether area-level factors were drivers of heterogeneity, the DIC, model fit statistic, was compared between models.

Results
The descriptive borrowing approach suggests that there is moderate spatial heterogeneity of CRC mortality across counties in Georgia by both NHW and NHB groups. Among NHW, excess deaths appeared more in the southeastern region of Georgia. Among NHB, excess deaths appeared more in the western region of Georgia. The CAR Bayesian model suggests that there is moderate spatial heterogeneity of CRC mortality among NHW but little variation among NHB. For both NHW and NHB, area-level poverty and area-level education did not explain the observed variation.

Conclusion
This analysis provides evidence of moderate spatial heterogeneity of CRC mortality in Georgia. Further research should establish other factors that may be contributing to heterogeneity to better target specific interventions that aim to improve the mortality disparity.

Description of the Spatial Heterogeneity in Colorectal Cancer (CRC) Mortality by Race across Georgia and an Investigation of the Association between Area-Level Factors and the Disparity

By

Nancy B Nguyen

B.S., University of Nebraska – Lincoln, 2020

Thesis Committee Chair: Lauren E. McCullough, PhD, MSPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2022

Acknowledgements

I would like to thank Dr. Lauren McCullough, my thesis advisor, for her guidance and support throughout my entire MPH experience. Even though my time at Rollins has been mostly virtual, I am grateful for the effort Lauren has made to make me feel connected and a part of her trainee team. Lauren is an amazing advisor and inspiring researcher in health disparities. I am thankful to have been able to learn and work alongside her.

I would also like to thank Rebecca Nash for her feedback and help with my spatial analysis. Rebecca was always willing to answer any concerns that I had and provided great insight when I was stuck on a coding block.

Finally, I would like to thank my family and friends for all their love and support. To my mom and my brothers, thank you for supporting my decision to move to another state and pursue this career. We are all so far away, but I always look forward to our texts and facetimes. To my partner, Phillip, thank you for being patient and an open ear on my hard days. To my thesis accountability group, I have never met such an admirable and hard-working group of women. I will remember our Thursdays forever.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER I: BACKGROUN/LITERATURE REVIEW

1.1 Mortality of Colorectal Cancer

In the United States (US), colorectal cancer (CRC) is the 3rd leading cause of cancer death in both men and women, with more than an estimated 50,000 CRC deaths in 2020 (2). CRC mortality has decreased in both sexes, where the male mortality rate decreased by 3.9% per year and the female mortality rate by 3.4% per year in 2002 – 2006 (3). The downward trend is due to changing risk behaviors, improved cancer treatments, and uptake in screening (3). The impact of early detection and screening has played a large role in decreasing mortality, especially since the stage at diagnosis is the strongest predictor of CRC survival. When detected early at the localized stage, the 5-year survival rate for CRC is approximately 91%, but this drops to as low as 14% for distant stage (2).

1.1.2 Age and CRC Mortality

CRC mortality increases with age. CRC predominately affects individuals greater than 50 years of age, with the median age at diagnosis being 66 and 69 years for men and women, respectively (2). Mortality rates for adults 50+ have been declining, yet, rates for those under 50 have increased by 1.3% per year since 2004 (2). Some studies have even seen increasing incidence among persons less than 50 years old, raising mixed support for screening age to change from 50 to 45 (3, 4). Promoting awareness for early-onset CRC is necessary, as the cause of early-onset CRC is still unknown and young patients are 58% more likely to present with distant versus localized disease than older patients, which we know to impact mortality (5).

1.1.3 Anatomic Site Differences and CRC Mortality

Research today still groups and analyzes CRC as one entity when several new studies have shown that different anatomical sites show different clinical, epidemiologic, and molecular traits. Tumors located in the proximal colon, which includes the cecum, ascending colon, and transverse colon, are often harder to detect through screening compared to distal colon cancers (6). Proximal colon cancer patients are more likely to be diagnosed at later stages and have even been shown to be 13% less likely to survive 5 years compared to distal tumors (7, 8). Epidemiologically, proximal colon cancers tend to develop more in women, older patients, and Black individuals (9, 10). Whereas among colon cancer diagnoses, distal colon patients were more likely to be White males (11). In addition to this, rectal cancer patients often experience worse prognoses compared to colon cancer patients (11). In a molecular lens, proximal colon cancers have a higher incidence of the microsatellite instability phenotype (12). Distal colon tumors and rectal tumors have a high incidence of chromosomal instability (13, 14).

1.2 Geographic Variation in CRC Mortality

In the US, CRC mortality varies geographically. In the 1980s, the Northeast region showed high CRC mortality rates but has now shifted to be highest in Midwestern and Southern states (2, 15, 16). In a study identifying county hotspots for early-onset CRC mortality, 92% of the identified hotspots were located in the South (15). The variations in mortality also differ within states. In a retrospective cohort study of 30,100 Georgia residents diagnosed with CRC, rural residents showed a 14% increased risk of death compared to urban residents (17). Rural residents nationwide are also more likely to live in poverty, live further away from cancer care services, are less likely to use screening methods, and have a higher risk of late-stage diagnoses – all factors that are associated with increased mortality (17–21).

1.3 Race and CRC Mortality

Negative social factors tend to be more prevalent among minority communities due to U.S. history, policies, and treatment of minorities. Compared to Non-Hispanic Whites (NHW), racial/ethnic minorities have higher rates of poverty, lower educational statuses, and less access to health care coverage (22). These disparities affect the entire cancer care continuum, including preventative screening, treatment, and survivorship (23). In the case of CRC, racial/ethnic minorities experienced lower 5-year survival compared to NHW, even after controlling for census-tract level poverty (24).

CRC mortality rates are highest among Non-Hispanic Blacks (NHB) compared to any other demographic group. Elevated mortality among NHB persists regardless of gender or age despite declining mortality rates for racial/ethnic groups, overall (2, 3, 22, 25). CRC mortality rates are roughly 40% higher for NHB (19.0 per 100,000 population) compared to NHW (13.8 per 100,000 population) (2). Studies have shown that there are racial differences in stage at diagnosis and location of diagnosis. For instance, NHB are more likely to be diagnosed with late-stage disease compared to NHW counterparts (2, 26). NHB are also more likely to be diagnosed with proximal colon cancer, which is harder to detect during screening and may, in part, account for differences in prognostic outcomes (9). Furthermore, NHB men and women are more likely to develop colorectal cancer at younger ages and among those with early-onset diagnoses, survival for NHB is significantly worse compared to NHW (27, 28).

1.4 Neighborhood and Structural Factors and CRC Mortality

In the U.S., historical racism and other structural and institutional barriers have affected racial/ethnic minorities and their health. Many of these historical, structural, and institutional

barriers manifest in modern-day neighborhood differences. There has been growing interest in how place, specifically neighborhood-level factors, affect health (29–31). An individual's neighborhood has been shown to have a large impact on individuals' health and health behavior. Specifically, a neighborhood's physical and social attributes can influence individuals' physical, psychological, and social well-being, contributing to race/ethnic health inequalities (29).

Additionally, social and structural neighborhood characteristics influence CRC mortality directly through lifestyle behaviors, health literacy, and access to cancer services (32). For example, areas with higher populations of poverty are more likely to be food deserts, areas with limited access to healthy and affordable foods (33, 34). In a population-based study, 5-year CRC survival for food desert residents was 4% lower than those not living in food deserts (34). Higher area poverty rate is also inversely related to CRC screening use (35). Furthermore, those living in lower SES census tracts were shown to have less spatial accessibility to colonoscopy services and were less likely to receive therapy or surgery (17, 36). Social and structural neighborhood characteristics thus can impact CRC mortality indirectly through stage at diagnoses and directly through access and quality of care (21, 22).

1.4 Significance of Thesis

There are few known drivers of CRC mortality disparities beyond age and stage of diagnosis. Outside of individual-level factors, emerging literature suggests that social and structural neighborhood characteristics may play an important role in CRC mortality disparities. Traditional methods lack the ability to characterize accurate measures of race-specific mortality due to small numbers and potential spatial dependence of observations. Spatial smoothing allows for the use of prior information from neighboring counties to combat this instability. Therefore, this study will be the first in Georgia to provide data on the spatial association of area-level

metrics on colorectal cancer mortality. The knowledge from this thesis may also assist in the targeting of interventions for specific counties and racial/ethnic groups in Georgia. Furthermore, this analysis may help answer future questions about how neighborhoods affect colorectal cancer outcomes in the Georgia and more broadly in the U.S.

1.7 Aims of Thesis

Aim 1: To characterize the spatial heterogeneity of CRC mortality across counties in Georgia by race.

Hypothesis: Rural counties in Georgia will have higher CRC mortality compared to urban counties for both Non-Hispanic Whites and Non-Hispanic Blacks.

Aim 2: To explore the relationship between area-level poverty and education measures as a potential driver of heterogeneity in CRC mortality.

Hypothesis: Areas with higher percentages of poverty will have higher CRC mortality compared to areas with lower area-level poverty. Area-level poverty will also explain CRC mortality. Areas with higher percentages of low educational attainment will also have higher CRC mortality.

Description of the spatial heterogeneity in colorectal cancer (CRC) mortality by race across Georgia and an investigation of the association between area-level factors and the disparity

Chapter II: SPATIAL HETEROGENEITY OF COLORECTAL CANCER MORTALITY

2.1 Abstract

Background

In the US and Georgia, colorectal cancer (CRC) is the 3rd leading cause of cancer death in both men and women (1, 2). Even as mortality rates decline each year, race disparities in CRC mortality persist. Many studies have investigated individual and area-level SES on colorectal cancer mortality and the racial disparity however few have investigated the relationship in a spatial lens. This analysis examines the variation in CRC mortality across counties in Georgia by race and area-level factors that may be drivers in the heterogeneity.

Methods

2,622 colorectal cancer deaths obtained from the 2005-2011 Georgia SEER cancer registries were aggregated at the county level and by race. A descriptive borrowing approach was implemented to describe the heterogeneity of CRC mortality. A conditional auto-regressive (CAR) Bayesian model – a disease mapping method – was used to examine spatial patterns of mortality and whether area-level factors were drivers in the variation seen. Area-level data were obtained from the American Community Survey 5-year estimates, for the period 2006-2010. To determine whether area-level factors were drivers of heterogeneity, the DIC, model fit statistic, was compared between models.

Results

The descriptive borrowing approach suggests that there is moderate spatial heterogeneity of CRC mortality across counties in Georgia by both NHW and NHB groups. Among NHW, excess deaths appeared more in the southeastern region of Georgia. Among NHB, excess deaths appeared more in the western region of Georgia. The CAR Bayesian model suggests that there is moderate spatial heterogeneity of CRC mortality among NHW but little variation among NHB. For both NHW and NHB, area-level poverty and area-level education did not explain the observed variation.

Conclusion

This analysis provides evidence of moderate spatial heterogeneity of CRC mortality in Georgia. Further research should establish other factors that may be contributing to heterogeneity to better target specific interventions that aim to improve the mortality disparity.

2.2 Introduction

In the US and Georgia, colorectal cancer (CRC) is the 3rd leading cause of cancer death in both men and women, with more than an estimated 50,000 CRC deaths in 2020 (1, 2). CRC mortality rates have been declining in the US since 1990 (2). However, in Georgia, mortality rate trends differ by age, race, and gender. Among adults less than 50 years of age, mortality rates had increased in 1990-2008, whereas adults above 50 experienced mortality rates that declined per year (37). CRC mortality has decreased in both Black males and White males, where the Black male mortality rate decreased by 0.5% per year and the White male mortality rate by 4.5% per year in 2002 – 2008 (37). CRC mortality has decreased in both Black females and White females, where the Black female mortality rate decreased by 0.5% per year and the White female mortality rate by 1.4% per year in 1990 – 2008 (37). Overall, these mortality rate trends from Georgia are comparable to those of the US more generally.

CRC survival is influenced by many factors, such as stage at diagnosis and individual characteristics such as age, race/ethnicity, and socioeconomic status (SES) (38). For both the US general population and the population of the state of Georgia, race disparities persist in CRC mortality. In the US in 2012-2016, CRC mortality rates among Non-Hispanic Blacks (NHB) were 40% higher compared to Non-Hispanic Whites (NHW) (2). This trend was seen in Georgia in 2006-2011, where Black males had the highest age-adjusted mortality rate of 29 deaths per 100,000, followed by Black women (19 deaths per 100,000), White males (18 deaths per 100,00), and the lowest in White females (12 deaths per 100,000) (1).

Many studies in the US have investigated the influence of individual-level risk factors such as SES, obesity, and health insurance on CRC mortality (22, 39, 40). In particular, studies using different measures of SES have consistently found that those who are from lower SES

have an increased risk of mortality from CRC (22, 41). A cohort study found that stage at diagnosis and SES partially explained the difference in mortality rates among Black individuals compared to Whites (42). Further examining racial disparities in CRC outcomes, one study found that ~50% of the Black-White disparity in mortality was explained by screening use and stage-specific CRC survival (43).

Georgia has one of the biggest metropolitan cities, Atlanta, in the US, and a large Black population, with roughly 2 million (32.6%) of its citizens identifying as Black (44, 45). Atlanta also has the largest income inequality in the nation (46). Low SES disproportionately affects Black individuals and reinforces racial health disparities across multiple aspects of the cancer care continuum (25, 47). To our knowledge, there have been few studies in Georgia focusing on how SES influences CRC mortality. In an exception, one study found that those living in lower-middle and low-SES areas in Georgia were at greater risk of death following CRC diagnosis compared to higher or upper-middle SES individuals (lower-middle: HR = 1.16, low: HR = 1.24) (17). Another study found a moderate negative correlation between living in worse social/economic environment and Mortality-Incidence-Ratios among Black CRC patients, whereas among White individuals a strong positive correlation was observed (48). One study did find that rural residence increased risk of death after CRC diagnosis, that was completely explained by census tract level SES (17). Though these studies looked at how individual factors influence risk of CRC mortality, they did not account for neighborhood environment. While one study assessed rural-urban differences, these proxies of neighborhood environment may not be sensitive to critical factors that influence behavior. Importantly, the environments in which people live and work shape individual choices that affect downstream health outcomes; therefore, considerations of the neighborhood environment must be accounted for in studies of

CRC outcomes. Additionally, understanding area-level drivers of CRC disparities can help guide public health policy and intervention, as they may be more amendable to intervention.

Geospatial analysis is a growing statistical method in public health surveillance that incorporates space to understand health outcomes. Geospatial analysis, therefore, provides a novel and valuable method for sensitively assessing neighborhood environmental effects on health outcomes. To date, several studies have focused on residential racial segregation, accessibility of screening services, and broad measures of socioeconomic status on late-stage diagnosis and CRC mortality from a geospatial lens (16, 49, 50). One study found that among those with low behavioral risk for CRC mortality, an increase in SES deprivation increased predicted CRC mortality rates by 11 people per 100,000 (16). However, there are few studies focusing on the spatial influence of area-level metrics on CRC outcomes as a contributor to racial disparities, and no studies done in the state of Georgia.

This thesis project aims to characterize the spatial structure of CRC mortality by race across counties in Georgia by comparing race-specific standardized mortality ratios (SMRs). Furthermore, we aim to explore area-level poverty and education as potential drivers of heterogeneity in CRC mortality across the state. By identifying areas with greater disparity, funding and interventions could be focused to increase equity in areas of need.

2.3 Methods

Data Sources

Individual patient, mortality, and population data were obtained using the information provided by Surveillance, Epidemiology, and End Results (SEER), an organization funded by the National Cancer Institute. The SEER cancer registry collects information on individual

demographics and tumors for all incident cancer cases. Data was obtained from SEER 18 and the SEER*Stat software, which includes the 3 SEER registries in Georgia at the time of diagnosis. The 3 Georgia registries, collected from the Georgia Center for Cancer Statistics (a division within Rollins School of Public Health's Epidemiology department), include Atlanta, Greater Georgia, and Rural Georgia. The Atlanta (Metropolitan) registry consists of Clayton, Cobb, DeKalb, Fulton, and Gwinnett County. The Rural Georgia registry consists of Glascock, Greene, Hancock, Jasper, Jefferson, Morgan, Putnam, Taliaferro, Warren, and Washington County. All remaining counties are included in the Greater Georgia registry.

Area-level information was obtained utilizing the US Census Bureau's American Community Survey (ACS) 5-year data, 2006-2010. The ACS is a nationwide cross-sectional survey that provides information on the nation's social, economic, housing, and demographic characteristics annually (51).

Study Population

Participants met the following criteria: individuals were diagnosed in Georgia with non-metastatic colon or rectal cancer from 2005-2011. The diagnosis of colon or rectal cancer had to be the first lifetime cancer diagnosis and diagnosed at the age of 20 or older. The aim of the study was to compare racial differences between non-Hispanic Black and non-Hispanic White participants, so study participants had to self-identify as 1 of these 2 racial/ethnic groups. This resulted in an initial study population of N=16,157 CRC cases.

Individual CRC-specific deaths were included in the final analysis if survival months were less than or equal to 60 months (5 years) from date of diagnosis to ensure equal follow up time. Death was considered colorectal cancer-specific if the cause of death was recorded as "Colon excluding Rectum" or "Rectum and Rectosigmoid Junction". If survival months and/or

cause of death were unknown, the participants were excluded from the study. This resulted in a final population of N = 2,622 CRC-specific deaths for analysis. Observed deaths were then aggregated at the county level and by race. The research question focuses on the race-specific spatial heterogeneity of CRC mortality across counties in Georgia. Therefore, individual CRC-specific deaths were aggregated at the county level using the county of residence at the time of cancer diagnosis of each patient.

Additional individual demographic information pulled from SEER included gender, race/ethnicity (NHW/NHB), age at diagnosis, county residence at diagnosis, primary site of cancer, tumor stage, survival months, and cause of death for descriptive use.

The International Classification of Disease for Oncology, Third Edition (ICD-O-3)/World Health Organization (WHO) 2008 guidelines were used to identify CRC cases. Colorectal was defined as the cecum, appendix, ascending colon, hepatic flexure, transverse colon, splenic flexure, descending colon, sigmoid colon, large intestines, rectosigmoid junction, and rectum. For tumor-specific information, each participant was classified according to SEER's summary staging manual 2000. According to the SEER website, the 2000 version contains the most precise clinical and pathological information obtained from medical records to describe the extent of disease (52). Tumor stage was classified as either localized or regional. Individuals were identified as having CRC as a primary cancer if their sequence number was defined as: one primary only, 1st of 2 or more primaries, one state registry-defined neoplasm, or 1st of 2 or more state-registered defined neoplasms.

Covariates

The research question focuses on the race-specific spatial heterogeneity of CRC mortality across counties in Georgia and whether area-level poverty and education explain the heterogeneity observed. Therefore, there are two covariates of interest: area-level poverty and area-level education.

Area-level poverty

To measure area-level poverty, county level data was used to approximate poverty. Area-level poverty was determined by estimating the percentage of persons living below the federal poverty line. Counties with greater than 20% of persons living below the federal poverty line were considered poor. Percentages were derived from the 2006 – 2010 ACS.

Area-level education

To measure area-level education, county level data was used to approximate educational attainment. Area-level education was determined by estimating the percentage of adults over 25 with less than 12^{th} grade education (no high school degree or GED equivalent). Percentages were derived from the 2006 – 2010 ACS.

Rurality

Rurality definitions followed the Georgia Department of Public Health guidelines (53) . Any county with a population of less than 50,000 according to the United States 2010 census was categorized as Rural. Any county with a population of 50,000 or more according to the United States 2010 census was categorized as non-Rural.

Outcomes

Standard Mortality Ratio

The standard mortality ratio (SMR) was used to measure CRC-specific mortality. Race/ethnicity-specific expected deaths were calculated for each county using population data extracted from SEER. Expected deaths were aggregated by county, race, and age for indirect age-standardization. Age groups mirrored Georgia Department of Public Health's CRC age categorization (1). The categories were defined as ages 20-49, 50-64, 65-74, and 75+.

$$Expected\ deaths\ for\ race_j = \sum_{i=1}^{4} pop_{ij} * Mortality\ Rate_{ij} * Mortality_{ij}$$

where:
i = age group,
j = race/ethnicity,
$Population_{ij}$ = count of individuals of race j, age i, for the diagnosis period 2005-2011

The crude SMR was then calculated by race and county, and indirectly adjusted for age, using this equation:

$$Standard\ Mortality\ Ratio_{jk} = \frac{Observed\ Deaths_{jk}}{Expected\ Deaths_{jk}}$$

where:
j = race/ethnicity,
k = county

Data Analysis

Descriptive statistics of all covariates and outcome in the overall population were obtained using the TABLES function in R.

Spatial Smoothing

To answer the first aim of describing the spatial heterogeneity of CRC mortality, a descriptive iterative borrowing technique was used to stabilize the SMR's of each county and

account for instability due to small numbers of CRC deaths. Under the assumption that near neighbors tend to be more similar than far, two borrowing techniques were implemented: different neighbor and common neighbor. The first smoothing approach, different neighbor, iteratively added neighboring county observed and expected CRC deaths to the county of interest until a minimum threshold of 30 observed deaths was met. Neighbors were added based on distances of county centroids, with the nearest neighboring county added first, and then sequentially from the next nearest county, until the threshold is met. The threshold of 30 observed deaths was met separately for NHW and NHB. The second smoothing approach, common neighbor, followed the same steps as above, but instead the county threshold of 30 observed deaths was met for both NHW and NHB groups concurrently.

Statistical Analysis

A conditional auto-regressive (CAR) Bayesian model was used to investigate area-level poverty and education as drivers of spatial heterogeneity of CRC mortality, separately. The CAR is a common prior used for spatial disease mapping, which suggests that the value of a given area can be estimated conditional on the level of the neighboring values (54). Bayesian disease mapping methods allow researchers to encode prior assumptions that may be unstated in a Frequentist approach. Unlike other traditional methods, the CAR model allows the estimation of disease by borrowing statistical information from neighbors to address instability due to small CRC mortality counts and accounts for spatial dependence of observations (54).

Leroux Model

The Leroux Model contains a single spatial random effect, ψ_i, that adjusts for the strength of the local neighborhood spatial autocorrelation constant, ρ.

A hierarchical model fit using the Markov chain Monte Carlo (MCMC) methods was used to obtain the posterior median estimates. There were 30,000 iterations discarded for MCMC burn-in, and then 30,000 additional samples were completed, with every 30th iteration kept, to determine the posterior distribution. A binary spatial weights matrix under the Queen contiguity neighbor definition was created.

The Leroux hierarchical model can be written as:

Likelihood:

$$Y_i|\theta_i \sim Poisson(E_i\theta_i)$$

where: Y_i = counts of death for county$_i$
E_i = expected death counts
θ_i = the multiplicative (relative) excess risk for county$_i$

$$\psi_i = \log(\theta_i)$$
$$Y_i|\beta, \psi_i \sim Poisson(E_i \exp(x_i\beta + \psi_i)$$

where: Y_i = counts of death for county$_i$
β = fixed effect for covariates
ψ_i = log(θ_i), spatial random effect for county$_i$
E_i = expected death counts

Prior definitions:

$$\beta \sim N(0, 100000)$$

where: β = covariate fixed effect

$$\psi_k|\boldsymbol{\psi}_{-k}, W, \tau^2, \rho \sim N\left(\frac{\rho\sum_{i=1}^{K} w_{ki}\psi_i}{\rho\sum_{i=1}^{K} w_{ki} + 1 - \rho}, \frac{\tau_u^2}{\rho\sum_{i=1}^{K} w_{ki} + 1 - \rho}\right)$$

where: k = county of interest
W = spatial weights matrix
τ^2 = variance
ρ = spatial autocorrelation constant

Hyperpriors definitions:

$$\tau^2 \sim Inverse - Gamma(1, 0.01)$$
$$\rho \sim Uniform(0, 1)$$

The three models mentioned above were run using the Leroux method to assess the association of the two covariates, area-level poverty and area-level education, and CRC-specific

mortality. The first model did not contain any covariates. The second model included area-level poverty. The third model included area-level education. The same three models were run a second time using the same methods, but for NHW and NHB, separately. The fitted values were then extracted from the models and mapped.

To compare models, the deviance information criterion (DIC) was used, with smaller values indicating a better model fit (55).

2.4 Results

In the final analysis of 2,622 individuals with CRC-specific mortality in Georgia, 1811 (69.1%) identified as Non-Hispanic White, and 811 (30.9%) identified as Non-Hispanic Black. Additionally, 1387 (52.9%) identified as male and 1235 (47.1%) identified as female. The mean age at diagnosis among NHW and NHB was 68.9 and 64.2, respectively. A greater percentage of NHB were diagnosed between 20-64 years old, whereas a greater percentage of NHW were diagnosed between 65-75+.

Crude Race-specific Age-adjusted Standard Mortality Ratios by County

After aggregating to the county level, 12 counties observed no NHW CRC-specific deaths, and 38 counties observed no NHB CRC-specific deaths. Calhoun and Taliaferro County both observed no NHW and NHB deaths within the inclusion requirements. Banks County had the highest SMR (10.8) amongst NHB (Figure 1). Madison County had the highest SMR (3.63) among NHW (Figure 1). There was more heterogeneity in CRC mortality among NHB across counties (SMR IQR: 0.80, 1.69) compared to NHW (0.79, 1.43) (Table 2).

NHB in Rural counties also observed greater CRC deaths than expected (1.30) compared to NHW (1.12) in Rural counties (Table 2). In non-Rural counties, SMR for NHW and NHB were similar (0.98 versus 0.91, respectively).

Descriptive Smoothing – Different Neighbor Approach

Median SMR for NHW overall using the different neighbor approach was 1.06 (IQR: 0.90, 1.25), whereas the median SMR for NHB overall was 1.03 (0.89, 1.18) (Table 2). Table 2 also shows NHB SMR deaths were smoothed more towards the null value of 1.0 compared to NHW. The map in Figure 2 suggests there is slight clustering of excess NHW deaths in the Northeastern and Southeastern corners of Georgia. Whereas spatial autocorrelation of NHB CRC excess deaths were more prominent in northeastern and southwestern corners of Georgia (Figure 4).

There were more NHW observed deaths than expected in Rural counties (Median SMR = 1.06; IQR 0.92, 1.27) compared to non-Rural counties (Median SMR = 0.99; IQR 0.85, 1.17) (Table 2). Similarly, NHB living in non-Rural counties (Median SMR = 0.90; IQR 0.81, 1.04) experienced less observed deaths than expected compared to Rural counties (Median SMR = 1.07; IQR 0.96, 1.22) (Table 2).

Descriptive Smoothing – Common Neighbor Approach

The map in Figure 3 suggests the mortality ratios were smoothed more towards the null among NHW compared to the different neighbor approach. Median SMR for NHW overall using the common neighbor approach was 1.06 (IQR 0.91, 1.14), whereas the median SMR for NHB overall was 1.03 (IQR 0.89, 1.16) (Table 2). Similarly, to the different neighbor approach,

clustering of excess NHW deaths in the Northeastern and Southeastern corners of Georgia (Figure 3). Clustering of NHB excess deaths using the common neighbor approach is almost exactly similar to the different neighbor approach (Figure 2, Figure 3).

NHW SMR among Rural counties (Median SMR = 1.05; IQR 0.91, 1.16) were higher than non-Rural counties (Median SMR = 0.94; IQR 0.86, 1.10) (Table 2). Similarly, NHB in Rural counties (Median SMR = 1.07; IQR 0.96, 1.21) exhibited more excess deaths than non-Rural counties (Median SMR = 0.89; IQR 0.81, 1.02) (Table 2).

CAR Bayes Leroux Model – Non-Hispanic Whites

The model fits along with the relative risk estimates are illustrated in Table 3. Using the Leroux Model, NHW CRC-specific deaths showed low spatial autocorrelation across counties ($\boldsymbol{\rho}$ = 0.12 – 0.15) (Table 3). For every 10% increase in area-level poverty, we observed a 3% increased risk of CRC-specific mortality among NHW (95% CI 0.92, 1.15) (Table 3). However, area-level education appears to be associated with CRC-specific deaths among NHW (RR = 1.18; 95% CI, 1.04, 1.30) (Table 3). Both area-level education and area-level poverty did not explain the spatial heterogeneity in CRC-mortality, as the DIC model fit statistic increases as covariates are added into the model (Table 3). Additionally, NHW SMRs remained largely similar across counties when covariates were added to the model (Figure 4).

CAR Bayes Leroux Model – Non-Hispanic Blacks

Using the Leroux Model, NHB CRC-specific deaths showed moderate spatial autocorrelation ($\boldsymbol{\rho}$ = 0.41 – 0.45) (Table 3). There is little spatial heterogeneity of CRC-mortality among NHB using the CAR Bayes smoothing technique (Median SMR = 1.01; IQR, 1.00, 1.01)

(Table 2, Figure 5). We observed modest associations between area-level poverty or area-level education, and CRC-specific deaths among NHB. For every 10% increase in area-level poverty, risk of CRC-specific mortality increased by 12% among NHB (95% CI 0.99, 1.26) (Table 3). For every 10% increase in area-level education appears, risk of CRC-specific deaths increased by 11% among NHB (RR = 1.11; 95% CI, 0.98, 1.24) (Table 3). Neither area-level education nor area-level poverty explained the spatial heterogeneity in CRC-mortality, as the DIC the model fit statistic remained relatively consistent as covariates were added into the model (Table 3). In fact, spatial heterogeneity increased slightly as covariates were added to the model (Figure 5).

2.5 Discussion

This study investigated whether CRC mortality varied across counties in Georgia by race and whether area-level SES measures explained the variation observed. The results of the descriptive borrowing techniques indicated that there was moderate heterogeneity of CRC mortality across counties for both NHW and NHB. For example, the counties near the Atlanta metro and Southwestern Georgia regions observed fewer cancer deaths among NHW than expected, whereas counties near Southeastern Georgia observed more cancer deaths among NHW (Figure 4). For NHB, patterns of low mortality were observed in the Atlanta area and Southeastern Georgia, whereas patterns of high mortality were observed in Western Georgia (Figure 5). Interestingly, there is little variation in CRC mortality across counties among NHB using the CAR Bayesian model, as the range of county SMRs is relatively close to 1.0.

In contrast to our hypothesis that poverty and education would explain the variation in CRC mortality, our findings from the fully Bayesian model disagreed with our hypothesis and area-level education and area-level poverty may not be drivers of the spatial heterogeneity of

CRC mortality in Georgia. Our null findings of county-level poverty contrast with several previous studies that did find an association between the covariate and CRC mortality. For instance, in a retrospective cohort study using the Georgia Comprehensive Cancer Registry (N = 20,444), census tract-level SES completely explained the remaining excess risk of death associated with rural residence (HR = 1.14, 95% CI 1.07, 1.22) (17). It is important to note that our study ran a spatial analysis, whereas Hines et al. use the Cox Proportional Hazards Model and census tract-level data instead of county-level. Census tract-level data may capture more heterogeneity in SES compared to the county-level. Additionally, a previous study using a spatial regression model found that lower SES index was associated with increased CRC mortality (56). Whilst we also observed modest associations between county-level SES indicators and CRC-specific deaths, the use of a single indicator for SES may be a contributor to our inconsistent findings.

There were notable limitations to this study. The study lacked individual-level information on poverty level and educational attainment. Using county-level poverty and education as individual-level proxies may introduce an ecologic fallacy by not accurately reflecting individuals' status. The use of spatial data aggregated to a different area unit also introduces the modifiable areal unit problem (MAUP), which may cause varying estimates in correlation coefficients (57). Data were aggregated to the county level as this scale is better for intervention targeting, however, the MAUP may not have been eliminated. In addition to this, area-level factors are geographically specific, and our findings may not be generalizable to other areas outside of Georgia.

Though the study aimed to compare the lived experiences of Non-Hispanic Whites to Non-Hispanic Blacks, there is not a direct comparison of the two groups together. Therefore, there

may be a disparity in mortality between the two groups that was not captured due to the analytical approach used. Additionally, the small sample size of NHB deaths may have affected the validity of the study. The borrowing method from both the descriptive approach and the fully Bayesian model should have created more stable measures, however, small numbers could have caused over-borrowing and over-smoothing of counts.

Despite the limitations, our study provides insight that CRC mortality is not homogenous across counties in Georgia and area-level SES measures may not explain the variation seen. The study also uses a method that allows for the estimation of disease by borrowing statistical information from neighbors to address instability due to small counts. Future research should also focus on other measures, such as screening prevalence and primary care physical access, that may explain the heterogeneity to aid with identification of areas in need of public health funding and interventions.

2.6 References

1. Georgia Department of Public Health. 2015. Colorectal Cancer Factsheet 2–4.
2. American Cancer Society. 2020. Colorectal Cancer Facts and Figures 2020-2022. Am Cancer Soc.
3. Edwards BK, Ward E, Kohler BA, Eheman C, Zauber AG, Anderson RN, Jemal A, Schymura MJ, Lansdorp-Vogelaar I, Seeff LC, van Ballegooijen M, Goede SL, Ries LAG. 2010. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer 116:544–73.
4. Siegel RL, Fedewa SA, Anderson WF, Miller KD, Ma J, Rosenberg PS, Jemal A. 2017. Colorectal Cancer Incidence Patterns in the United States, 1974-2013. J Natl Cancer Inst 109:27–32.
5. Abdelsattar ZM, Wong SL, Regenbogen SE, Jomaa DM, Hardiman KM, Hendren S. 2016. Colorectal cancer outcomes and treatment patterns in patients too young for average-risk screening. Cancer 122:929–34.
6. Samadder NJ, Curtin K, Tuohy TMF, Pappas L, Boucher K, Provenzale D, Rowe KG, Mineau GP, Smith K, Pimentel R, Kirchhoff AC, Burt RW. 2014. Characteristics of missed or interval colorectal cancer and patient survival: A population-based study. Gastroenterology 146:950–960.
7. Wu X, Cokkinides V, Chen VW, Nadel M, Ren Y, Martin J, Ellison GL. 2006. Associations of subsite-specific colorectal cancer incidence rates and stage of disease at diagnosis with county-level poverty, by race and sex. Cancer 107:1121–1127.
8. Wong R. 2010. Proximal tumors are associated with greater mortality in colon cancer. J Gen Intern Med 25:1157–1163.
9. Wong RJ. 2009. Marked Variations in Colon Cancer Epidemiology: Sex-specific and Race/Ethnicity-specific Disparities. Gastroenterol Res 2:268–276.
10. Yang J, Du XL, Li ST, Wang BY, Wu YY, Chen ZL, Lv M, Shen YW, Wang X, Dong DF, Li D, Wang F, Li EX, Yi M, Yang J. 2016. Characteristics of differently located colorectal cancers support proximal and distal classification: A population-based study of 57,847 patients. PLoS One 11:1–12.
11. Li FY, Lai M De. 2009. Colorectal cancer, one entity or three. J Zhejiang Univ Sci B 10:219–229.
12. Sanz-Pamplona R, Cordero D, Berenguer A, Lejbkowicz F, Rennert H, Salazar R, Biondo S, Sanjuan X, Pujana MA, Rozek LS, Giordano TJ, Ben-Izhak O, Cohen HI, Trougouboff P, Bejhar J, Sova Y, Rennert G, Gruber SB, Moreno V. 2011. Gene expression differences between colon and rectum tumors. Clin Cancer Res 17:7303–12.
13. Miyakura Y, Sugano K, Konishi F, Ichikawa A, Maekawa M, Shitoh K, Igarashi S, Kotake K, Koyama Y, Nagai H. 2001. Extensive methylation of hMLH1 promoter region predominates in proximal colon cancer with microsatellite instability. Gastroenterology 121:1300–1309.
14. Frattini M, Balestra D, Suardi S, Oggionni M, Alberici P, Radice P, Costa A, Daidone MG, Leo E, Pilotti S, Bertario L, Pierotti MA. 2004. Different genetic features associated with colon and rectal carcinogenesis. Clin Cancer Res 10:4015–4021.
15. Rogers CR, Moore JX, Qeadan F, Gu LY, Huntington MS, Holowatyj AN. 2020. Examining factors underlying geographic disparities in early-onset colorectal cancer

survival among men in the United States. Am J Cancer Res 10:1592–1607.
16. Kuo T-M, Meyer AM, Baggett CD, Olshan AF. 2019. Examining determinants of geographic variation in colorectal cancer mortality in North Carolina: A spatial analysis approach. Cancer Epidemiol 59:8–14.
17. Hines RB, Markossian TW, Johnson A, Dong F, Bayakly R. 2014. Geographic residency status and census tract socioeconomic status as determinants of colorectal cancer outcomes. Am J Public Health 104:63–71.
18. Cole AM, Jackson JE, Doescher M. 2013. Colorectal cancer screening disparities for rural minorities in the United States. J Prim Care Community Heal 4:106–111.
19. Baldwin L-M, Cai Y, Larson EH, Dobie SA, Wright GE, Goodman DC, Matthews B, Hart LG. 2008. Access to cancer services for rural colorectal cancer patients. J Rural Heal 24:390–399.
20. Hines RB, Markossian TW. 2012. Differences in Late-Stage Diagnosis, Treatment, and Colorectal Cancer-Related Death Between Rural and Urban African Americans and Whites in Georgia. J Rural Heal 28:296–305.
21. Singh GK, Jemal A. 2017. Socioeconomic and Racial/Ethnic Disparities in Cancer Mortality, Incidence, and Survival in the United States, 1950-2014: Over Six Decades of Changing Patterns and Widening Inequalities. J Environ Public Health 2017.
22. Ward E, Jemal A, Cokkinides V, Singh GK, Cardinez C, Ghafoor A, Thun MJ. 2010. Cancer disparities by race/ethnicity and socioeconomic status. CA Cancer J Clin 54:78–93.
23. Peppercorn JM, Smith TJ, Helft PR, DeBono DJ, Berry SR, Wollins DS, Hayes DM, Von Roenn JH, Schnipper LE. 2011. American society of clinical oncology statement: Toward individualized care for patients with advanced cancer. J Clin Oncol 29:755–760.
24. Singh GK, Miller BA, Benjamin Hankey DF, Brenda Edwards SK. 1975. SEER Area Socioeconomic Variations and Cancer, 1975–1999. NCI Cancer Surveill Monogr Ser 1975–1999.
25. Carethers JM, Doubeni CA. 2020. Causes of Socioeconomic Disparities in Colorectal Cancer and Intervention Framework and Strategies. Gastroenterology 158:354–367.
26. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal A. 2020. Colorectal cancer statistics, 2020. CA Cancer J Clin 70:145–164.
27. Holowatyj AN, Ruterbusch JJ, Rozek LS, Cote ML, Stoffel EM. 2016. Racial/ethnic disparities in survival among patients with young-onset colorectal cancer. J Clin Oncol 34:2148–2156.
28. Augustus GJ, Ellis NA. 2018. Colorectal Cancer Disparity in African Americans: Risk Factors and Carcinogenic Mechanisms. Am J Pathol.
29. Diez Roux A V., Mair C. 2010. Neighborhoods and health. Ann N Y Acad Sci 1186:125–145.
30. Macintyre S, Ellaway A, Cummins S. 2002. Place effects on health: How can we conceptualise, operationalise and measure them? Soc Sci Med 55:125–139.
31. O'Campo P. 2003. Invited Commentary: Advancing Theory and Methods for Multilevel Models of Residential Neighborhoods and Health. Am J Epidemiol 157:9–13.
32. Gomez SL, Shariff-Marco S, DeRouen M, Keegan THM, Yen IH, Mujahid M, Satariano WA, Glaser SL. 2015. The impact of neighborhood social and built environment factors across the cancer continuum: Current research, methodological considerations, and future

directions. Cancer 121:2314–30.
33. Dutko P, Ploeg M Ver, Farrigan T. 2012. Characteristics and Influential Factors of Food Deserts.
34. Fong AJ, Lafaro K, Ituarte PHG, Fong Y. 2021. Association of Living in Urban Food Deserts with Mortality from Breast and Colorectal Cancer. Ann Surg Oncol 28:1311–1319.
35. Schootman M, Jeffe DB, Baker EA, Walker MS. 2006. Effect of area poverty rate on cancer screening across US communities. J Epidemiol Community Health 60:202–207.
36. Charlton ME, Matthews KA, Gaglioti A, Bay C, McDowell BD, Ward MM, Levy BT. 2016. Is Travel Time to Colonoscopy Associated With Late-Stage Colorectal Cancer Among Medicare Beneficiaries in Iowa? J Rural Heal 32:363–373.
37. Solomon I, Davis V, McNamara C, Bayakly A, Moon T. 2014. Colorectal Cancer in Georgia, 2007-2011. Georg Dep Public Heal Heal Prot Off Chronic Dis Heal Behav Inj Epidemiol.
38. Jemal A, Ward EM, Johnson CJ, Cronin KA, Ma J, Ryerson AB, Mariotto A, Lake AJ, Wilson R, Sherman RL, Anderson RN, Henley SJ, Kohler BA, Penberthy L, Feuer EJ, Weir HK. 2017. Annual Report to the Nation on the Status of Cancer, 1975–2014, Featuring Survival. JNCI J Natl Cancer Inst 109.
39. Murphy TK, Calle EE, Rodriguez C, Kahn HS, Thun MJ. 2000. Body Mass Index and Colon Cancer Mortality in a Large Prospective Study. Am J Epidemiol 152:847–854.
40. Tawk R, Abner A, Ashford A, Brown CP. 2016. Differences in Colorectal Cancer Outcomes by Race and Insurance. Int J Environ Res Public Health 13.
41. Steinbrecher A, Fish K, Clarke CA, West DW, Gomez SL, Cheng I. 2012. Examining the Association Between Socioeconomic Status and Invasive Colorectal Cancer Incidence and Mortality in California. Cancer Epidemiol Biomarkers Prev 21:1814–1822.
42. Gomez SL, O'Malley CD, Stroup A, Shema SJ, Satariano WA. 2007. Longitudinal, population-based study of racial/ethnic differences in colorectal cancer survival: Impact of neighborhood socioeconomic status, treatment and comorbidity. BMC Cancer 7:193.
43. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, Van Ballegooijen M, Zauber AG, Jemal A. 2012. Contribution of screening and survival differences to racial disparities in colorectal cancer rates. Cancer Epidemiol Biomarkers Prev 21:728–736.
44. U.S. Census Bureau.
45. US Census Bureau Public Information Office. Majority of African Americans Live in 10 States; New York City and Chicago Are Cities With Largest Black Populations - Census 2000 - Newsroom - U.S. Census Bureau.
46. Atlanta Georgia income inequality: Atlanta worst in US by Bloomberg.
47. Laiyemo AO, Doubeni CA, Pinsky PF, Doria-Rose VP, Bresalier R, Lamerato LE, Crawford ED, Kvale P, Fouad M, Hickey T, Riley T, Weissfeld J, Schoen RE, Marcus PM, Prorok PC, Berg CD. 2010. Iace and colorectal cancer disparities: Health-care utilization vs different cancer susceptibilities. J Natl Cancer Inst 102:538–546.
48. Wagner SE, Hurley DM, Hébert JR, McNamara C, Bayakly AR, Vena JE. 2012. Cancer mortality-to-incidence ratios in Georgia: Describing racial cancer disparities and potential geographic determinants. Cancer 118:4032–4045.
49. Zhou Y, Bemanian A, Beyer KMM. 2017. Housing Discrimination, Residential Racial Segregation, and Colorectal Cancer Survival in Southeastern Wisconsin. Cancer Epidemiol Prev Biomarkers 26:561–568.

50. Zahnd WE, Josey MJ, Schootman M, Eberth JM. 2021. Spatial accessibility to colonoscopy and its role in predicting late-stage colorectal cancer. Health Serv Res 56:73.
51. U.S. Census Bureau. American Community Survey (ACS).
52. Young JJ, Roffers S, Ries L, Fritz A, Hurlbut A. 2001. Summary Staging Manual 2000: Codes and Coding Instructions. Natl Cancer Institute, NIH Pub.
53. SORH Maps of Georgia | Georgia Department of Community Health.
54. Leroux BG, Leit X, Breslowt N. 2022. ESTIMATION OF DISEASE RATES IN SMALL AREAS : A NEW MIXED MODEL FOR SPATIAL DEPENDENCE * 179–191.
55. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. 2002. Bayesian measures of model complexity and fit. J R Stat Soc Ser B (Statistical Methodol 64:583–639.
56. Thatcher EJ, Camacho F, Anderson RT, Li L, Cohn WF, DeGuzman PB, Porter KJ, Zoellner JM. 2021. Spatial analysis of colorectal cancer outcomes and socioeconomic factors in Virginia. BMC Public Health 21.
57. Gotway CA, Young LJ. 2002. Combining Incompatible Spatial Data. J Am Stat Assoc 97:632–648.

2.7 Tables and Figures

Table 1. Characteristics of observed CRC-specific deaths in Georgia, 2005-2011 (N = 2622)

Sample characteristics	Non-Hispanic White n = 1811 (69.1%)		Non-Hispanic Black n = 811 (30.9%)		Total n = 2622	
	n	%	n	%	n	%
Anatomic location of cancer						
Colon	1347	74.4	638	78.7	1985	75.7
Rectum	464	25.6	173	21.3	637	24.3
Sex						
Female	835	46.1	400	49.3	1235	47.1
Male	976	53.9	411	50.7	1387	52.9
Age at diagnosis, years (mean) (SD)	68.9	13.6	64.2	15.0	67.4	14.2
Age group						
20-49	158	8.7	132	16.3	290	11.1
50-64	501	27.7	293	36.1	794	30.3
65-74	458	25.3	162	20.0	620	23.6
75+	694	38.3	224	27.6	918	35.0
Tumor stage						
Localized	505	27.9	248	30.6	753	28.7
Regional	1306	72.1	563	69.4	1869	71.3

Table 2. Race-stratified colorectal cancer SMRs in Georgia counties according to urban or rural designation

	Non-Rural Counties (N = 41) Median SMR (IQR)	Rural Counties (N = 118) Median SMR (IQR)
Non-Hispanic White		
Crude	0.98 (0.81, 1.28)	1.12 (0.77, 1.56)
Iterative smoothing: different neighbors	0.99 (0.85, 1.17)	1.06 (0.92, 1.27)
Iterative smoothing: common neighbors	0.94 (0.86, 1.10)	1.05 (0.91, 1.16)
CARBayes smoothing		
Leroux: intercept only	1.02 (0.93, 1.15)	1.06 (1.00, 1.15)
Leroux covariate adjusted: poverty	1.02 (0.92, 1.16)	1.08 (1.02, 1.17)
Leroux covariate adjusted: education	1.01 (0.90, 1.09)	1.13 (1.07, 1.23)
Non-Hispanic Black		
Crude	0.91 (0.75, 1.10)	1.30 (0.84, 1.84)
Iterative smoothing: different neighbors	0.90 (0.81, 1.04)	1.07 (0.96, 1.22)
Iterative smoothing: common neighbors	0.89 (0.81, 1.02)	1.07 (0.96, 1.21)
CARBayes smoothing		
Leroux: intercept only	1.00 (0.99, 1.01)	1.01 (1.00, 1.01)
Leroux covariate adjusted: poverty	0.96 (0.93, 0.99)	1.03 (1.00, 1.08)
Leroux covariate adjusted: education	0.99 (0.95, 1.03)	1.08 (1.04, 1.11)

Table 2. Continued…

	Georgia Overall (N = 159) Median SMR (IQR)
Non-Hispanic White	
Crude	1.09 (0.79, 1.43)
Iterative smoothing: different neighbors	1.06 (0.90, 1.25)
Iterative smoothing: common neighbors	1.01 (0.91, 1.14)
CARBayes smoothing	
Leroux: intercept only	1.05 (0.99, 1.15)
Leroux covariate adjusted: poverty	1.08 (1.00, 1.16)
Leroux covariate adjusted: education	1.11 (1.04, 1.21)

Non-Hispanic Black		
Crude	1.10	(0.80, 1.69)
Iterative smoothing: different neighbors	1.03	(0.89, 1.18)
Iterative smoothing: common neighbors	1.03	(0.89, 1.16)
CARBayes smoothing		
Leroux: intercept only	1.01	(1.00, 1.01)
Leroux covariate adjusted: poverty	1.02	(0.97, 1.07)
Leroux covariate adjusted: education	1.06	(1.02, 1.10)

Table 3. CAR-Bayesian Leroux model fitting results and relative risk of CRC mortality

	Effect(s)			
	Posterior Median	95% Bayesian Credible Intervals	10% RR	95% RR CI
Non-Hispanic White				
Leroux: intercept only	0.0503	(-0.0107, 0.1070)	1.05	(0.99, 1.11)
Leroux covariate adjusted: poverty	0.0028	(-0.0084, 0.0144)	1.03	(0.92, 1.15)
Leroux covariate adjusted: education	0.0162	(0.0041, 0.0262)	1.18	(1.04, 1.30)
Non-Hispanic Black				
Leroux: intercept only	0.0030	(-0.0672, 0.0770)	1.00	(0.93, 1.08)
Leroux covariate adjusted: poverty	0.0111	(-0.0014, 0.0228)	1.12	(0.99, 1.26)
Leroux covariate adjusted: education	0.0101	(-0.0015, 0.0217)	1.11	(0.98, 1.24)

Table 3. Continued…

	DIC	ρ
Non-Hispanic White		
Leroux: intercept only	799.43	0.1363
Leroux covariate adjusted: poverty	800.11	0.1221
Leroux covariate adjusted: education	802.66	0.1514
Non-Hispanic Black		
Leroux: intercept only	538.84	0.4145
Leroux covariate adjusted: poverty	538.12	0.4455
Leroux covariate adjusted: education	538.37	0.4178

Figure 1. Race-stratified crude colorectal cancer age-adjusted standard mortality ratio across counties in Georgia. Data was obtained from SEER for individuals diagnosed in Georgia 2005-2011

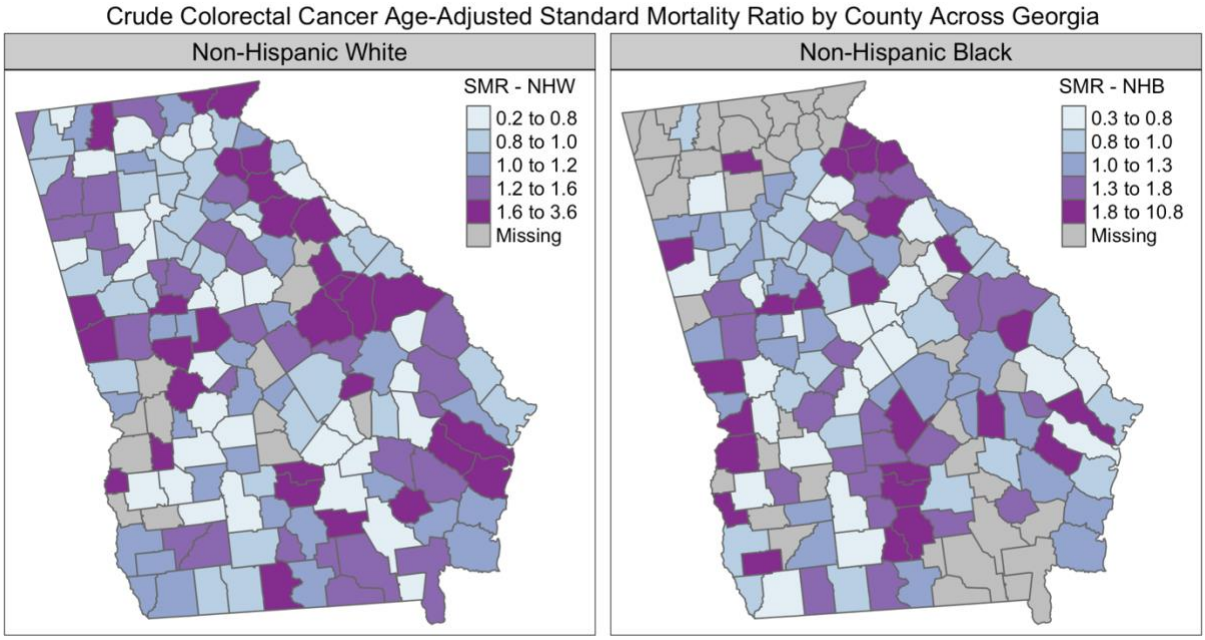


Figure 2. Colorectal cancer standard mortality ratio by county in Georgia using the different neighbor approach

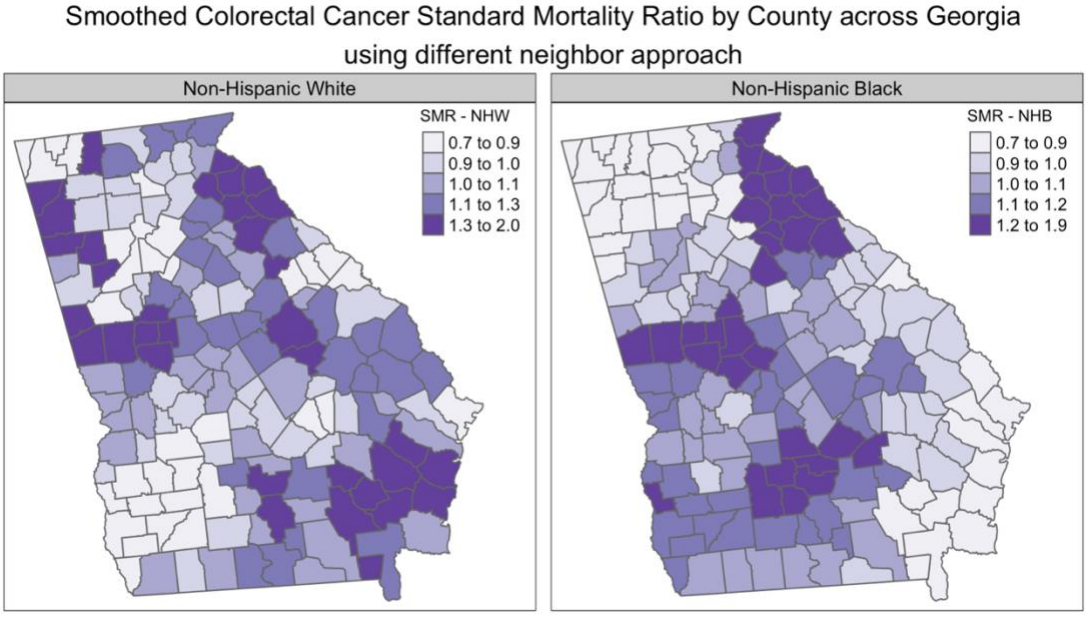


Figure 3. Colorectal cancer standard mortality ratio by county in Georgia using the common neighbor approach

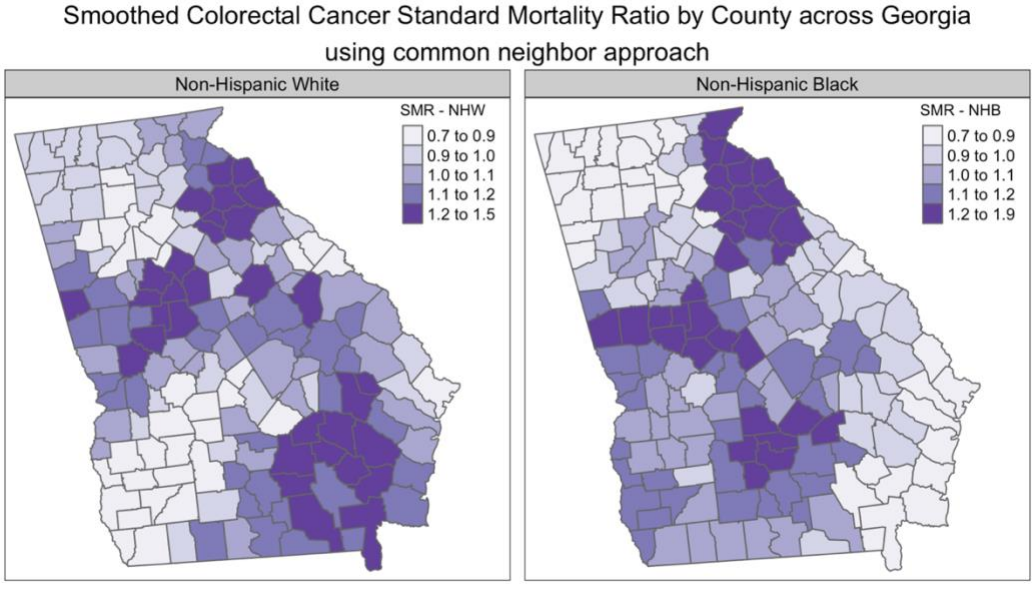


Figure 4. Leroux Model fitted NHB standard mortality ratio across counties in Georgia. A binary spatial weights matrix using the Queen contiguity neighbor definition was used

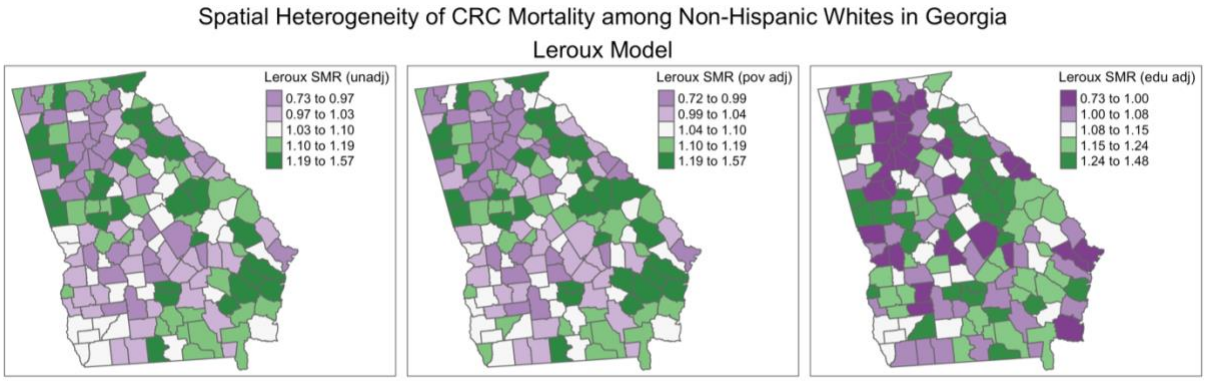
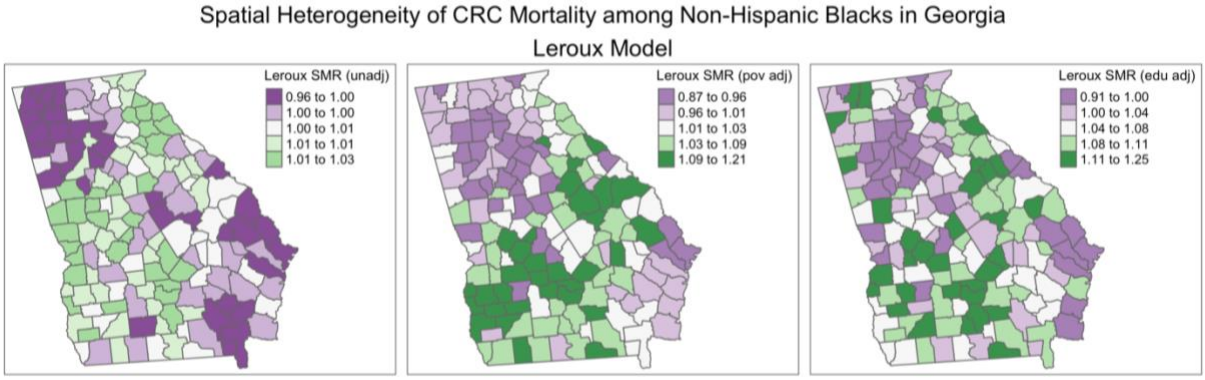


Figure 5. Leroux Model fitted NHB standard mortality ratio across counties in Georgia. A binary spatial weights matrix using the Queen contiguity neighbor definition was used



CHAPTER III: SUMMARY, PUBLIC HEALTH IMPLICATIONS, POSSIBLE FUTURE DIRECTIONS

The goal of this thesis was to characterize the variation in CRC mortality across counties in Georgia by race. Additionally, it was to explore if area-level poverty and education were drivers in the variation seen. The study showed moderate heterogeneity of CRC mortality in Georgia by both NHW and NHB. Although area-level poverty and education did not influence heterogeneity, the findings from this study still provide valuable information on where CRC mortality prevalence was greatest or lowest in Georgia.

After characterizing CRC mortality in Georgia, the results of this thesis also indicated that Rural counties in Georgia experience slightly greater mortality compared to non-Rural counties. Additional research is needed to identify if area-level measures influence CRC mortality by rurality. Understanding these mechanisms could further identify areas in need of public health funding and intervention.

Epidemiologic cartography and other GIS approaches map raw data to visualize health statistics. It is a necessary tool for descriptive epidemiology, but it is not as sufficient as disease mapping and spatial analysis. Instead, disease mapping is driven by core epidemiologic questions and focuses on fixing fundamental epidemiologic and statistical problems. Disease mapping is a tool that can characterize the distribution of health within a geographic area, whereas spatial analysis allows researchers to estimate the determinants of health. These tools are increasingly being incorporated within population health and health surveillance studies.

This thesis focused mainly on disease mapping, where we looked at different techniques that would give the best estimate of spatial heterogeneity in disease intensity and where specifically intensity was higher or lower. Future research could explore alternative ways to answer this research idea. For instance, spatial cluster analysis could answer the following

questions: Is there significant clustering of CRC mortality, and where does the significant clustering exist? Another approach could be a spatial regression analysis to answer the following question: Is area-level poverty or education associated with CRC mortality after adjusting for spatial correlation?

This thesis also highlights the need for further research to examine other neighborhood characteristics that may be drivers in CRC mortality heterogeneity. While studies have looked at spatial access to colonoscopy and late-stage CRC diagnosis, a future approach can incorporate how the accessibility to screening services is related to CRC mortality (50). Furthermore, other research has examined SES as an index instead of single indicator, making it more difficult to identify effective interventions. Additional research should focus on other indicators used in SES indices, such as the percentage of insured or unemployed persons.

In conclusion, the findings from this thesis characterized the variation in CRC mortality across counties in Georgia and by race. The descriptive borrowing techniques indicated that counties near Southeastern Georgia observed more cancer deaths among NHW persons, and patterns of high mortality were observed in Western Georgia among NHB persons. The findings also exploit remaining questions about whether socioeconomic status explains the heterogeneity seen in CRC mortality. Future research is needed to understand other mechanisms that may be drivers of variation in CRC mortality. By identifying these drivers and areas with greater disparity, funding, and interventions could be focused on increasing equity in areas of need.

4.0 Appendices

4.1 Appendix A

Literature Review

Author, Year	PMID	Population			Research Question	Assessment
		N	Race/Ethnicity	Study Period		Independent Var
Yang J (2016)	27936129	57,847	NA	2000-2012	Identify the most useful method for grouping colorectal cancer by tumor location according to both baseline and survival characteristics	Primary tumor site
Hines RB (2012)	22757954	15,174	White, Black	1992-2007	What are the CRC outcomes for a sample of residents of the state of Georgia according to geographic residence (rural vs urban) and race?	race, ethnicity, county of residence at time of diagnosis (rural vs urban)
Hines RB (2014)	24432920	20,444	NA	2000-2007	We examined the impact of geographic residency status and census tract (CT)-level socioeconomic status (SES) on colorectal cancer (CRC) outcomes.	Race/Hispanic ethnicity, Gender, Age at diagnosis, Date of diagnosis, First course of treatment received, Last date of follow-up, Vital status at last follow-up, Census Tract of patient's residential address, SES at CT level
Hinshaw (2021)	33394205	37,803	NA	2008-2016	What are CRC incidence and mortality across counties in eastern North Carolina, by stage, and are there racial disparities?	NA
Kruse-Diehr (2021)	33600304	252	White, Black	1999-2018	What is the relationship between racial residential segregation and CRC mortality, do these effects of segregation differ by race and rurality?	racial residential segregation, rurality

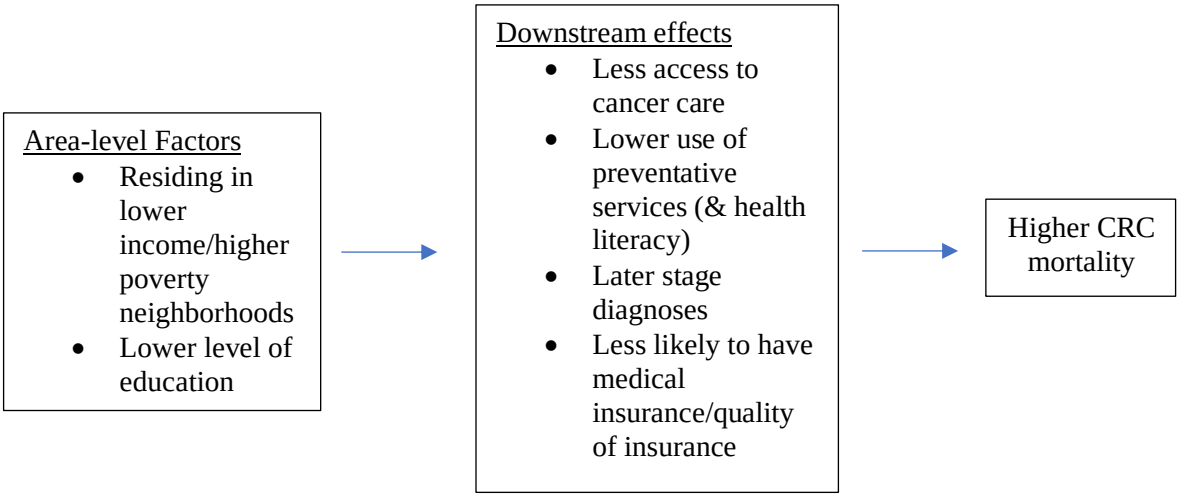
Author, Year	Assessment		Statistical analysis	Results (report main effect estimates only)	Overarching Conclusions
	Covariates	Dependent Var			
Yang J (2016)	age at diagnosis, year of diagnosis, ethnicity, sex, stage, tumor grade, mucinous histology, treatment	Disease-specific survival (DSS), time to DSS	Cox Proportional Hazards regression model, Kaplain Meier curves	Compared with LCC and ReC, RCC was significantly affected older patients (median age 75 years) and women (55.4%), to be advanced stage (stage II and above; 72.1%), and to have mucinous histology (14.9%).	LCC and ReC share characteristics but differ from RCC with regard to baseline characteristics and DSS. Reasonable to divide CRC into 3 entities.
Hines RB (2012)	gender, age at diagnosis, county level variables (median household income, % of individuals living below poverty level, % high school graduates, % of adults unemployed, % of adults with professional occupations), SES	late stage of disease at diagnosis (stage III and IV), receipt of treatment, cancer-specific mortality	chi-squared test, t-test, multilevel logit model, Cox proportional hazards model	African Americans had 40% increased odds of late-stage CRC diagnosis (OR, 1.40; 95% CI, 1.30-1.51) compared to their white counterparts. Rural/urban county-level designation was not associated with late-stage CRC diagnosis	Rural residents experienced increased risk of death due to colon cancer. African Americans have decreased odds of receiving surgery.
Hines RB (2014)	NA	late-stage disease at diagnosis, receipt of treatment (chemo, surgery), survival	Kaplan Meier method, Chi-squared statistic, odds ratios, mutilevel hierarchical models, Cox proportional hazards model	For COLON cancer: Residents of low-SES CTs also had 17% decreased odds (AOR = 0.83; 95% CI = 0.72, 0.96) of receiving chemotherapy. In model 2, compared with urban residents, rural residents had 14% higher risk of death (hazard ratio [HR] = 1.14; 95% CI = 1.07, 1.22) following diagnosis.	There was no association between late-stage disease at diagnosis and geography or SES.
Hinshaw (2021)	NA	incidence and mortality rates	Just calculated rates	Overall mortality rates were significantly higher in the hotspot (18.1, 95% CI 16.6–19.7) and Eastern NC (15.9, 95% CI 15.3–16.6) compared to Non-ENC (13.9, 95% CI 13.7–14.2) areas. By stage (localized, regional, and distant) were also higher in hotspots.	The paper says that spatial mapping identified distinct "hotspots" in certain counties in NC, however, information is from a prior study. The methods and figures don't show they conducted the spatial analysis themselves.
Kruse-Diehr (2021)	SES (e.g. low income, low education, overcrowding)	mortality	Mixed linear regression model	Urban Delta Region counties with low and high, but not moderate, levels of racial segregation had higher CRC mortality rates among Black residents but not as evident in rural counties.	Racial segregation was not significantly associated with CRC mortality among White residents in urban counties

Author, Year	PMID	Population			Research Question	Assessment
		N	Race/Ethnicity	Study Pop Period		Independent Var
Torress (2018)	30425965	1,120	NA	2000-2010	What are the geographic distributions of breast, cervical, and CRC incidence among female residents in Baltimore City, MD, and the neighborhood characteristics associated with those distributions?	cancer type, mean age at diagnosis, tumor grade, race, and street address of patient at diagnosis
Kuo (2019)	30640041	NA	NA	2003-2013	How does spatial autocorrelation work to obtain unbiased estimates for the association between CRC mortality and county-level determinants in NC?	county level (socio-demographic, access and quality of health care, behavioral risk factors, and urbanicity)
Rogers CR (2020)	32509399	32,447	NHW, NHB, Hispanic adults, Hispanic adolescents 15-49	1999-2016	What are mortality hospots specific to men with Early Onset CRC? What are the differences in individual- and county-level characteristics between EOCRC hotspots and non-hotspots?	individual and county-level determinants
Veach (2014)	25426487	3108 counties	Caucasian, African-American, Hispanic/Latino	2005-2007	What risk factors impact CRC death for each racial group?	median county income, % below poverty level, % urban, avg diabetes rate, avg obesity rate, % age pop, % race
Geyer (2020)	33221647	106 ASC locations	NA	2013-2017	What is the spatial relationship between CRC mortality and ambulatory surgery center density?	Ambulatory Surgery Center Density
Carroll (2018)	30713133	82,828	NA	1973-2013	What age-group specific survival following CRC diagnosis? How did it change over time? What are the differences in younger vs older ppl?	age(early onset, older onset)

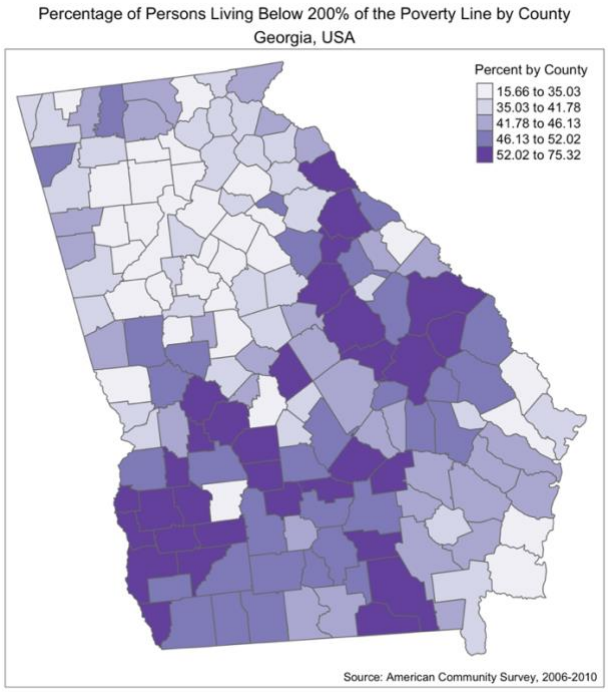
Author, Year	Assessment		Statistical analysis	Results (report main effect estimates only)	Overarching Conclusions
	Other Important Vars	Dependent Var			
Torress (2018)	% race, % racial diversity, % female-headed household, % household <$25k, % housing violation, crime, domestic violence, teen birth, % employed, % tree coverage	spatial clusters/"hot spots" of cancer incidence	hot spot analysis, ordinary least squares regression models for neighborhood-level variables	There was evidence of spatial variation in incidence of cancers. Small area estimates are needed to detect local patterns of disease. There was a relationship between colorectal cancer incidence and % African American in Baltimore City, however other neighborhood-level covariates were not significant.	NA
Kuo (2019)	NA	CRC-specific death	cluster analysis, spatial econometric models, ordinary least squares model	The average total effect of SES deprivation for each risk group: 60.78 (low), 0.76 (moderate) and 1.32 (high risk). 1 sd increase in SES deprivation associated with, on average, an increase of 61 CRC deaths per 100,000 for the low-risk group.	Negligible effect in areas where behavioral risk was moderate or high.
Rogers CR (2020)	NA	CRC specific survival probability, hazard in hotspots versus non-hotspots	Cox proportional hazards model, empirical Bayes (EB) smoothed model, LISA	Men residing in hotspots were more likely to be diagnosed with metastatic disease (stage IV CRC) compared to those residing in non-significant spots (2.58% vs 1.94%).	Hotspot counties were more likely to have higher poverty rates, greater prevalence of adult obesity, more physical inactivity, lower college completion rates, higher adult smoking rates, increased rurality.
Veach (2014)	NA	CRC mortality	Moran's I and simultaneous autoregressive model	CRC mortality rate among African Americans were positively correlated with average % population obese (0.016)	Defined education as “at least high school degree” OR otherwise.
Geyer (2020)	NA	CRC mortality	global, local, regional Moran's I	CRC mortality rates (median: 15.30 per 100,000 of the US 2000 standard million population) exhibited hot spots in rural Pennsylvania counties.	CRC mortality rates clustered in Rural PA counties. Surgery center density clusters were in Urban SE counties.
Carroll (2018)	race, marital status, cancer grade, malignant history, surgery, therapy, and county level SES related factors	incidence and survival	Bayesian Poisson Knorr-Held Model	Improved survival for individuals in counties with higher % higher education and % persons living in poverty. There is age-group specific difference in CRC incidence and survival.	They also could not adjust for race due to low non-White population

4.2 Appendix B

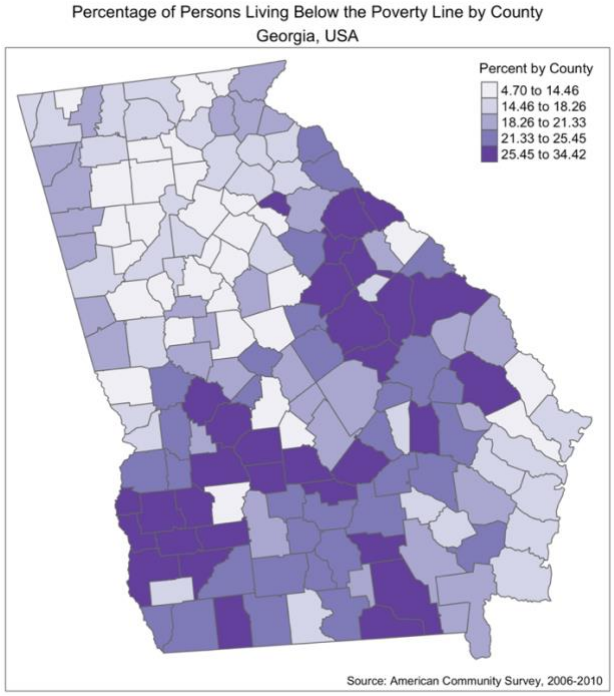
Directed acyclic graph (DAG) of the association between area-level SES (exposure) and CRC mortality (outcome)



4.3 Appendix C

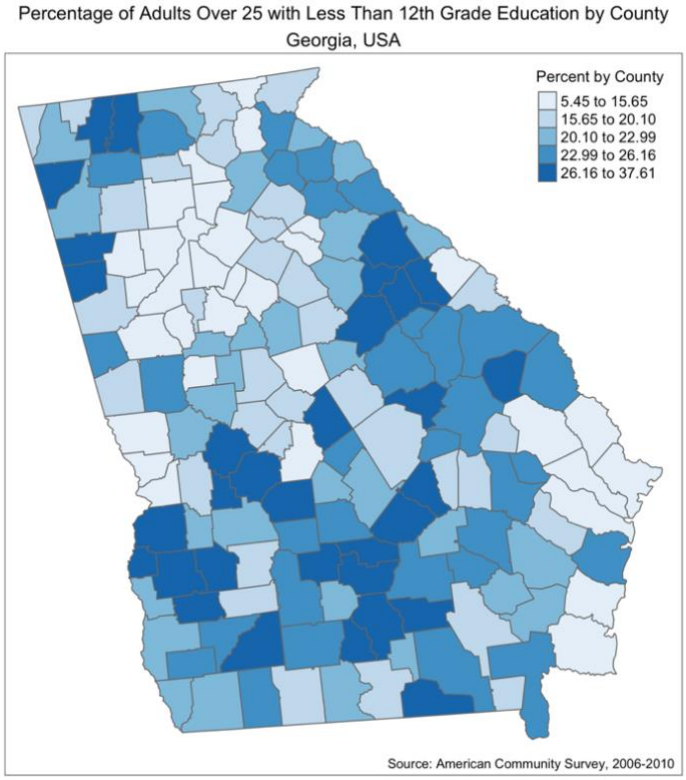


Appendix C. Percentage of persons living below 200% of the poverty line by county in Georgia. American Community Survey, 5-year estimates, 2006-2010.



Appendix C. Percentage of persons living below the poverty line by county in Georgia. American Community Survey, 5-year estimates, 2006-2010.

4.4 Appendix D



Appendix D. Percent of adults over 25 with less than 12th grade education (no HS degree or GED equivalent) by county in Georgia. American Community Survey, 5-year estimates, 2006-2010.

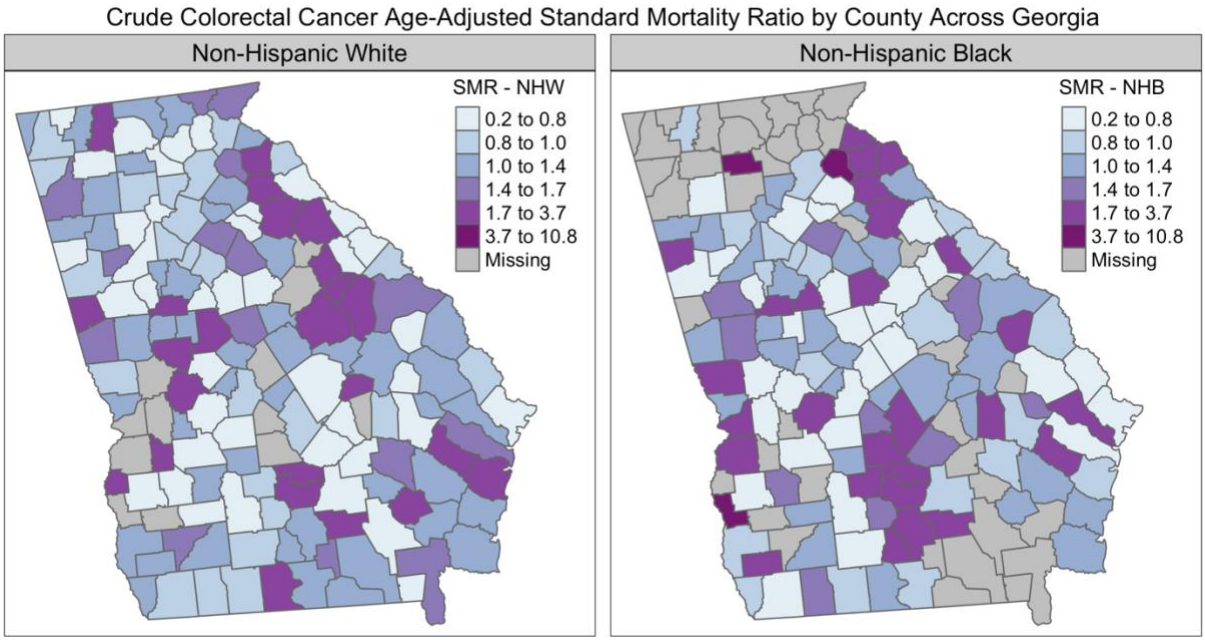
% living below poverty was calculated using this equation:

Persons below the poverty line: $\frac{(\mathrm{C17002e02} + \mathrm{C17002e03})}{C17002e01} * 100$

% educational attainment was calculated using this equation:

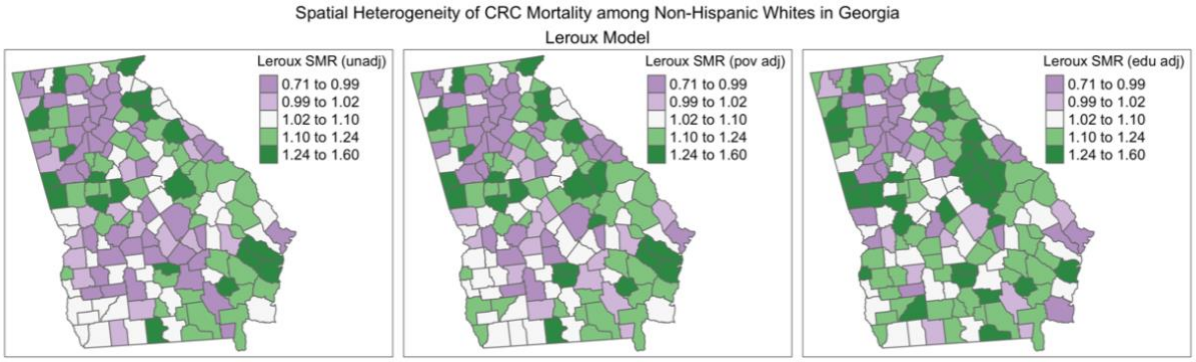
% < HS grad: $\frac{(\mathrm{B15002e03} + \ldots + \mathrm{B15002e10} + \mathrm{B15002e20} + \ldots + \mathrm{B15002e27})}{\mathrm{B15002e01}} * 100$

4.5 Appendix E



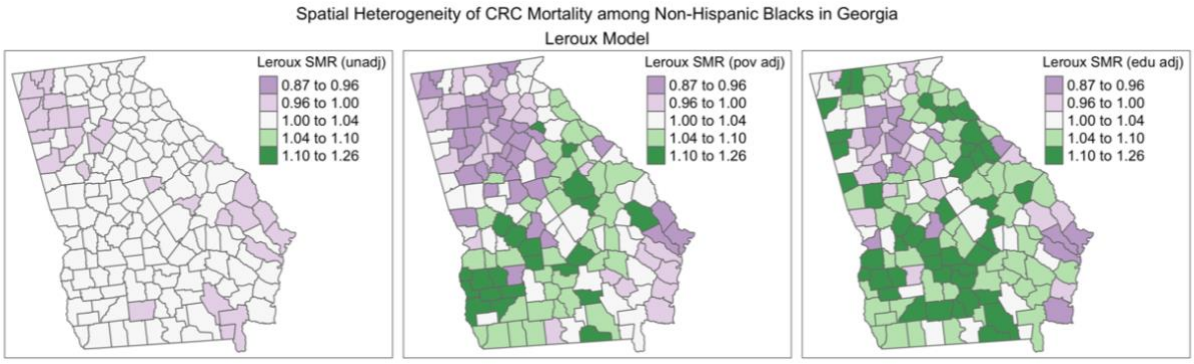
Appendix E. Race-stratified crude colorectal cancer age-adjusted standard mortality ratio across counties in Georgia. Data was obtained from SEER for individuals diagnosed in Georgia 2005-2011. Legend cut points are same in both race/ethnicity maps.

4.6 Appendix F



Appendix F. Legends with same cut points in each map. Leroux Model fitted NHW standard mortality ratio across counties in Georgia. A binary spatial weights matrix using the Queen contiguity neighbor definition was used.

4.7 Appendix G



Appendix G. Legends with same cut points in each map. Leroux Model fitted NHB standard mortality ratio across counties in Georgia. A binary spatial weights matrix using the Queen contiguity neighbor definition was used.