EVALUATING THE METHODOLOGY AND CLINICAL UTILITY OF SPATIAL ACCESS TO HEALTH CARE MEASURES IN APPALACHIA

by

Joseph J. Donohoe

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Bioinformatics)
in The University of Michigan
2015

Dissertation Committee:

Associate Professor Rajesh Balkrishnan, Chair
Professor Roger T. Anderson, University of Virginia
Associate Professor David A. Hanauer
Assistant Professor Maureen A. Sartor
Associate Professor Kai Zheng

DEDICATION

I dedicate this dissertation to my wife Whitney, for all her love and support.

ACKNOWLEDGMENTS

Many people have helped me during the course of this dissertation. I first want to thank my advisor Dr. Rajesh Balkrishnan for his terrific leadership and guidance. Graduate school is largely about exploration. Which career is the best fit? Which research direction is most appropriate? Which collaboration is most helpful? I explored as much as anyone. Dr. Balkrishnan was patient throughout, giving any support he could. He also achieved that delicate balance between structured guidance and independent work. I am grateful for the support of my committee members: Dr. Roger Anderson, Dr. David Hanauer, Dr. Maureen Sartor, and Dr. Kai Zheng. Although different parts of my dissertation were sometimes outside of their various research specialties, each committee member was more than willing to participate in any way needed. I also want to thank the Department of Computational Medicine and Bioinformatics, and particularly the graduate coordinators Dr. Margit Burmeister and Dr. Daniel Burns, for encouraging my interdisciplinary research and career interests. In completing much of my analyses, I am indebted to collaborators Vince Marshall, Fabian Camacho, and Xi Tan. Vince and Fabian are superb statisticians, data analysts, and database experts. Without their expertise, my research would have taken far longer and been less focused. Xi Tan willingly provided any clinical and pharmaceutical input that I requested.

I am blessed to have too many loving friends and family members to thank here. My parents, Dwayne and Janet, have loved me through years of degrees, travels, and detours, and I have no doubt they will continue standing by me wherever life leads. My in-laws, Ken and Sandy, welcomed me into their family and took an interest in all my academic pursuits, despite my jumbled explanations of what I was working on. Finally, and most importantly, I thank my wife Whitney. She sacrificed the most during this dissertation, yet never failed to offer her support when I needed it. This dissertation, along with so many other professional and personal accomplishments in my life, would not be possible were Whitney not the thoughtful and loving person she is.

This research was part of a larger project funded by the National Cancer Institute and the NIH Office on Women's Health through grant 1 R21 CA168479 (Balkrishnan, PI).

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGMENTS iii

LIST OF FIGURES vii

LIST OF TABLES ix

LIST OF APPENDICES x

ABSTRACT xi

CHAPTER

1. Introduction

1.1 Statement of the Problem 1
1.2 Nature of the Study 8
1.3 Study Aims 10
1.4 Significance of the Study 12
1.5 References 14

2. Evaluating and Comparing Methods for Measuring Spatial Access to Mammography Centers in Appalachia

2.1 Abstract 18
2.2 Introduction 19
2.3 Study Area and Data 23
2.4 Methodology 24
2.5 Results 28
2.6 Discussion 32
2.7 References 37

3. Spatial Access to Primary Care Providers in Appalachia: Evaluating Current Methodology 48

3.1 Abstract 48

3.2 Introduction..49
3.3 Study Area and Data 52
3.4 Methodology 53
3.5 Results 56
3.6 Discussion 60
3.7 References..65

4. Predicting Late-Stage Breast Cancer Diagnosis and Receipt of Adjuvant Therapy: Applying Current Spatial Access to Care Methods in Appalachia..74

4.1 Abstract 74
4.2 Introduction..75
4.3 Methods..76
4.4 Results 81
4.5 Discussion 84
4.6 References..88

5. Overall Conclusion...99

5.1 Research Question 99
5.2 Research Goals..99
5.3 Study Results 100
5.4 Limitations of the Study..104
5.5 Future Research Possibilities 104
5.6 Overall Study Impact 105
5.7 References..107

APPENDICES 108

LIST OF FIGURES

FIGURES

2.1. Population density across the counties of PA, OH, KY, and NC. 41

2.2. Spatial Access to Mammography centers by (a) county provider to population ratio, broken into quintiles, and (b) closest driving time from census block groups 43

2.3. Spatial access of census block groups in PA, OH, KY, and NC to mammography centers by (a) original 2SFCA scores; (b) slow-decay 2SFCA scores; (c) fast-decay 2SFCA scores; (d) slow-decay and variable catchment 2SFCA scores; (e) fast-decay and variable catchment 2SFCA scores; (f) continuous-decay and variable catchment 2SFCA scores; (g) slow-decay and variable catchment SPAR scores; (h) fast-decay and variable catchment SPAR scores. 44

2.4. Comparison of spatial accessibilty across each census block group between (a) slow-decay and fast-decay 2SFCA scores; (b) Slow-decay, variable catchment and fast-decay, variable catchment 2SFCA scores; (c) Slow-decay, variable catchment and fast-decay, variable catchment SPAR scores...45

2.5. Comparison of spatial accessibilty across rural and urban census block groups between (a-b) slow-decay and fast-decay 2SFCA scores; (c-d) Slow-decay, variable catchmentand and fast-decay, variable catchment 2SFCA scores; (e-f) Slow-decay, variable catchment and fast-decay, variable catchment SPAR scores. 46

3.1. Population densities from the 2010 U.S. Census in Appalachia counties of PA, OH, KY, and NC. 68

3.2. Spatial access to primary care providers by (a) county provider to population ratio, broken into quintiles, and (b) closest driving time from census block groups. 70

3.3. Spatial access of census block groups in Appalachia PA, OH, KY, and NC to primary care providers by (a) original 2SFCA scores; (b) slow-decay 2SFCA scores; (c) fast-decay 2SFCA scores; (d) slow-decay and variable catchment 2SFCA scores; (e) fast-decay and variable catchment 2SFCA scores; (f) continuous-decay and variable catchment 2SFCA scores; (g) slow-decay and variable catchment SPAR scores; (h) fast-decay and variable catchment SPAR scores. Scores are broken into quintiles, with larger quintiles representing greater spatial access. 71

3.4. Comparison of spatial accessibilty to primary care providers across rural and urban census block groups of Appalachia PA, OH, KY, and NC between (a-b) slow-decay and fast-decay 2SFCA scores; (c-d) Slow-decay, variable catchmentand and fast-decay, variable catchment 2SFCA scores; (e-f) Slow-decay, variable catchment and fast-decay, variable catchment SPAR scores...72

4.1. Local coefficient estimates for the logistic GWR model predicting late stage breast cancer diagnosis across variables (a) age, (b) insurance status, (c) 2SFCA score, and (d) county primary care to population ratio. 98

LIST OF TABLES

TABLES

2.1. Spearman correlations* between 10 spatial access to care measures across PA, OH, KY, and NC. 42

2.2. Descriptive statistics of mammography center spatial access scores between rural and urban census block groups of PA, OH, KY, and NC. 47

2.3. Descriptive statistics of mammography center spatial access scores between Appalachia and non-Appalachia census block groups of PA, OH, KY, and NC. 47

3.1. Spearman correlations* between 10 spatial access to prmiary care measures across Appalachia PA, OH, KY, and NC. 69

3.2. Descriptive statistics of primary care spatial access scores between rural and urban census block groups of Applachia PA, OH, KY, and NC. 73

4.1. Characteristics of study patients by breast cancer stage at diagnosis. 93

4.2. Characteristics of study patients by receipt of adjuvant hormonal therapy (AT). 94

4.3. Parameter estimates for the global logistic regression model predicting late stage breast cancer diagnosis* and receipt of adjuvant hormonal therapy^{+}. 95

4.4. Summary statistics of the parameter estimates for the logistic GWR models for late stage breast cancer diagnosis*^{a} and receipt of adjuvant hormonal therapy^{+b}. 96

4.5. Comparison of fit between global logistic regression and GWR models for late stage breast cancer diagnosis*^{a} and receipt of adjuvant hormonal therapy^{+b}. 97

LIST OF APPENDICES

APPENDIX

1: The University of Michigan IRB Approval 108

2: Data User Agreement 111

ABSTRACT

The Appalachia region of the U.S. has noted socioeconomic disparities, elevated rates and mortality for many cancers, and substandard cancer treatment patterns. These disparities, combined with Appalachia's largely rural population, suggest that the region has reduced access to health care. This research investigated the methodology behind spatial access to healthcare using population-based clinical data and geographic information systems (GIS) software. The dissertation's goal was to provide a guide of the latest spatial access methods in Appalachia and to demonstrate how those methods can be incorporated into models studying cancer disparities in the region. Accredited mammography centers and primary care physicians in 2008 from Pennsylvania, Ohio, Kentucky, and North Carolina, along with U.S. Census population data, were geocoded into GIS software. Methods compared included ratios of mammography centers and physicians to county populations, travel time to closest mammography centers and physicians, and several versions of the newer two-step floating catchment area (2SFCA) method, which has never been evaluated in Appalachia. As a measure of predictive validity, spatial access methods were then used to predict two important breast cancer clinical indicators: stage at diagnosis and receipt of adjuvant hormonal therapy after a qualifying surgery. Urban and non-Appalachia areas had consistently shorter travel times than the rural and Appalachia areas of the same states, across both travel times to mammography centers and primary care physicians. The 2SFCA measures that included variable catchment sizes appeared distinct from the original 2SFCA method and 2SFCA methods that included distance decay weights but no variable catchments. Predictors of late-stage breast cancer diagnosis included age, insurance status, county primary care to population ratio, and primary care 2SFCA score. Geographically weighted logistic regression revealed that the effect of the predictor variables varied throughout the study region. Predictors of adjuvant hormonal therapy included the presence of multiple diseases, county economic status, and mammography center 2SFCA score. Overall, the 2SFCA method with variable catchment sizes offered the greatest predictive validity of the access measures and offers theoretical improvements over the other access to care measures. Nonetheless, further research is needed to validate the 2SFCA method parameters with patient healthcare utilization data.

CHAPTER 1

Introduction

1.1 Statement of the Problem

Appalachia is a 205,000 square mile area that spans the Appalachian Mountains from southern New York to northern Mississippi. The region consists of parts of 13 states: Alabama, Georgia, Kentucky, Maryland, Mississippi, New York, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, and West Virginia. Appalachia lags behind national averages across many socioeconomic indicators (Pollard & Jacobsen, 2013). The percentage of high school and college graduates is less in Appalachia than nationally. Per capita income is less in Appalachia than nationally, and the percentage of Appalachian residents below the poverty line is also greater than national averages (Pollard & Jacobsen, 2013). Residents of Appalachia also engage in more risky health behaviors. Rates of smoking (Appalachia Regional Commission, 2008) and obesity (Herath & Brown, 2013) are higher in Appalachia than nationally.

In addition to the socioeconomic disparities, Appalachia residents suffer from an increased cancer burden. Since the early 1990s, the National Cancer Institute has recognized this burden and funded a network of community partners and academic researchers to reduce cancer health disparities in the region (Appalachia Community Cancer Network, 2013). Incidence of several cancers is higher in Appalachia than non-Appalachia parts of the same states and is higher in Appalachia than nationally. For example, Appalachia Kentucky, Ohio, and Virginia have higher incidence of cervical cancer than non-Appalachia parts of those states (Appalachia Community Cancer Network, 2009). Appalachia Kentucky, Ohio, and Pennsylvania have higher incidence of colon cancer than non-Appalachia areas of those states (Appalachia Community Cancer Network, 2009). Appalachia Kentucky, Ohio, and Virginia also have higher incidence of lung cancer than non-Appalachia parts of those states (Appalachia Community Cancer Network, 2009). The Appalachia region as a whole has a higher incidence rate of lung and colon cancer that the rest of the U.S. (Winger et al., 2003).

Cancer mortality is similar, with many cancers having higher mortality rates in Appalachia than non-Appalachia parts of the same states, and higher rates in Appalachia than nationally. The mortality rate for all cancers combined is higher in Appalachia than the rest of the U.S. (Blackley, Behringer, & Zheng, 2012). Mortality rates specifically for lung and bronchus cancer, colon cancer, and cervical cancer are also all higher in Appalachia than in non-Appalachia states (Blackley, Behringer, & Zheng, 2012). Within Ohio, the Appalachia region has increased mortality rates for lung and bronchus cancer, colon cancer, cervical cancer, and melanoma skin cancer compared to the non-Appalachia region (Fisher et al., 2008). Disparities exist even for cancers where mortality rates are similar between Appalachia and non-Appalachia areas, such as female breast cancer. Breast cancer mortality rates are declining faster in non-Appalachian counties in Appalachian states and in non-Appalachia national counties compared to Appalachian counties (Yao, Lengerich, & Hillemeier, 2012).

Consistent with these disparities, there is evidence that cancer treatment and screening is sub-optimal in Appalachia. Rates of the preferred early stage breast cancer treatment—breast conserving surgery and adjuvant radiation—are lower for Appalachian women than national rates (Freeman, Huang, & Dragun, 2012). Rates of mammograms and clinical breast exams, the standard breast cancer screening tools, are lower for Appalachian women than among other U.S. women (Hall et al., 2002). Cervical cancer screening rates are also lower for Appalachian women compared to national rates (Hall et al., 2002). Outside of Appalachia there is a much larger body of research identifying substandard cancer care and implicating factors like race and ethnicity, urban or rural residence, and socioeconomic status. Much of this research uses data from the Surveillance, Epidemiology, and End Results (SEER) program, which largely excludes Appalachia. Kentucky was added to the SEER database in 2001 and is currently the only Appalachia area represented.

Cumulatively, the increased incidence, mortality, and examples of substandard care suggest that the Appalachia region has reduced access to cancer care (Pasket et al., 2011). Access to care refers to an individual's ability to obtain the medical care they need (Gulliford et al., 2002). Both spatial and non-spatial factors can impact access to care. Spatial factors focus on the geographic relationship between patient and healthcare provider. Measures like travel time or travel distance between patient and provider are examples of spatial factors. Non-spatial factors include socioeconomic and demographic variables such as household income, educational attainment, and insurance status.

Research both within and outside of Appalachia demonstrates that reduced access to care negatively impacts cancer outcomes. In England, for example, the likelihood of receiving recommended lung cancer care decreased as travel time to treatment providers increased (Jones et al., 2008). Among black women in Detroit, late stage breast cancer diagnosis was associated with greater travel distance to mammography facilities (Dai, 2010). An example of non-spatial factors is demonstrated by research showing that lung cancer patients with Medicaid insurance were less likely to receive recommended treatments than their peers with private insurance

(Harlen et al., 2005). In Appalachia Ohio, children without a regular primary care physician had poor general health, and the parents of those children reported that access to primary care was the contributing issue (Smith & Holloman, 2011).

This thesis focuses on the methodology behind the spatial, or geographic, component of access to care, often referred to as spatial access. The simplest method for measuring spatial access focuses on the density of healthcare providers. Here, researchers calculate the ratio of healthcare providers per population count in a defined geographic boundary. For example, the Department of Health and Human Services (DHHS) defines Health Professional Shortage Areas (HPSAs) as rational service areas where the population to primary care physician ratio is 3,500:1 or higher (DHHS, 2013). Calculating travel time or travel distance between a patient and their healthcare provider is another relatively straightforward approach for measuring spatial access. Researchers input patients' addresses and healthcare provider addresses into a geographic information systems (GIS) program and the software maps the addresses as latitude and longitude coordinates, a process called geocoding. A researcher can then determine if travel time to a healthcare provider impacts an outcome of interest, such as whether or not a patient received the latest recommended chemotherapy.

Both of these approaches—population to provider ratios and travel time—have limitations. Population to provider ratios use fixed geographic boundaries, such as census tracts or zip codes. If a resident lives in the center of the boundary, the approach may capture the providers they could reasonably access. But, residents who live near the edge of boundaries are likely to travel into a neighboring boundary to access health care (Guagliardo, 2004). Another issue is travel within boundaries. Population to provider ratios make no distinction between providers adjacent to a resident and providers located at the opposite boundary edge to the resident. This missing feature is referred to as distance decay. Conversely, travel time or distance does capture distance decay. One provider is recorded as five minutes from a resident while another provider is 30 minutes from the resident. Travel time and distance measures are limited, however, by failing to account for supply and demand of providers and residents (Luo & Wang, 2003). For example, two patients may each be 30 minutes from their closest provider, but one of the providers serves twice as many patients. The providers are similarly located (supply), but their demand is different.

The two step floating catchment area (2SFCA) method was designed to overcome these limitations (Wang & Luo, 2005). This method also begins by geocoding healthcare provider and population addresses into GIS software. The first step in the 2SFCA identifies each healthcare provider and calculates a population to provider ratio within a defined service area, or catchment, around that provider. For example, one provider may have 3,000 residents and 12 other providers in the one-hour travel time catchment area that surrounds them, while another provider may have 500 residents and 10 other providers in their surrounding one-hour catchment area. The second step of the 2SFCA identifies each resident or group of residents (e.g., if census tract population data is being used) and searches for all the healthcare providers within that resident's catchment area. Each provider found in a resident's catchment area will have a corresponding ratio, calculated in step one. The patient's spatial access score is found by summing all the ratios of the healthcare providers found within their catchment. Thus, the 2SFCA method overcomes the boundary problem of provider to population ratios by creating a new catchment for each provider and resident. Step one of the 2SFCA method overcomes the supply and demand problem of travel time measures by incorporating the ratio of residents to providers for each physician's catchment area.

Despite the improvements of the original 2SFCA method, there are noted limitations. Similar to traditional provider to population ratios, the 2SFCA method does not account for distance decay (Wang, 2012). If a catchment area is set as a travel time of 30 minutes, all providers within a 30 minute drive of a resident are considered equally accessible. This implies that residents are as likely to receive care from a provider 1 minute away compared to a provider 30 minutes away. The other drawback of the original 2SFCA method is the fixed catchment size (Wang, 2012). In urban areas, defining a catchment as a 15 minute travel time might be appropriate, but in rural areas residents may regularly travel longer than 15 minutes for healthcare.

Enhancements to the 2SFCA method have addressed both problems. Researchers have added a weighting function to account for distance decay within catchments (Lou & Qi, 2009). The weighting function can be discrete, where every 10 minutes the weight will increase until the catchment limit is reached. If the catchment size was set as a 30 minute travel time, providers within 10 minutes of a resident might receive the highest weight because the resident has the best access to those providers. Providers between 10 and 20 minutes from a resident would receive a smaller weight, and providers between 20 and 30 minutes of a resident would receive the smallest weight. The weighting function can also be continuous, where the weight is applied smoothly across the entire 30 minute catchment size (Wang, 2012). Luo and Qi (2009) originally proposed the three distinct weights (corresponding to travel times of 0-10 minutes, 10-20 minutes, and 20-30 minutes) without any empirical evidence that those travel times match healthcare utilization behaviors. There is a similar lack of empirical justification behind the choice of a continuous distance decay function. Researchers have implemented a continuous function to prevent the sudden weighting drop of at the end of each discrete travel zone, but how quickly or slowly that function weights travel time has yet to be determined (McGrail, 2012). Wan et al. (2012) recently proposed a relative spatial access ratio (SPAR) for use when researchers do not have empirical evidence for distance decay. The method is mathematically simple—calculating the ratio between a census tract's 2SFCA spatial access score and the mean of all census tract scores from the study area—but it did provide stable results across a variety of distance decay functions. This reduces the importance of choosing decay function parameters that mimic actual healthcare utilization behaviors, while still allowing for comparison between different geographic areas.

The other improvement to the 2SFCA method was the addition of variable catchment sizes (McGrail, 2012). Since there are two steps to the 2SFCA method, the catchments need to vary twice. First a catchment is determined for providers (referred to as the service catchment size), and then a catchment is determined for residents (referred to as the population catchment size). Some researchers suggest that the service catchment size should vary according to the setting the healthcare provider practices in (McGrail & Humphreys, 2009). This approach accounts for the reality that metropolitan providers are likely to service different size areas than rural providers. Other researchers suggest that the service catchment size should expand until a certain population to provider threshold is met (Luo & Whippo, 2012).

There are also several options for varying the population catchment size. One possibility is to cap the population catchment size at a set number of providers (McGrail & Humphreys, 2009b). If the cap is 50 providers, the catchment size for an urban resident may be a 15 minute driving time while the catchment size for a rural resident may be an hour driving time. Another approach is to use a population to provider ratio as the threshold, where the catchment size increases until a desired ratio (e.g., the HSPA ratio of 3,500:1) is reached (Luo & Whippo, 2012). As with the choice of distance decay functions, there is a lack of empirical justification for the different methods to vary catchment size. Ideally, researchers would determine both distance decay within catchments and catchment sizes by referencing medical claims data and clinical records to identify how far different proportions of patients travel to receive different types of care. Mao and Nekorchuk (2013) recently used data from a travel survey to determine that Florida households traveled on average 22.8 minutes for a medical or dental service, with the first quartile at 10 minutes and the third quartile at 30 minutes. They then set both the service and population catchment sizes at a max of 30 minutes travel time. Although the data collection and implementation were simplistic—for example, there was no distance decay within the 30 minute catchment, the type of service was broadly defined, and the sample only included 118 households—the technique is an example of using actual healthcare utilization behaviors to define the 2SFCA parameters.

Overall, cancer disparities in Appalachia warrant further research. Appalachia residents' reduced access to healthcare contributes to those cancer disparities. Unfortunately, there is no published research on Appalachia residents' access to care using the latest measure of spatial access, the 2SFCA method. This likely stems from the variety of functional parameters that the 2SFCA offers and the lack of clear guidance as to which are most appropriate. As a result, research in Appalachia examining cancer disparities and access to care is not using the most comprehensive approach, thereby limiting the conclusiveness of its results.

1.2 Nature of the Study

This thesis is a retrospective cohort study that investigates spatial access to care in the Appalachia region and its impact on breast cancer care. The research uses a linked dataset of combined patient clinical and treatment information, healthcare provider geographic information, and census-level demographic information. Study participants are drawn from the Appalachia counties of Kentucky, North Carolina, Ohio, and Pennsylvania during the 2006-2008 calendar years.

Initially the thesis focuses on the methodology behind the measurement of spatial access to care. Population and provider locations in the study area were geocoded into GIS software. Population locations and counts are from Census blocks of the U.S. Census 2000 Tiger/Line files. Primary care physician locations are obtained from the American Medical Association (AMA) Physician Masterfile. Spatial access to more specialized cancer care is also evaluated. All U.S. Food and Drug Administration (FDA) accredited mammography facilities in the study area during 2008 are geocoded. Corresponding population counts and locations are drawn from Census block groups of the U.S. Census 2000 Tiger/Line files. Census block groups, instead of the smaller Census blocks, are used with the mammography facilities because block groups allow researchers to determine the population of women age 40 and above, which was the recommended age for mammography screening during the study years (U.S. Preventive Services Task Force, 2009). Census blocks are more precise because they cover a smaller area, but age and gender are not available at the block level. Thus, Census blocks are used when evaluating access to primary care physicians because all members of the population use primary care services and no further population distinctions are necessary.

Once population data and healthcare provider locations are input into GIS software, spatial access can be computed for each geographic boundary (i.e., Census blocks or Census block groups) in the study area. Access is determined using each of the methods referenced in section 1.1, including the traditional methods of population to provider ratios and travel time to the nearest provider location, as well the 2SFCA method and its iterations. These methods will determine spatial access to primary care physicians and spatial access to mammography facilities. Spatial statistics were then used to evaluate each of the spatial access methods.

After identifying a best practice approach for measuring spatial access to care in Appalachia, I examine how access to care impacts breast cancer stage at diagnosis, an important clinical indicator for patient outcomes. I also examine how access to care predicts receipt of adjuvant hormonal therapy after breast cancer surgery. Street addresses of breast cancer patients from the Appalachia counties of Kentucky, North Carolina, Ohio, and Pennsylvania during 2006-2008 were geocoded in GIS software and matched to corresponding Census blocks or Census block groups. Each patient is associated with the spatial access score computed for their corresponding Census designated geographic area. Regression modeling was used to examine whether spatial access to care predicts late-stage breast cancer diagnosis or receipt of adjuvant therapy. Additional predictor variables tested include provider characteristics, facility characteristics, patient socioeconomic data, and patient demographic data. The regression analysis serves as a measure of predictive validity for the recommended measure of spatial access.

1.3 Study Aims

This research investigates spatial accessibility in reference to primary care, where travel times and service areas are smaller, and cancer care, where travel times and service areas are larger. I then apply the refined spatial accessibility methods to two cancer care questions in Appalachia: 1) How does spatial access to care impact stage of breast cancer at diagnosis? 2)

How does spatial access to care impact whether or not patients receive recommended adjuvant hormonal therapy? The following three objectives will achieve these goals:

Aim 1: **Investigate methods for determining spatial accessibility to cancer care in the Appalachia region and identify best practices for future research.**

Specialty health care, including cancer care facilities, often requires further travel than primary care, especially in rural areas (Huang et al., 2009). Using the 2SFCA method to measure spatial access to cancer care may require further refinement of distance decay and catchment size functions. This aim evaluates the following methods for determining spatial access to mammography centers in the Appalachia region: 1) provider to population ratios, 2) travel time to providers, 3) 2SFCA method with three distance decay functions (quick-step discrete weighting function, slow-step discrete weighting function, and a continuous weighting function) and two variable catchment size functions (McGrail, 2012), and 4) relative spatial access ratio (SPAR) derived from the 2SFCA method (Wan et al., 2012). The results will lead to a recommendation of the most appropriate method for measuring spatial access to cancer care as compared in Appalachia.

Aim 2: Investigate methods for determining spatial accessibility to primary care in the Appalachia region and identify best practices for future research.

Primary care access is a crucial component of population health (Guagliardo, 2009). Methods to measure spatial access to primary care in Appalachia must perform effectively across rural, urban, and suburban settings in a large geographic area. The same spatial access methods will be evaluated as in Aim 1. The results will lead to a recommendation of the best performing method for measuring spatial access to primary care in Appalachia, with updated distance decay and catchment size functions if necessary.

Aim 3: Apply recommended spatial accessibility method to examine how access to health care affects breast cancer stage at diagnosis and receipt of adjuvant hormonal therapy. Breast cancer stage at time of diagnosis is an important predictor for overall survival, with later staged diagnoses presenting fewer treatment options (Henry et al., 2011). There is controversy in the literature as to whether travel time is an important predictor of late stage diagnosis. A recent study examined the issue across 10 state cancer registries and found no impact of travel time, but failed to use the more rigorous 2SFCA method (Henry et al., 2011). Adjuvant hormonal therapy is recommended for hormone receptor positive patients after either breast conserving surgery or mastectomy, and was also evaluated as a predictor of guideline concordant treatment. Aim three will model breast cancer stage at diagnosis and receipt of adjuvant hormonal therapy as a function of access to care variables, including spatial accessibility determined from the recommended methods in aims one and two. This aim will generate separate models for access to primary care and access to mammography facilities.

1.4 Significance of the Study

The significance of this work stems from the methodologies evaluated, the importance of the cancer care question investigated, and the broad potential applicability of the work. The recommended methodology has potential application to other disadvantaged population groups outside of Appalachia and to other cancer care and chronic disease outcomes within Appalachia.

The 2SFCA method has never been used to study access to care in the Appalachia region, despite its improvements over existing methods. Research on cancer treatment and outcomes in Appalachia needs to employ current best practice approaches for measuring spatial access to care, particularly because poor access to care is often implicated in Appalachia's cancer disparities. Here, I provide the groundwork for future public health researchers in Appalachia to use the most accurate methods for determining spatial access to care.

The research also addresses an unresolved cancer care question in Appalachia: How does access to cancer care impact stage of breast cancer diagnosis? Research in Kentucky determined that longer travel times to mammography facilities predicted later stage diagnoses (Huang et al., 2009). Subsequent research across 10 states found no effect of travel time to mammography facility on breast cancer stage at diagnosis (Henry et al., 2011). In this thesis, I use the more rigorous 2SFCA method to examine the question.

The recommended methodology's potential application to other disadvantaged population groups outside of Appalachia is also significant. Previous research has measured access to care for ethnic minorities, individuals with disabilities, veterans, and many other disadvantaged groups. Clarification of the most accurate measures for determining spatial access to care will assist research in each of those populations, potentially leading to health policy changes with positive impacts. The results of this thesis will assist further public health research on additional cancer and chronic disease outcomes within Appalachia. This research applies spatial access methods to breast cancer care, but access to care is implicated in treatments for many other cancers and chronic diseases in Appalachia.

1.5 References

Appalachian Regional Commission (ARC). (2010). Economic Assessment of Appalachia. An Appalachia Community Cancer Network. 2009. *The Cancer Burden in Appalachia*. Retrieved from http://www.accnweb.com/DisparitiesData.aspx

Appalachian Development Initiative Report. Retrieved from http://www.arc.gov/images/newsroom/publications/EconomicAssessmentofAppalachiaJune2010.pdf

Blackley D., Behringer B., & Zheng S. (2012). Cancer mortality rates in Appalachia: descriptive epidemiology and an approach to explaining differences in outcomes. *J Community Health*. 37(4): 804-813.

Dai, D. (2010). Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. Health Place, 16(5), 1038–1052.

Department of Health and Human Services: Health Resources and Services Administration. (2013). Primary Medical Care HPSA Designation Overview. Retrieved from http://www.hrsa.gov/shortage/

Fisher J, Engelhardt H, Stephens J, et al. (2008). Cancer-related disparities among residents of Appalachia Ohio. J Health Dispar Res Pract. 2:61–74.

Freeman, A.B., Huang, B., & Dragun, A.E. (2012). Patterns of Care With Regard to Surgical Choice and Application of Adjuvant Radiation Therapy for Preinvasive and Early Stage Breast Cancer in Rural Appalachia. Am J Clin Oncol. 35(4): 358-63.

Guagliardo MF. (2004). Spatial accessibility of primary care: concepts, methods and challenges. Int J Health Geogr. 3:3.

Gulliford, M., Figueroa-Munoz, J., Morgan, M. et al. (2002). What does `access to health care' mean? J Health Serv Res Policy. 7;186-188.

Hall, H.I., Uhler, R.J., Coughlin, S.S., & Miller, D.S. (2002). Breast and cervical cancer screening among Appalachian women. Cancer Epidemiol Biomarkers Prev. 11:137–142.

Harlan, L.C., Greene, A.L., Clegg, L.X., et al. (2005). Insurance status and the use of guideline therapy in the treatment of selected cancers. J Clin Oncol. 23(36):9079–9088.

Henry K.A., Boscoe, F.P., Johnson C.J., Goldberg D.W., Sherman R., & Cockburn M. (2011). Breast cancer stage at diagnosis: Is travel time important? J. Community Health. 1;36:933–942.

Huang B., Dignan M., Han D., & Johnson O. (2009). Does distance matter? Distance to mammography facilities and stage at diagnosis of breast cancer in Kentucky. The Journal of Rural Health. 25(4): 366–371.

Jones, A.P., Haynes, R., Sauerzapf. V. et al. (2008b). Travel times to health care and survival from cancers in Northern England. Eur J Cancer. 44(2): 269–274.

Lian M, Struthers J, Schootman M (2012). Comparing GIS-Based Measures in Access to Mammography and Their Validity in Predicting Neighborhood Risk of Late-Stage Breast Cancer. PLoS ONE 7(8): e43000. doi:10.1371/journal.pone.0043000.

Luo, W. & Qi, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. Health Place. 15:1100–1107.

Luo, W. & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. Environment and Planning B. 30(6): 865–884.

Luo, W. & Whippo, T.L. (2012). Variable catchment sizes for the two-step floating catchment area (2SFCA) method. Health Place. 18:789–795.

Mao, L. & Nekorchuk, D. (2013). Measuring spatial accessibility to healthcare for populations with multiple transportation modes. Health Place. 24:115-122.

McGrail, M.R. (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. Int J Health Geogr, 16:11-50.

McGrail, M.R. & Humphreys, J.S. (2009). The index of rural access: an innovative integrated approach for measuring primary care access. BMC Health Serv Res. 9:124.

McGrail, M.R. & Humphreys, J.S. (2009b). A new index of access to primary care services in rural areas. Aust NZ J Pub Health. 33:418–423.

Smith & Hollomon, 2011. Health Status and Access to Health Care Services: A comparison between Ohio's rural Non-Appalachian and Appalachian families. *Fam Community Health.* 34(2): 102-110.

Paskett, E.D., Fisher, J.L., Lengerich, E.J. et al. (2011). Disparities in underserved white populations: the case of cancer-related disparities in Appalachia. Oncologist. 16(8): 1072-81.

Pollard, K. & Jacobsen, L.A. (2013). The Appalachian Region: A Data Overview from the 2007-2011 American Community Survey Chartbook. Retrieved from http://www.prb.org/pdf13/appalachia-census-chartbook-2012.pdf

U.S. Preventive Services Task Force. (2009). Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 151: 716–726, W-236.

Wan, N., Zhan, F.B., Zou, B., & Chow, E. (2012). A relative spatial access assessment approach for analyzing potential spatial access to colorectal cancer services in Texas. App Geography. 12:291–299.

Wang, F. (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. Ann Assoc Am Geogr. 102(5): 1104–1112.

Wang F. & Luo W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. Health Place. 11:131–146.

Winger, P.A., Tucker, T.A., Jamison PM, et al. (2003). Cancer in Appalachia, 2001-2003. Cancer; 112:181–192.

Yao N, Lengerich EJ, Hillemeier MM. (2012). Breast cancer mortality in Appalachia: reversing patterns of disparity over time. J Health Care Poor Underserved. 23(2):715-25.

CHAPTER 2

Evaluating and Comparing Methods for Measuring Spatial Access to Mammography Centers in Appalachia

2.1 Abstract

Purpose

This study evaluated spatial access to mammography centers in Appalachia using both traditional access measures and the two-step floating catchment area (2SFCA) method. Spatial Access to care was compared between urban and rural and Appalachia and non-Appalachia areas of the same states.

Methods

The study region included the entire states of Pennsylvania, Ohio, Kentucky, and North Carolina. Accredited mammography centers during 2008 and women age 45 and older at the census block group level were geocoded. Ratios of county mammography centers to women age 45 and older, driving time to nearest mammography facility, and various 2SFCA approaches were compared.

Results

Closest travel time measures favored urban areas. Mammography center to population ratios produced abrupt results based on rigid county boundaries. The 2SFCA method produced varied results depending on the parameters chosen. Overall, Appalachia areas had greater travel times to their

closest mammography center. Appalachia areas in OH and NC had worse 2SFCA scores than non-Appalachia areas of the same states.

Conclusion

Significant differences exist between 2SFCA measures depending on which function parameters are chosen. A relative 2SFCA approach, the spatial access ratio (SPAR) method, was recommended because it helped minimize the differences between various 2SFCA approaches.

2.2 Introduction

The Appalachia region of the U.S. spans a 13 state region from Alabama to New York. The region has been designated as a special population of interest by the National Cancer Institute because of disparities in cancer outcomes and treatment patterns (Lane et al., 2012; Lengerich et al., 2005). An example of these disparities are higher mortality rates for lung and bronchus, colon, and cervical cancer (Blackley, Behringer, & Zheng, 2012); reduced mammography screening (Bhanegaonkar et al., 2012); and lower rates of breast conserving surgery, rather than mastectomy (Freeman, Huang, & Dragun, 2012). Human population in Appalachia is sparse, with 42 percent of the population classified as rural, as compared to a national average of 20 percent (ARC, 2014). The region also has a lower per capita income and a higher poverty rate than the national average (Pollard & Jacobson, 2014). Due to Appalachia's rural population, distressed economic indicators, and mountainous terrain, reduced access to care has been implicated in the region's cancer disparities (Paskett et al., 2011).

Access to care is a multidimensional concept that consists of both spatial and nonspatial factors. Nonspatial factors included age, socioeconomic status, and ethnicity, for example (Wang & Luo, 2005). Spatial access is focused on the geographic distribution of healthcare providers and the populations they serve (Wang, 2012). Some researchers further distinguish between potential spatial access, referring to the geographic possibility that a service will be utilized, and realized access, which refers to the actual service used (McGrail, 2012). This study focuses on potential spatial access (further termed ‘spatial access’ for brevity.).

Traditionally, spatial access was calculated as a provider to population ratio, often using counties as the geographic boundaries (Wang, 2012). The method is simple to compute, requiring only population statistics (e.g., U.S. Census data) and the number of healthcare providers working in that county. The current Health Professional Shortage Area (HPSA) designation by the U.S. Department of Health and Human Services (DHHS) relies on this technique (DHHS, 2010). DHHS sets a threshold of one physician per 3,000 people in a rational service area to define HPSAs. With more widespread use of geographic information systems (GIS) technologies, researchers began using travel time or travel distance between population and providers as a measure of spatial access (Wang & Lou, 2005). This allowed more specificity, as individual patient and provider addresses could be geocoded into GIS software.

Although these two methods are the most widely used, they have noted limitations (Guagliardo, 2004). The fixed boundaries of provider to population ratios do not match real world healthcare utilization. Ratios also fail to account for travel impediments within their fixed boundary, giving equal weight to providers and populations on opposite ends of a boundary as those geographically close. While travel time or distance GIS calculations overcome the latter problem of travel impediments, the calculations fail to capture the supply and demand element of healthcare.

The two-step floating catchment area (2SFCA) has emerged as an additional measure of spatial access to care that overcomes some of the limitations of the original methods (Luo &

Wang, 2003; Wang & Luo, 2005). The first step of the 2SFCA identifies the populations that each healthcare provider serves. This area is called a catchment and can range from a 30 to 60 minute travel time, or longer, depending on the type of care. A provider to population ratio is computed for each physician's catchment. Step two moves to each population area and searches for all service providers within that population's catchment, also set at 30 to 60 minutes, or longer. The step one ratios within a population's catchment are then summed, resulting in an access score for each population location.

There are limitations to the original 2SFCA method, however (Luo & Whippo, 2012; Wang, 2012). The step one and step two catchments still suffer from the boundary problem described for traditional provider to population ratios. For some service providers a 30 minute catchment may be accurate, but others will serve a smaller area (downtown urban providers) or larger area (rural community providers). Also, within the step one and step two catchments, the original 2SFCA fails to account for travel impediments, also termed distance decay. During step one, for example, a population 5 minutes from a provider is deemed as likely to use the service as a population 25 minutes away.

In response to these limitations, several researchers have modified the original 2SFCA (Luo & Qi, 2009; McGrail & Humphreys, 2009). A distance decay function was added to distinguish between different geographic distances within a catchment. The catchment can be broken into distinct zones (e.g., 0-10 min., 10-20 min, 20-30 min.), with a unique weight applied at each zone (Luo & Qi, 2009). A continuous function has also been used (McGrail & Humphreys, 2009) where there is no drop at the end of the zone. Weights are applied smoothly at each increase in travel time or distance. Another addition has been the use of variable catchment sizes. One technique is to define a desired provider to population ratio (e.g., 1:3,000) and expand a catchment until that ratio is met (Luo & Whippo, 2012). Another technique is to cap the number of providers a population can access (e.g., 100), and when that limit is reached the catchment ends (McGrail & Humphreys, 2009). Taking a different approach, another group of researchers (Wan et al., 2012) created the spatial access ratio (SPAR) score, which can use any of the modified 2SFCA methods mentioned by taking a population's (i.e., census tract) 2SFCA score and dividing it by the mean 2SFCA score of an entire region. The SPAR score represents a relative comparison that may be useful when researchers are unsure how much of a travel impediment exists and therefore which decay weights to use (Wan et al., 2012).

The 2SFCA method, despite its noted benefits, has never been used to study access to care in Appalachia. Outside of Appalachia, only a limited amount of research has compared and contrasted the effects of the newer 2SFCA parameters (i.e., distance decay and variable catchment sizes) on resulting access to care scores (McGrail, 2012; Wang, 2012). These comparisons generally occur in the context of primary care. Thus, researchers in Appalachia who study spatial access to more specialized cancer care need guidance on which methods are most appropriate and which parameters to choose.

The primary goal of this study was to explore patterns of potential spatial access to mammography centers in Appalachia. We compared results across the measures of provider to population ratios, travel times between populations and providers, and the newer 2SFCA method. Within the 2SFCA method, we compared the effects of added distance decay functions and variable catchment sizes on the resulting access scores. Finally, we evaluated the differences in access to mammography centers between Appalachia and non-Appalachia areas of the same states, using all three spatial access measures.

2.3 Study area and data

We focused on a four state region of Appalachia: PA, OH, KY, and NC. All U.S. Food and Drug Administration (FDA)-accredited non-mobile mammography facilities in the four states during the year 2008 were obtained. There were 1,181 mammography centers in the four state study region, and each center had a street address that was able to be geocoded into GIS software. We did not include mammography centers of bordering states, which can affect access scores of populations near state borders.

Population data were extracted from the 2010 U.S. Census at the census block group level. Block groups usually have between 600 and 3,000 people, with an ideal size of 1,500. Much prior research measuring spatial access to care aggregated population data at larger geographic areas, generally zip codes (Dai, 2010) and census tracts (Wan et al., 2012). Using census block groups is more computationally intensive but allows greater precision in geographic estimation. Following previous research (Luo & Wang, 2003), we used the population weighted centroid of each block group, rather than the geographic centroid (U.S. Census Bureau, 2014). Centroid locations are given as latitude and longitude, which allows for integration into GIS software. Unlike mammography centers, we did include populations in neighboring states to simulate more accurate spatial access along state borders. We explore this edge effect in greater detail in the discussion.

The Appalachia Regional Commission's (ARC) county designations were used to distinguish Appalachia and non-Appalachia areas. There were 28,418 total block groups, although only 28,357 had residents. Of those, there were 8,707 block groups in Appalachia regions of the four states and 19,650 block groups in non-Appalachia regions. There were 10,717,421 people living in Appalachia regions and 27,392,443 people living in non-Appalachia regions. Further designation of rural and non-rural areas was used when comparing spatial access scores. Urban-rural designation was done by census tract, with census RUCA codes of 7-10 considered rural and codes 1-6 considered urban-suburban (Weeks et al., 2004). The rural population was 2,971,896 people and the urban population was 35,137,968 people. We used descriptive statistics and t-tests to compare access scores between Appalachia and non-Appalachia regions and between rural and urban regions.

2.4 Methodology

Provider to population ratios

We calculated provider to population ratios at the county level. The number of mammography centers in each county was divided by the population of women age 45 and older in that county. Guidelines called for mammography screening beginning at age 40 (U.S. Preventive Service Task Force, 2009). However, the 2010 Census used age groups 35-44 and 45-54. We chose to use women age 45 and older, consistent with prior research (Anderson et al., 2014). Each census block group, our primary unit of analysis, was then assigned a county ratio based on the geographic county it resided in.

Travel time

Mammography centers and census block groups were geocoded into ArcGIS (Version 10.1, ESRI Inc., Redlands, CA). We used the Origin-Destination (OD) cost matrix function of the ArcGIS Network Analyst extension to compute travel times between the population weighted centroid of each block group and each mammography center, only including pairs within one hour of each other. A recent 10-state cancer registry study (Henry et al., 2011) with over 160,000 breast cancer patients found that the average travel time to a patient's diagnosing mammography facility was 16.5 minutes and 97% of patients lived within one hour of their diagnosing facility. Along with travel time to the closest facility, we also computed the average travel time to the three closest facilities.

Two-step floating catchment area

The first step of the original 2SFCA method identifies all populations (k) within a service provider's (j) maximum catchment (d_o), set as a maximum of 60 minutes driving time. Although other 2SFCA mammography studies have used a 30 minute maximum catchment size (Lian et al., 2012), we felt the rural nature of Appalachia warranted a 60 minute maximum catchment size (McGrail & Humphreys, 2009). A population to provider ratio (R_j) was then computed using the total service providers (S_j) and total population (P_k) within catchment j:

Step 1: $R_j = S_j / \sum k \in \{d_{jk} < d_o\}\ P_k$

Step 2 of the 2SFCA identifies all service providers (j) within the maximum travel time (d_o) of a population location (i) and sums all of the eligible service providers' step 1 ratios (R_j), resulting in that population location's access score (A_i):

Step 2: $A_i = \sum j \in \{d_{ij} < d_o\}\ R_j$

We also examined three additional versions of the 2SFCA. The first included an added distance decay function across catchment time zones (Luo & Qi, 2009). Catchments were divided into four time zones: 0-10 minutes, 11-20 minutes, 21-30 minutes, and 31-60 minutes. Weights were applied at each zone in order to simulate the travel impediment as travel time increased. We used fast-step weightings and slow-step weightings, chosen from prior research (McGrail, 2012; Wan et al., 20120):

- Fast-decay weights (*w)* = 1, 0.60, 0.25, 0.05
- Slow-decay weights (*w*) = 1, 0.80, 0.55, 0.15

Thus, each population location (k) in step one and each physician ratio (R_j) during step two are discounted by an above weight, depending on the travel time between population and provider (i.e., above weights correspond to 0-10, 11-20, 21-30, and 31-60 minute time zones):

- Step 1: $R_j = S_j / \sum k \in \{d_{jk} < d_o\}\ P_k\ w$
- Step 2: $A_i = \sum j \in \{d_{ij} < d_o\}\ R_j\ w$

The second version of the 2SFCA we examined included both distance decay weights and varying catchment sizes. For step one of the 2SFCA, we used local population characteristics to determine how far to expand a service provider's catchment, and thereby whether or not to apply the distance decay weights mentioned above (McGrail & Humphreys, 2009b). If a population was within 10 minutes of a service provider, no decay weight was applied. If the service provider was one of the five closest providers, no weight was applied because that provider will likely provide service to populations with few other choices. Prior research (McGrail & Humphreys, 2009b) set this limit to the closest 25 services, but that study examined primary care, not mammography centers, which are less abundant. Finally, if the service provider's town population was greater than 5,000 and the population center's town population was less than half of the service provider's, no weight was applied. This rule represents situations where the service provider is in a larger town that likely provides services to smaller nearby communities. To estimate town populations, we linked mammography centers and census block groups with their closest Census Designated Place (CDP). CDPs are meant to represent the settled concentration of populations as either a city, town, village, or borough, and can cross county boundaries (U.S. Census Bureau, 2014). For step two of the 2SFCA, we capped the number of services a population center could access at 20. Thus, the first 20 services were weighted according to the distance decay rules above, but services after the first 20 closest were not included in that population's catchment. The cap represented the point where additional services do not offer increasing access. Again, prior primary care research set this cap at 100 (McGrail & Humphreys, 2009), but due to lesser numbers of mammography centers we set the cap at 20.

When evaluating the combination of 2SFCA distance decay and variable catchment sizes, we also included a continuous decay weighting scheme to contrast with the slow-decay and fast-decay weights (McGrail & Humphreys, 2009). The continuous weighting scheme followed the same rules and caps as above. But, when travel times (*d*) between population and provider were greater than 10 minutes and less than 60 minutes, and a slow or fast-decay weight would be applied, the following continuous weighting (*w*) was applied instead:

- Continuous-decay weights (*w*) = ((60-*d*) / (60-10)) ^ 1.5

The final 2SFCA version we considered was the spatial access ratio (SPAR) score (Wan et al., 2012). SPAR was created in response to the subjectivity of distance decay weights and varying catchment sizes. SPAR is simple to calculate, being the ratio between a block group's 2SFCA score and the mean of all block group 2SFCA scores.

In sum, each census block group was assigned 10 spatial access scores for comparison: 1) County mammography facility to women age 45 years and older ratio; 2) Travel time to closest mammography center; 3) Original 2SFCA score; 4) Slow-decay 2SFCA score; 5) Fast-decay 2SFCA score; 6) Slow-decay and variable catchment 2SFCA score; 7) Fast-decay and variable catchment 2SFCA score; 8) Continuous-decay and variable catchment 2SFCA score; 9) Slow-decay and variable catchment SPAR score; 10) Fast-decay and variable catchment SPAR score. We mapped spatial access scores over their geographic area to visualize the resulting access within our study region.

2.5 Results

The regions with the highest population density included Philadelphia and surrounding southeastern Pennsylvania; the Pittsburgh area of western Pennsylvania; the Ohio cities of Cleveland, Akron, Columbus, and Cincinnati; Louisville and Lexington in Kentucky, and the smaller cities of central North Carolina (Fig. 2.1). The least populated areas were north central Pennsylvania and eastern Kentucky.

Spearman correlations were statistically significant between each of the 10 spatial access measures (Table 2.1). Increases in the provider to population ratio and the 2SFCA scores signal increasing spatial access, thus the negative correlation of both those measures with travel time to the closest mammography center, where decreasing travel times signal increasing spatial access. Travel time and population to provider ratio had the weakest correlation (rho = -0.115). Slightly stronger correlations occurred between travel time and the 2SFCA measures ($-0.522 \leq$ rho ≤ -0.180) and between provider to population ratio and the 2SFCA measures ($0.351 \leq$ rho ≤ 0.5450). Expectedly, the strongest correlations occurred between the various 2SFCA scores. The original 2SFCA score had a moderately strong correlation ($0.599 \leq$ rho ≤ 0.753) with the modified 2SFCA approaches, while each of the modified 2SFCA approaches were strongly correlated ($0.777 \leq$ rho ≤ 1).

County provider to population ratios are shown in Figure 2.2, with darker colors representing larger ratios in the higher quintile groups, and therefore greater spatial access. Many of the highest scores are found in Kentucky, where lower populations result in higher ratios.

The largest population centers—the Pittsburgh and Philadelphia regions—have mediocre scores, despite having the greatest number of mammography facilities. Rigid boundary differences are present, especially in Kentucky, where scores from the largest and smallest quintiles border each other.

The largest driving times between census block groups and their closest mammography facility occurred in north central Pennsylvania, eastern Ohio, and eastern Kentucky (Fig. 2.2). The Cleveland-Akron to Pittsburgh corridor, the Philadelphia area, and central North Carolina had the shortest travel times. The shortest travel time measure had more gradual transitions from higher to lower access, compared to the county ratio measure.

The census block groups with the highest original 2SFCA scores were in the Cleveland-Akron-Pittsburgh corridor, as well as central and eastern Kentucky (Fig. 2,3a). When slow (Fig. 2.3b) and fast (Fig. 2.3c) distance decay functions were added, higher access scores became more dispersed and less clustered in contiguous areas, with the greatest difference occurring between the original and fast-decay 2SFCA scores. The differences were more subtle when varying the catchment sizes on the slow-decay (Fig. 2.3d) and fast-decay (Fig 2.3e) scores. For the slow-decay scores, using a variable catchment decreased access in the northern corridor of Cleveland-Youngstown-Pittsburg because populations were only allowed to access the first 20 providers. These metropolitan areas had more than 20 mammography centers (e.g., the greater Pittsburgh area has about 60), and without a limit on the number of providers their populations can access their scores increase. Conversely, using variable catchments comparatively increased the slow-decay access for the NC cities Greensboro and Charlotte because those cities do not have a surplus of mammography centers (e.g., the greater Greensboro area has about 20), and thus their access scores did not drastically decrease when limiting catchment sizes.

The most noticeable difference when varying catchment sizes within fast-decay scores (Fig. 2.3e) was reduced access in northwest Ohio, the Youngstown-Pittsburg corridor, and rural central PA. Using a continuous-decay function along with varying catchments (Fig. 2.3f) resulted in scores more similar to the slow-decay and variable catchment 2SFCA scores, rather than the fast-decay and variable catchment 2SFCA scores. The greatest difference with the continuous-decay scores was the increase in scores throughout much of NC. The difference between slow-decay, variable catchment and fast-decay, variable catchment SPAR scores mirrored the differences between their corresponding 2SFCA scores, despite previous evidence of SPAR scores being robust against changes in decay weightings (Wan et al., 2012).

It is also helpful to compare the actual change in value for a 2SFCA score at the same population center after changing the decay weighting scheme. Figure 4 plots the different 2SFCA and SPAR values at each census block group by the varying slow- and fast-decay weightings. When using only the distance decay approach (Fig. 2.4a), there was no uniform increase or decrease based on slow or fast decay weightings. The only exceptions were the highest outlier scores, where the fast-decay weighting always increased scores. The distance decay and variable catchment size approach (Fig. 2.4b) produced a uniform result where nearly all block groups had reduced scores when using the fast-decay, variable catchment technique. The SPAR distance decay and variable catchment approach (Fig. 2.4c) moderated scores to some extent, compared to the similar 2SFCA approach, but a noticeable decrease in scores after using the fast decay, variable catchment technique still occurred for a majority of block groups.

Figure 2.5 distinguishes between urban and rural census block groups when comparing the change in value for 2SFCA scores after changing the decay weighting scheme. The most noticeable difference is that the highest scores across each approach were always from rural block groups (the scales are the same between urban and rural). The lower population pressure of rural areas in step one of the 2SFCA likely results in their increased scores compared to urban areas. For urban block groups, the SPAR distance decay and variable catchment approach (Fig. 5f) was closer to achieving its intended effect of not letting decay weights uniformly impact scores. For rural block groups, however, scores were uniformly decreased with the fast-decay, variable catchment technique (Fig. 2.5e). Mammography centers are generally farther distances from rural block groups, thus faster decay weightings appear to have more impact.

Continuing with the urban-rural divide, Table 2.2 compares the mean values across all spatial access measures between urban and rural block groups. Rural block groups had a higher provider to population ratio across each state. Conversely, travel time to the closest mammography center favored urban block groups, who on average had between an 11.15 minute (PA) and 6.17 minute (KY) shorter travel time than rural block groups. In PA, the 2SFCA and SPAR measures were either very similar between urban and rural block groups, or urban block groups had larger scores (e.g., slow-decay 2SFCA and fast-decay 2SFCA). In OH and NC, scores were higher for urban block groups across each 2SFCA and SPAR measure. In KY, rural block groups generally had larger scores, with the exception of the fast-decay, variable catchment 2SFCA and the fast-decay, variable catchment SPAR. Kentucky's urban areas are smaller than the other study states, and the increasing population demands are not balanced with substantially greater numbers of mammography centers.

Finally, the mean values of spatial access scores between Appalachia and non-Appalachia block groups were compared (Table 2.3). Appalachia block groups had a higher provider to population ratio across each state, similar to the rural block groups in Table 2.2. Travel time, though, was greater for Appalachia areas, ranging from an increase of only 1.02 minutes in NC to 7.23 minutes in KY. In PA and KY, Appalachia block groups had higher 2SFCA and SPAR scores than non-Appalachia block groups. The difference in scores was closer in OH, but generally the non-Appalachia scores were higher. In NC, all 2SFCA and SPAR measures were higher in non-Appalachia block groups.

2.6 Discussion

Substandard access to care is implicated in cancer incidence and treatment patterns throughout Appalachia, yet no prior studies have evaluated the latest spatial access to care measures in the region. This study evaluated spatial access methods across a four-state region of Appalachia, compared differences between measures, and evaluated access between Appalachia and non-Appalachia areas of the same states. All access measures were significantly correlated with each other, although the strongest correlations occurred between variations of the 2SFCA method. Measuring travel time to closest mammography centers produced expected results. Urban-core areas were all within 10 minutes of their closest mammography center, while rural parts of north central PA, eastern OH, and eastern KY comprised the majority of areas with travel times greater than 40 minutes. Appalachia populations across each state had longer average travel times than non-Appalachia populations. Provider to population ratios were more unpredictable and abrupt. Adjacent areas of rural KY, for example, comprised the lowest and highest quintile scores because of rigid county boundaries. Clear geographic differences emerged when comparing the various iterations of the 2SFCA method. It was difficult, however, to discern easily identifiable patterns corresponding to the 2SFCA parameters because of the study region's size and varying population distribution.

A consistent finding of the original 2SFCA, in regards to primary care physicians, was that urban areas had the highest access scores (Wang & Luo, 2005). In regards to mammography centers in Appalachia, our results partly confirm this trend. Urban areas in western PA, throughout Ohio, and in central NC all had among the highest original 2SFCA access scores. The largest urban area, the Philadelphia region, contradicted that trend with mostly poor to moderate scores. Also surprising was that rural eastern KY comprised a large portion of the highest access scores, likely because of the low population density and a geographically consistent distribution of mammography centers through the state.

The concept of distance-decay weights within catchments was originally proposed because researchers wanted to limit accessibility at the edge of physician catchments, which often occurred between major population centers (Luo & Qi, 2009). In our original 2SFCA scores, for example, the entire corridor between northeastern OH cities received the highest quintile of access scores. After adding distance-decay weights, though, distinct regions of lower access suburban and rural areas emerged between these urban areas, particularly when using faster decay weights.

When creating methods to vary catchment sizes in the 2SFCA approach, researchers wanted to further refine access scores, especially in rural areas (Luo & Whippo, 2012; McGrail, 2012). Within a realistic outer limit (e.g., 60 minutes), a rural population should be expected to travel farther for a service if that service is one of only several available. Following McGrail (2012), we did not apply distance decay weights for those closest provider and population connections when evaluating the total population that a provider serves. Our distance decay, variable catchment 2SFCA scores changed accordingly. In rural areas of each study state, mean access scores decreased when adding the variable catchments compared to only using distance-decay weights. Providers were assumed to serve more rural populations, thus the step 1 2SFCA provider-to-population ratios were lower. In step 2, summed provider ratios around each population center were lower, resulting in lower overall access scores.

Wan et al. (2012) created the SPAR technique due to the lack of guidance in the 2SFCA method for choosing decay weights within catchments. Across a nine county region of Texas, they showed that SPAR scores remained comparatively stable between populations despite using different decay weights. They applied SPAR to a 2SFCA model using only distance decay, not variable catchments, and in reference to primary care. Our application of SPAR in Appalachia to a distance decay, variable catchment model focusing on mammography centers largely achieved SPAR's intended effect. For urban populations, which represented 92% of our study population, the SPAR technique reduced the effects of choosing between a fast or slow decay weight. (Our urban-rural classification placed many suburban areas in the urban category, for simplicity.) For these urban populations, there was no standard increase or decrease in SPAR scores corresponding to fast or slow decay weights. Conversely, for rural populations, moving from the slow to fast distance-decay, variable catchment SPAR technique resulted in a uniform decrease in scores. The average travel time to the closest mammography center for rural census block groups ranges from approximately 14 to 18 minutes, per state, which is outside of the initial 10 minute zone where no distance decay is applied. For urban block groups the average times to the closest center are all approximately less than 10, making the choice of fast or slow decay weights irrelevant. We hypothesize that this population feature makes the SPAR technique less effective for rural block groups in our study region.

Conceptually, our study had a number of strengths. To the best of our knowledge, it is the first time that the 2SFCA method has been used for Appalachia. Modern GIS software is making spatial access to care easier to measure, and public health research in the region needs to incorporate the latest techniques. Within broader work on the 2SFCA method, this study adds to a growing body of research applying the 2SFCA method to cancer care, rather than primary care

(Diad, 2010; Lian et al., 2012; Wan et al., 2012). Methodologically, our study benefited from the use of a smaller population area (census block group) than many previous studies that used either zip code areas or census tracts when employing the 2SFCA method (Wang, 2012). Although more computationally intensive, this technique provided more specificity when interpreting results. Another methodological strength was our inclusion of variable catchment sizes, which are both a theoretical and empirical improvement over distance-decay functions alone. Many studies published since the creation of variable catchment sizes did not utilize the technique (Lian et al., 2012; Mao & Nekorchuk, 2013; Tao et al., 2014).

There were some important limitations to this study. We included populations from bordering states when computing service catchments in step 1 of the 2SFCA, but we did not include mammography centers from bordering states. This likely created an edge effect in our study region, particularly when adjacent urban areas likely included many additional mammography facilities (e.g., in New Jersey and Delaware for populations in Philadelphia). In other rural areas, such as western NC, there are unlikely to be many neighboring mammography centers along the rural Tennessee border. Another methodological limitation was our lack of information about the actual capacity of each mammography center. It is plausible that urban facilities have the capacity to serve many more women simultaneously than smaller rural facilities. By not including this distinction, there is a possibility that we underestimated 2SFCA access scores in urban areas. This omission may explain the Philadelphia region's mediocre 2SFCA scores.

Another limitation, and an area we believe warrants additional research, concerns our lack of representative patient healthcare utilization data. It is generally agreed that the 2SFCA needs to include travel impediments and service area variability (McGrail, 2012). Yet, the choice of which distance-decay weight to apply, or how to vary catchment sizes, is arbitrary. For instance, research needs to empirically validate that a service provider 40 minutes from a population should receive 1/12th of the weight of a provider 15 minutes away, as our fast-decay weighting scheme assumed. Similarly, it is difficult to decide where to cap the number of services that a population will consider, as we did when using 20 mammography centers as the upper limit when varying catchment sizes.

In conclusion, we recommend the SPAR technique when researchers do not have data on how far their study population travels or what their population's travel preferences are. The SPAR technique, particularly for urban and suburban areas, reduces the impact of choosing a faster or slower weighting.

2.7 References

Anderson RT, Yang TC, Matthews SA, Camacho F et al. (2014). Breast Cancer Screening, Area Deprivation, and Later-Stage Breast Cancer in Appalachia: Does Geography Matter?. Health Services Research, 49: 546–567.

Appalachia Regional Commision (ARC) The Appalachia region. [Accessed October 10, 2014]. Available at: http://www.arc.gov/appalachian_region/TheAppalachianRegion.asp

Bhanegaonkar A, Madhavan S, Khanna R, & Remick SC (2012). Declining mammography screening in a State Medicaid Fee-for-Service Program: 1999–2008. *Journal of Women's Health*, 21(8): 821–829.

Blackley D, Behringer B, & Zheng S (2012). Cancer mortality rates in Appalachia: descriptive epidemiology and an approach to explaining differences in outcomes. *J Community Health*. 37(4): 804-813.

Dai D (2010). Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. *Health Place*. 16(5), 1038–1052.

Department of Health and Human Services (2008). Shortage Designation: Health Professional Shortage Areas & Medically Underserved Areas/Populations [Accessed October 15, 2014]. Available at: http://www.hrsa.gov/shortage/

Freeman AB, Huang B, & Dragun AE (2012). Patterns of Care With Regard to Surgical Choice and Application of Adjuvant Radiation Therapy for Preinvasive and Early Stage Breast Cancer in Rural Appalachia. *Am J Clin Oncol*. 35(4): 358-63.

Guagliardo MF (2004). Spatial accessibility of primary care: concepts, methods and challenges. *Int J Health Geogr*. 3:3.

Henry KA, Boscoe FP, Johnson CJ, Goldberg DW, Sherman R, & Cockburn M (2011). Breast cancer stage at diagnosis: Is travel time important? *J. Community Health*. 1;36:933–942.

Lane N, Konrad R, Ricketts T et al. (2012). *Health Care Costs and Access Disparities in Apppalachia*. Washington, DC: Appalachia Regional Commission.

Lengerich EJ, Tucker TC, Powell RK et al. (2005). Cancer Incidence in Kentucky, Pennsylvania, and West Virginia: Disparities in Appalachia. *Journal of Rural Health.* 21(1): 39–47.

Lian M, Struthers J, & Schootman M (2012). Comparing GIS-Based Measures in Access to Mammography and Their Validity in Predicting Neighborhood Risk of Late-Stage Breast Cancer. *PLoS ONE.* 7(8): e43000. doi:10.1371/journal.pone.0043000.

Luo W & Qi Y (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health Place.* 15:1100–1107.

Luo, W & Wang F (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B.* 30(6): 865–884.

Luo W & Whippo TL (2012). Variable catchment sizes for the two-step floating catchment area (2SFCA) method. *Health Place.* 18:789–795.

Mao, L. & Nekorchuk, D. (2013). Measuring spatial accessibility to healthcare for populations with multiple transportation modes. Health Place. 24:115-122.

McGrail MR & Humphreys JS (2009). The index of rural access: an innovative integrated approach for measuring primary care access. *BMC Health Serv Res*. 9:124.

McGrail MR & Humphreys JS (2009b). A new index of access to primary care services in rural areas. *Aust NZ J Pub Health.* 33:418–423.

McGrail MR (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *Int J Health Geogr*. 16:11-50.

Paskett ED, Fisher JL, Lengerich EJ et al. (2011). Disparities in underserved white populations: the case of cancer-related disparities in Appalachia. *Oncologist.* 16(8): 1072-81.

Pollard K & Jacobson LA (2014) The Appalachian Region: A Data Overview from the 2008-2012 American Community Survey. *Appalachia Regional Commission.* Retrieved from http://www.arc.gov/research/researchreportdetails.asp?REPORT_ID=109.

Tao Z, Cheng Y, Dai T, & Rosenberg MW. (2014). Spatial optimization of residential care facility locations in Beijing, China: maximum equity in accessibility. *International Journal of Health Geographics.* 14:13-33.

U.S. Census Bureau (2014). Centers of Population [accessed June 5, 2014]. Available at: https://www.census.gov/geo/reference/centersofpop.html

U.S. Census Bureau (2014). Geographic Terms and Concepts – Place [Accessed June 5, 2014]. Available at: https://www.census.gov/geo/reference/gtc/gtc_place.html

U.S. Preventive Services Task Force. (2009). Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 151: 716–726, W-236.

Wan N, Zhan FB, Zou B, & Chow E (2012). A relative spatial access assessment approach for analyzing potential spatial access to colorectal cancer services in Texas. *App Geography*. 12:291–299.

Wang F (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Ann Assoc Am Geogr*. 102(5): 1104–1112.

Wang F & Luo W (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health Place*. 11:131–146.

Weeks WB, Kazis LE, Shen Y, et al. (2004). Differences in Health-Related Quality of Life in Rural and Urban Veterans. *American Journal of Public Health.* 94(10):1762-1767.

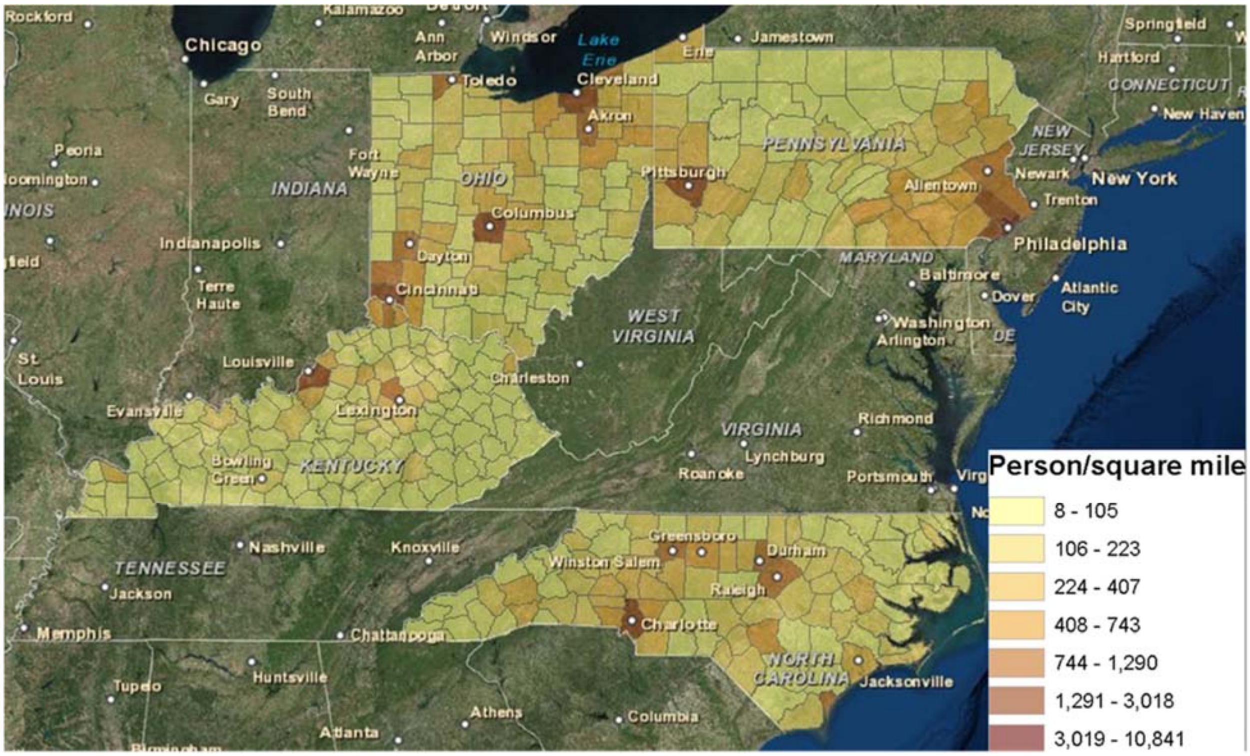

Fig 2.1 Population density across the counties of PA, OH, KY, and NC.

Table 2.1 Spearman correlations* between 10 spatial access to care measures across PA, OH, KY, and NC

	Prov. - Pop. Ratio	Travel Time	Original 2SFCA	Slow 2SFCA	Fast 2SFCA	Slow - Var. Catch. 2SFCA	Fast - Var. Catch. 2SFCA	Cont. - Var. Catch. 2SFCA	Slow - Var. Catch. SPAR	Fast - Var. Catch. SPAR
Prov. - Pop. Ratio	1									
Travel Time	-0.115	1								
Original 2SFCA	0.351	-0.180	1							
Slow 2SFCA	0.510	-0.416	0.753	1						
Fast 2SFCA	0.545	-0.522	0.640	0.948	1					
Slow - Var. Catch. 2SFCA [a]	0.541	-0.243	0.619	0.835	0.818	1				
Fast - Var. Catch. 2SFCA [b]	0.503	-0.485	0.603	0.882	0.909	0.901	1			
Cont. - Var. Catch. 2SFCA [c]	0.537	-0.212	0.599	0.790	0.777	0.953	0.836	1		
Slow - Var. Catch. SPAR	0.541	-0.243	0.619	0.835	0.818	1.000	0.901	0.953	1	
Fast - Var. Catch. SPAR	0.503	-0.485	0.603	0.882	0.909	0.901	1.000	0.836	0.901	1

2SFCA, Two-Step Floating Catchment Area; *SPAR*, Spatial Access Ratio

* All correlations significant at the 0.01 level

[a] Slow - Var. Catch. 2SFCA: Both slow-decay weightings and varying catchment size rules

[b] Fast - Var. Catch. 2SFCA: Both fast-decay weightings and varying catchment size rules

[c] Cont. - Var. Catch. 2SFCA: Both continuous-decay weightings and varying catchment size rules

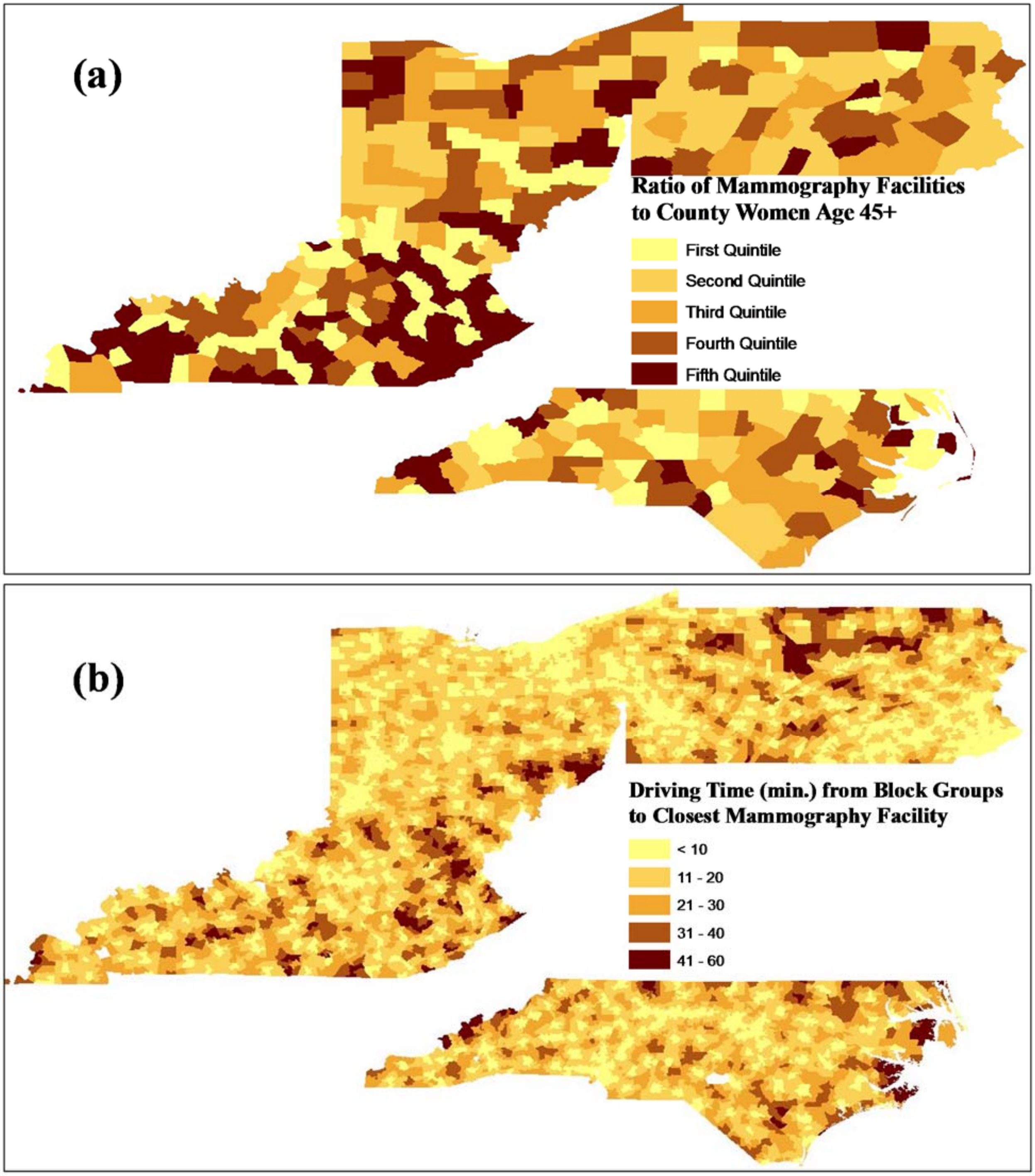

Figure 2.2 Spatial Access to Mammography centers by (a) county provider to population ratio, broken into quintiles, and (b) closest driving time from census block groups

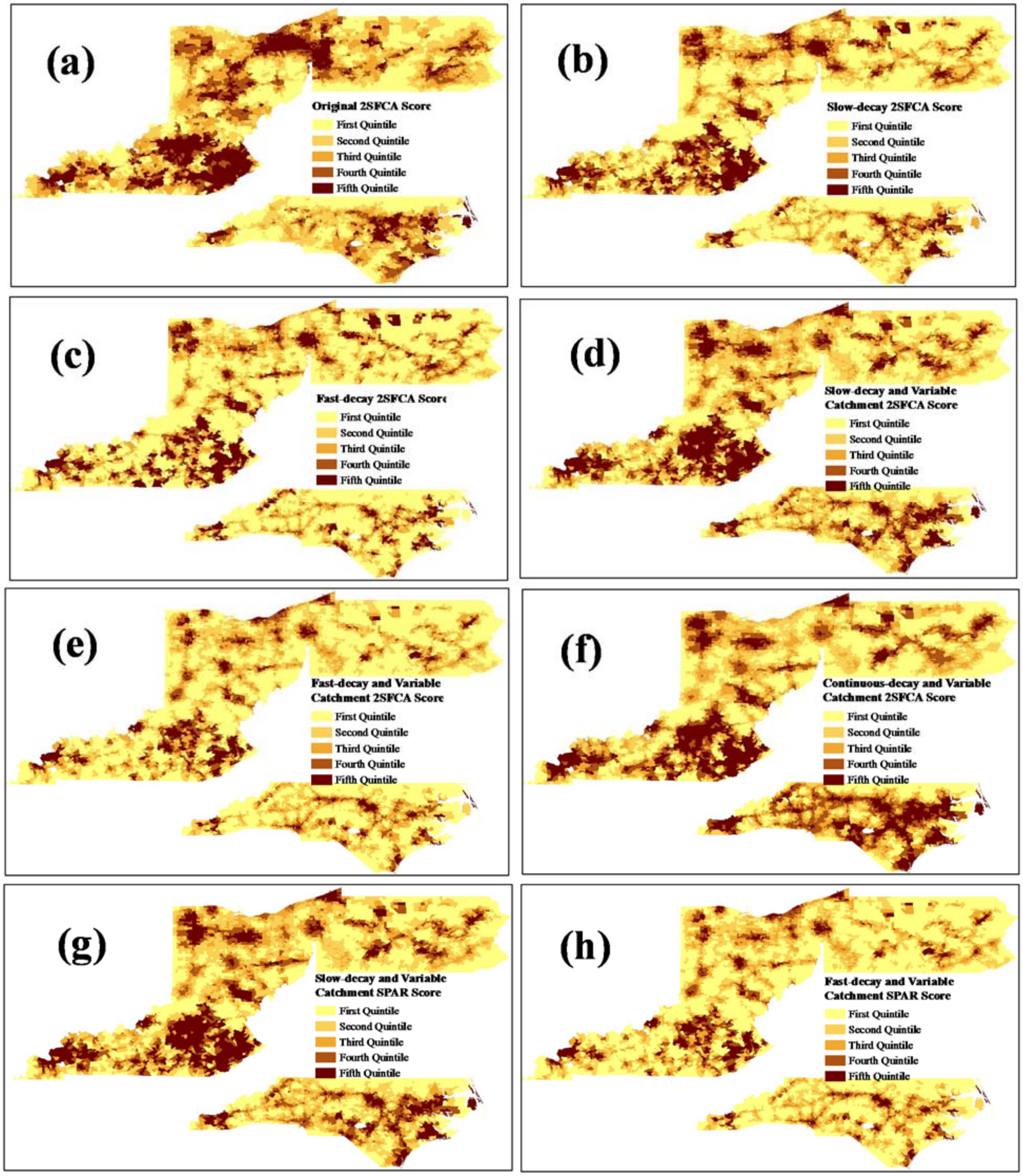

Figure 2.3 Spatial access of census block groups in PA, OH, KY, and NC to mammography centers by (a) original 2SFCA scores; (b) slow-decay 2SFCA scores; (c) fast-decay 2SFCA scores; (d) slow-decay and variable catchment 2SFCA scores; (e) fast-decay and variable catchment 2SFCA scores; (f) continuous-decay and variable catchment 2SFCA scores; (g) slow-decay and variable catchment SPAR scores; (h) fast-decay and variable catchment SPAR scores. Scores are broken into quintiles, with larger quintiles representing greater spatial access.

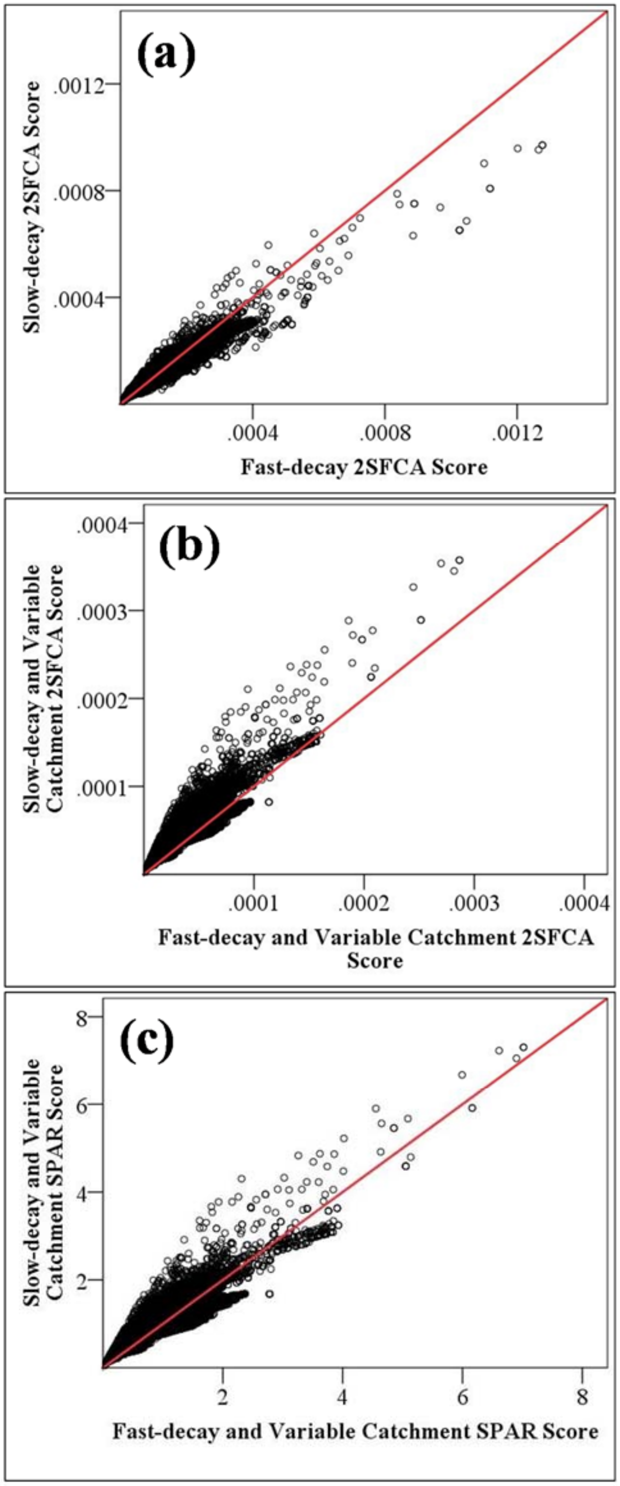

Figure 2.4 Comparison of spatial accessibilty across each census block group between (a) slow-decay and fast-decay 2SFCA scores; (b) Slow-decay, variable catchment and fast-decay, variable catchment 2SFCA scores; (c) Slow-decay, variable catchment and fast-decay, variable catchment SPAR scores.

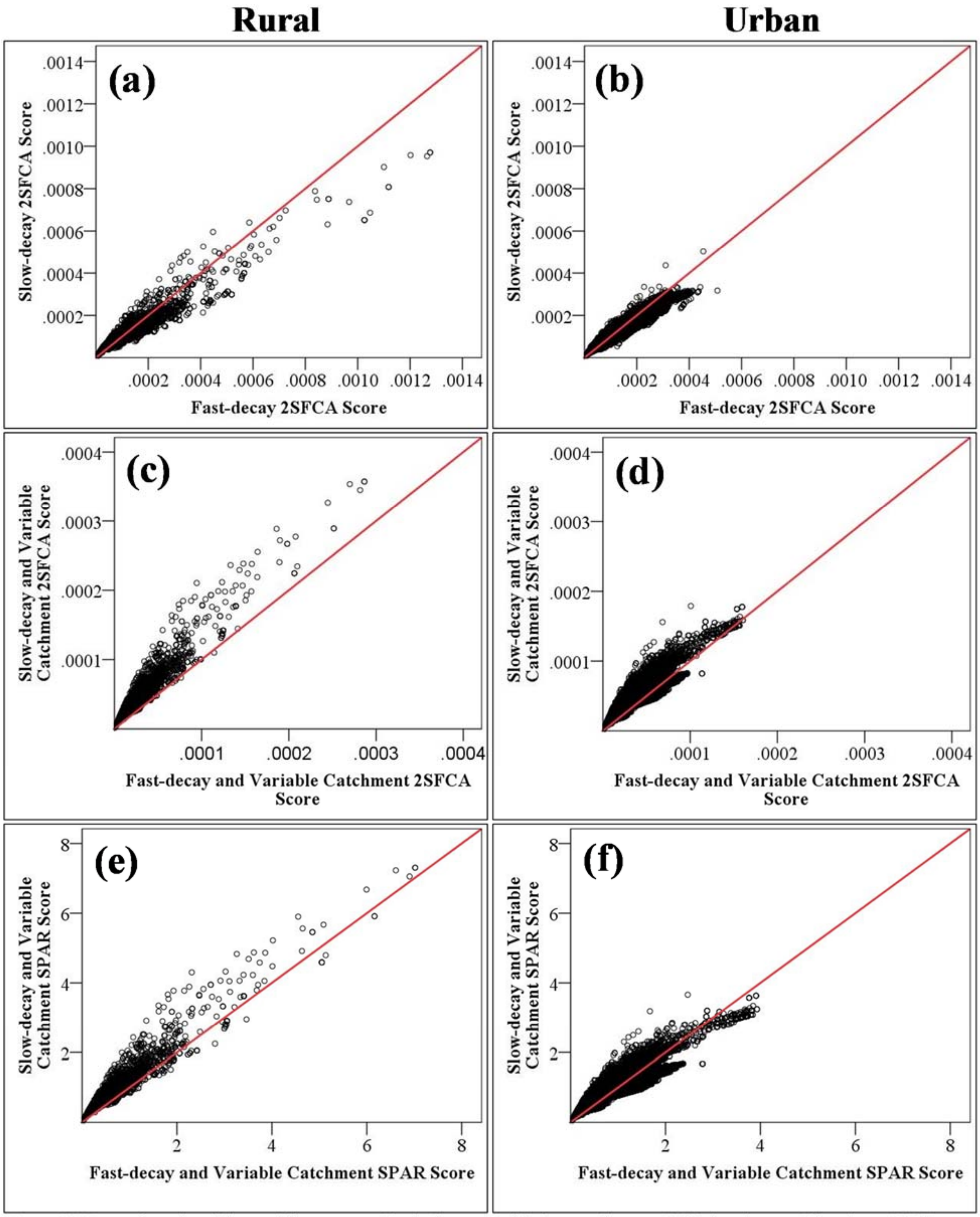

Figure 2.5 Comparison of spatial accessibilty across rural and urban census block groups between (a-b) slow-decay and fast-decay 2SFCA scores; (c-d) Slow-decay, variable catchmentand and fast-decay, variable catchment 2SFCA scores; (e-f) Slow-decay, variable catchment and fast-decay, variable catchment SPAR scores.

Table 2.2 Descriptive statistics of mammography center spatial access scores between rural and urban census block groups of PA, OH, KY, and NC.

Spatial Access Measures	All States	Pennsylvania		Ohio		Kentucky		North Carolina	
	Mean (SD)	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Provider to Population Ratio	0.0000148 (0.000073)	0.000154 (0.000070)	0.000134 (0.000039)	0.000174 (0.000084)	0.000158 (0.000066)	0.000283 (0.000194)	0.000164 (0.000093)	0.000164 (0.000111)	0.000123 (0.000052)
Travel Time to Closest Mammograpy Center (min.)	8.67 (7.83)	17.93 (11.58)	6.78 (6.39)	14.08 (9.95)	6.81 (5.85)	16.42 (11.36)	10.25 (8.95)	17.21 (10.81)	10.48 (7.71)
Origianl 2SFCA	0.000129 (0.000049)	0.000101 (0.000037)	0.000115 (0.000039)	0.000109 (0.000034)	0.000144 (0.000040)	0.000197 (0.000116)	0.000147 (0.000058)	0.000098 (0.000041)	0.000120 (0.000038)
Slow-decay 2SFCA	0.000136 (0.000058)	0.000105 (0.000052)	0.000124 (0.000046)	0.000113 (0.000047)	0.000150 (0.000049)	0.000207 (0.000145)	0.000156 (0.000066)	0.000108 (0.000061)	0.000124 (0.000047)
Fast-decay 2SFCA	0.000140 (0.000070)	0.000112 (0.000077)	0.000128 (0.000053)	0.000122 (0.000066)	0.000154 (0.000060)	0.000221 (0.000188)	0.000159 (0.000079)	0.000120 (0.000090)	0.000127 (0.000058)
Slow-decay, variable catchment 2SFCA	0.000049 (0.000024)	0.000038* (0.000018)	0.000038 (0.000020)	0.000046 (0.000019)	0.000050 (0.000014)	0.000078 (0.000054)	0.000069 (0.000033)	0.000044 (0.000023)	0.000054 (0.000021)
Fast-decay, Variable Catchment 2SFCA	0.000041 (0.000023)	0.000023 (0.000015)	0.000032 (0.000017)	0.000029 (0.000015)	0.000047 (0.000017)	0.000049 (0.000042)	0.000056 (0.000035)	0.000028 (0.000019)	0.000042 (0.000021)
Continuous-decay, Variable Catchment 2SFCA	0.000046 (0.000024)	0.000036 (0.000018)	0.000034 (0.000019)	0.000041 (0.000018)	0.000043 (0.000013)	0.000074 (0000054)	0.000064 (0.000029)	0.000048 (0.000025)	0.000058 (0.000020)
Slow-decay, Variable Catchment SPAR	1.00 (0.49)	0.77* (0.37)	0.77 (0.41)	0.93 (0.39)	1.03 (0.28)	1.60 (1.11)	1.41 (0.68)	0.89 (0.47)	1.10 (0.42)
Fast-decay, Variable Catchment SPAR	1.00 (0.55)	0.57 (0.37)	0.79 (0.41)	0.71 (0.37)	1.15 (0.42)	1.20 (1.03)	1.38 (0.85)	0.68 (0.47)	1.03 (0.52)

2SFCA, Two-Step Floating Catchment Area; *SPAR*, Spatial Access Ratio
* Not siginificant at $p < 0.05$; all other Rural and Urban comparisons were significant at $P < 0.01$.

Table 2.3 Descriptive statistics of mammography center spatial access scores between Appalachia and non-Appalachia census block groups of PA, OH, KY, and NC.

Spatial Access Measures	All States	Pennsylvania		Ohio		Kentucky		North Carolina	
	Mean (SD)	App	Non-App	App	Non-App	App	Non-App	App	Non-App
Provider to Population Ratio	0.0000148 (0.000073)	0.00146 (0.000048)	0.000125 (0.000031)	0.000189 (0.000124)	0.000153 (0.000043)	0.000244 (0.00019)	0.000170 (0.000093)	0.000128* (0.000082)	0.000126 (0.000056)
Travel Time to Closest Mammograpy Center (min.)	8.67 (7.83)	9.77 (8.68)	5.00 (4.19)	11.31 (9.36)	6.39 (5.20)	16.79 (11.77)	9.56 (8.14)	11.98 (8.17)	10.96 (8.35)
Origianl 2SFCA	0.000129 (0.000049)	0.000132 (0.000037)	0.000097 (0.000033)	0.000129 (0.000044)	0.000145 (0.000039)	0.000202 (0.000109)	0.000141 (0.000052)	0.000099 (0.000032)	0.000122 (0.000039)
Slow-decay 2SFCA	0.000136 (0.000058)	0.000138 (0.000050)	0.000109 (0.000038)	0.000144 (0.000073)	0.000149 (0.000042)	0.000204 (0.000134)	0.000153 (0.000064)	0.000109 (0.000044)	0.000126 (0.000049)
Fast-decay 2SFCA	0.000140 (0.000070)	0.000140 (0.000061)	0.000115 (0.000043)	0.000157 (0.000098)	0.000151 (0.000049)	0.000208 (0.000172	0.000159 (0.000079)	0.000115 (0.000059)	0.000129 (0.000062)
Slow-decay, Variable Catchment 2SFCA	0.000049 (0.000024)	0.000046 (0.000021)	0.000029 (0.000014)	0.000048 (0.000018)	0.000050 (0.000013)	0.000079 (0.000050)	0.000068 (0.000033)	0.000047 (0.000018)	0.000055 (0.000022)
Fast-decay, Variable Catchment 2SFCA	0.000041 (0.000023)	0.000035 (0.000019)	0.000028 (0.000013)	0.000037 (0.000021)	0.000048 (0.000016)	0.000049 (0.000039)	0.000057 (0.000036)	0.000032 (0.000016)	0.000043 (0.000022)
Continuous-decay, Variable Catchment 2SFCA	0.000046 (0.000024)	0.000043 (0.000020)	0.000025 (0.000012)	0.000043* (0.000016)	0.000043 (0.000013)	0.000075 (0.000049)	0.000063 (0.000030)	0.000050 (0.000019)	0.000059 (0.000021)
Slow-decay, Variable Catchment SPAR	1.00 (0.49)	0.95 (0.42)	0.60 (0.29)	0.98 (0.37)	1.03 (0.26)	1.62 (1.02)	1.39 (0.68)	0.95 (0.36)	1.12 (0.44)
Fast-decay, Variable Catchment SPAR	1.00 (0.55)	0.87 (0.47)	0.69 (0.33)	0.91 (0.52	1.17 (0.39)	1.21 (0.95)	1.39 (0.87)	0.78 (0.40)	1.04 (0.54)

2SFCA, Two-Step Floating Catchment Area; *SPAR*, Spatial Access Ratio
* Not siginificant at $p < 0.05$; all other Appalachia and Non-Appalachia comparisons were significant at $P < 0.01$.

CHAPTER 3

Spatial Access to Primary Care Providers in Appalachia: Evaluating Current Methodology

3.1 Abstract

Purpose

The goal of this research was to examine spatial access to primary care physicians in Appalachia using both traditional access measures and the two-step floating catchment area (2SFCA) method. Spatial Access to care was compared between urban and rural regions of Appalachia.

Methods

The study region included Appalachia counties of Pennsylvania, Ohio, Kentucky, and North Carolina. Primary care physicians during 2008 and total census block group populations were geocoded into GIS software. Ratios of county physicians to population, driving time to nearest physician, and various 2SFCA approaches were compared.

Results

Urban areas of the study region had shorter travel times to their closest primary care physician. Provider to population ratios produced results that varied widely from one county to another due to strict geographic boundaries. The 2SFCA method produced varied results depending on the chosen distance decay weight and variable catchment size technique. 2SFCA scores showed greater access to care in urban areas of PA, OH, and NC.

Conclusion

The different parameters of the 2SFCA method—distance decay weight and variable catchment size—have a large impact on the resulting spatial access to primary care scores. We recommend using a relative 2SFCA approach, the spatial access ratio (SPAR) method, when detailed patient travel data are unavailable. The 2SFCA method shows promise for measuring access to care in Appalachia, but more research on patient travel preferences is needed to inform implementation.

3.2 Introduction

Appalachia is a mountainous region that spans the northern portions of Mississippi and Alabama in the south and reaches into the southern portion of New York at its northern end. Appalachia is largely rural, with 42% of its population classified as rural compared to the national average of 20% (Appalachia Regional Commission, 2014). Socioeconomically, the region has a lower per capita income and a higher poverty rate than the national average (Pollard & Jacobson, 2014). Access to adequate healthcare is an ongoing concern in Appalachia, largely due to the mountainous terrain, rural population distribution, and socioeconomic disparities (Lengerich et al., 2005).

Access to primary care services is especially important given primary care providers' role as a gateway to health systems (Starfield et al., 2005). In Appalachia, regular primary care encounters have been shown to increase early cancer detection and reduce mortality (Camacho et al., 2014). In Appalachia Ohio, children with irregular primary care visits had poorer general health outcomes, and the parents of those children reported that lack of access to primary care prevented regular contact (Smith & Holloman, 2011). Regular, quality primary care encounters can also counteract the negative effects that economic disparities have on heath (Shi, 2012), a particularly important outcome given Appalachia's generally reduced economic status.

Accurately measuring access to primary care is important in Appalachia. Spatial access is one component of access to care, distinct from non-spatial factors such as insurance status or level of education (Wang & Luo, 2005). One traditional measure of spatial access is a county provider to population ratio. The technique is self-explanatory, deriving a ratio of practicing healthcare providers to the total county population, often drawn from U.S. Census data. The U.S. Department of Health and Human Services (DHHS) relies on this technique when designating Health Professional Shortage Areas (HPSA; DHHS, 2010). Another common strategy for measuring spatial access to care is to use geographic information systems (GIS) software to calculate the shortest travel time between population points and healthcare providers. Instead of generalizing across an entire geographic area, shortest travel time measures focus more specifically on a smaller population unit, often census tracts, census block groups, or individual patients (Wang, 2012).

There are noted limitations to these standard measures of spatial access to care (Guagliardo, 2004). The geographic boundaries used in provider to population ratios are not representative of typical healthcare use. Patients often cross county boundaries, and providers within a county boundary may not actually be accessible, especially in large, rural counties (Wang, 2012). Travel time calculations overcome the latter problem of travel cost between patients and providers. However, travel time measures do not account for the supply and demand factors that impact healthcare (Luo & Wang, 2003).

A more recent method for measuring spatial access to care is the two step floating catchment area (2SFCA) method, which was designed to overcome some of the limitations of provider to population ratios and travel time (Luo & Wang, 2003; Wang & Luo, 2005). The 2SFCA method begins by setting a catchment area (usually 30 or 60 minutes) around each healthcare provider and identifying all the populations within that provider's catchment. A provider to population ratio is then calculated for each provider's catchment. In step two, each population becomes the center of a catchment, and the step one ratios associated with each provider in that population's catchment are summed.

Several additional parameters have been added to the original 2SFCA framework (Luo & Qi, 2009; McGrail & Humphreys, 2009). Catchments at both steps can be further refined by adding a distance decay function, where distinct zones (e.g., 0-10 min., 10-20 min, 20-30 min.) each receive unique weights (Luo & Qi, 2009). Instead of separate zones, a continuous function can also be applied to travel distance within catchments (McGrail & Humphreys, 2009). Another improvement added to the original 2SFCA method is the technique of varying catchment sizes at both step one and two. This approach attempts to capture the fact that different healthcare providers serve patients at varying geographic distances, such as in rural compared to urban settings (McGrail, 2012). Thus, catchment sizes should not remain uniform across all providers and populations, and can be limited by capping the number of providers that a population can access, for example (McGrail & Humphreys, 2009). Lastly, owing to the uncertainty in how to weight travel within catchments, or how exactly to vary catchments, a relative approach called the Spatial Access Ratio (SPAR) was created (Wan et al., 2012). This technique divides a population's 2SFCA score by the mean 2SFCA score of all populations, which researchers demonstrated minimizes the differences in 2SFCA scores resulting from using different decay weights (Wan et al., 2012).

Despite its theoretical advantages and its empirical validity in predicting clinical outcomes (Dai, 2010; Lian et al., 2012), the 2SFCA method has never been used to study access to primary care in Appalachia. The primary focus of this study was to examine spatial access to primary care in Appalachia using both the traditional measures of provider to population ratios and closest travel time, as well as the 2SFCA method and its various parameters. Within Appalachia, results were compared between urban and rural areas.

3.3 Study Area and Data

The study examined the Appalachia regions of Pennsylvania, Ohio, Kentucky, and North Carolina. Appalachia counties were determined using the Appalachia Regional Commission's (ARC) county designations. Primary care physicians were derived from the 2008 American Medical Association (AMA) Physician Masterfile, using specialties of Family Practice, General Practice, Internal Medicine, and General Pediatrics (Camacho et al., 2014). Office addresses at the street level were available for 8,039 of the 9,483 physicians in the study area. The population weighted centroid of the physician's office census tract was used for the remaining 1,444 physicians. Primary care physicians from neighboring states and neighboring in-state, non-Appalachia areas were not included, creating possible edge effects.

The 2010 U.S. Census was used to derive population data at the census block group level. Block groups are ideally comprised of 1,500 people, with a range from 600 to 3,000. The population weighted centroid (U.S. Census Bureau, 2014) of each block group was used as the geographic reference, similar to previous research (Luo & Wang, 2003). Many previous studies measuring spatial access to care used the larger geographic areas of zip codes (Dai, 2010) or census tracts (Wan et al., 2012), which provide less geographic specificity than the smaller census block groups. We did include neighboring block groups within one hour travel time of our Appalachia study area. There were 8,721 populated block groups in Appalachia regions of the four states, resulting in a total population of 10,717,421 people.

Within Appalachia, populations were also dichotomized as urban and rural. Designation occurs at the census block level by assigning RUCU codes 7-10 as rural and RUCU codes 1-6 as urban (Weeks et al., 2004). Codes 1-6 also include suburban areas, but these areas were grouped as urban for ease of interpretation (Weeks et al., 2004). Descriptive statistics and t-tests were used to examine the differences between rural and urban regions.

3.4 Methodology

Provider to population ratios

Primary care provider to population ratios were calculated at the county level. The total primary care providers in each county were divided by that counties total population. Although primary care data was current as of 2008, and population data was from the 2010 Census, the 2010 Census offered a closer approximation than 2000 Census data.

Travel Time

Primary care providers and census block groups were geocoded into ArcGIS (Version 10.1, ESRI Inc., Redlands, CA). The Origin-Destination (OD) cost matrix function of the ArcGIS Network Analyst extension was used to determine travel times between the population weighted centroid of each block group and the closest primary care provider. Similar to previous research (Wang, 2012), and corresponding with the maximum time used in the 2SFCA method, a maximum travel time of 60 minutes was set.

Two-step floating catchment area

The 2SFCA method begins by identifying all populations (k) within a service provider's (j) maximum catchment (d_o), which we set as 60 minutes driving time. The rural characteristics of Appalachia necessitated the 60 minute maximum time, rather than a 30 minute maximum

(McGrail & Humphreys, 2009). After identifying all populations, a population to provider ratio (R_j) is then computed using the total service providers (S_j) and the total population (P_k) within catchment j:

Step 1: $R_j = S_j / \sum k \in \{d_{jk} < d_o\} P_k$

The next step of the 2SFCA method searches for all service providers (j) within the maximum travel time (d_o) of a population location (i). The sum of all the eligible service providers' step 1 ratios (R_j) is then calculated, resulting in that population location's access score (A_i):

Step 2: $A_i = \sum j \in \{d_{ij} < d_o\} R_j$

Several additional versions of the 2SFCA were also considered. A distance decay function was added across four time zones within catchments, corresponding to travel times of 0-10 minutes, 11-20 minutes, 21-30 minutes, and 31-60 minutes (Luo & Qi, 2009). A different weight was applied at each zone, reflecting the travel cost associated with increasing time. Two weighting schemes were compared (McGrail, 2012; Wan et al., 2010):

- Fast-decay weights (w) = 1, 0.60, 0.25, 0.05
- Slow-decay weights (w) = 1, 0.80, 0.55, 0.15

Weights are applied at each step and correspond to the 0-10, 11-20, 21-30, and 31-60 minute time zones within each catchment. Thus, during step one all populations (k) are reduced by the above weights, depending on the travel time to that catchment's provider (j). In step two, each provider ratio (R_j) is reduced by its corresponding weight. This yields the following update step one and two equations:

- Step 1: $R_j = S_j / \sum k \in \{d_{jk} < d_o\} P_k\, w$
- Step 2: $A_i = \sum j \in \{d_{ij} < d_o\} R_j\, w$

Another 2SFCA version was compared that included both distance decay weights and varying catchment sizes. The varying catchment method used a different technique for step one and two of the 2SFCA (McGrail, 2012). Local population distributions were used in step one to decide whether or not the distance decay weights were applied, thereby expanding or contracting the provider's catchment. Populations within the first 10 minutes of a provider had no decay weight. If a provider was one of the closest 25 providers to a population it was assumed that population would travel because it had fewer choices, thus no decay weight was applied to that population. Also, no decay weight was applied if the provider's town had a population of 5,000 or greater and the population's town had a total population of less than half the provider's. This scenario attempts to capture smaller towns that will travel to access providers in larger nearby communities (McGrail & Humphreys, 2009b). Primary care providers and block groups were linked to their closest Census Designated Place (CDP) to estimate the population of their closest towns. The Census Bureau describes CDPs as representative of a settled concentration of populations in a city, town, village, or borough (U.S. Census Bureau, 2014b).

In determining step two catchments of the 2SFCA, block groups were only allowed to access their closest 100 primary care providers, simulating the point at which additional providers are unlikely to be utilized. The first 100 providers, however, would have distance decay weights applied as usual. For the distance decay and variable catchment size approach, a continuous distance decay weighting scheme was also applied to contrast with the fast and slow weighting scheme (McGrail & Humphreys, 2009). Thus, the same rules as above were applied, but when a decay weight was to be used for travel times (d) between 10 and 60 minutes, the following weight (w) was applied:

- Continuous-decay weight (w) = $((60-d) / (60-10))$ ^ 1.5

The last 2SFCA method evaluated in our study region was the spatial access ratio (SPAR) score (Wan et al., 2012). Each of the other 2SFCA methods makes assumptions about patient healthcare behavior when assigning certain weights or capping a catchment at a certain number of providers. SPAR was created to minimize the subjectivity when choosing distance decay weights and varying catchment sizes. A SPAR score is given by dividing a census block group's 2SFCA score by the mean 2SFCA score of all the block groups in the entire study area (Wan et al., 2012).

A total of 10 spatial access scores were applied to each census block group in the Appalachia regions in PA, OH, KY, and NC: 1) Ratio of county primary care providers to total county population; 2) Travel time to closest primary care provider; 3) Original 2SFCA score; 4) Slow-decay 2SFCA score; 5) Fast-decay 2SFCA score; 6) Slow-decay and variable catchment 2SFCA score; 7) Fast-decay and variable catchment 2SFCA score; 8) Continuous-decay and variable catchment 2SFCA score; 9) Slow-decay and variable catchment SPAR score; 10) Fast-decay and variable catchment SPAR score.

3.5 Results

The greater Pittsburgh area had the highest population density of our four state Appalachia region (Fig. 3.1). The Youngstown area along the PA-OH border, the Erie area in northwest PA, and the Greensboro and Ashland areas in NC all also had among the most populated counties in our study region. North-central PA and eastern KY were the least populated areas.

All spearman nonparametric correlations were significant between the 10 spatial access measures, except for the correlation between travel time and continuous-decay, variable catchment scores (Table 3.1). For reference, the provider to population ratios, 2SFCA scores, and SPAR scores are all similar in that increasing values signify greater spatial access. Travel time is opposite, where lower times indicate greater spatial access, which is why travel time was negatively correlated with each of the other measures. There was a marked difference between the 2SFCA measures that included a variable catchment size function and those that did not. For example, provider to population ratios and travel time were more strongly correlated with the original, slow, and fast-decay 2SFCA scores ($0.426 \leq$ rho ≤ 0.765) than with the slow, fast, and continuous-decay 2SFCA scores that also had varying catchment sizes ($0.017 \leq$ rho ≤ 0.34). The correlations within these groups were strong—the original, slow, and fast-decay 2SFCA scores between 0.719 and 0.970 and the slow, fast, and continuous-decay 2SFCA scores with varying catchment sizes between 0.805 and 0.972—but the correlation between groups was not as strong ($-0.174 \leq$ rho ≤ 0.463). As expected, the correlations were identical between the slow and fast-decay 2SFCA scores with varying catchment sizes and their corresponding SPAR scores.

Several of the areas with the greatest population density also had the highest ratio of primary care physicians to county population, including the Pittsburgh, Ashland, and Greensboro areas (Fig. 3.2a). There were also rural regions with higher ratios, such as in eastern KY, northeastern PA, and eastern OH. Many of the counties with ratios in the highest quintile were bordering counties with ratios in the lowest quintile, demonstrating the impact that county boundary lines can have when using provider to population ratios.

The majority of census block groups were within a 10 minute drive time of their closest primary care physician (Fig. 3.2b). The regions with the largest travel times were generally those with the lowest population density, as shown in Figure 3.1. Block groups with the largest travel times were found throughout north central PA, the entire Appalachia region of KY, and western

NC. It is difficult to interpret the presence of the largest travel times (i.e., the group of 41-60 minutes) in the far eastern part of our KY and NC regions because those areas border regions where we did not have access to primary care data. Thus, it is possible that those region's true time to their closest primary care physician is less than shown.

The census block groups with the highest original 2SFCA scores were located primarily in the Pittsburgh area (Fig. 3.3a). After adding distance decay weights, 2SFCA scores were reduced in the suburban areas surrounding Pittsburg but increased in several other population centers of our study region, including near the Greensboro, Ashland, and Youngstown areas (Fig. 3.3b, 3.3c). Overall, decay weights appear to have dispersed scores, rather than concentrating them across one large, adjoining area, as with the original 2SFCA scores. This was especially true for the fast-decay weighting scheme (Fig. 3.3c). Adding variable catchment sizes with decay weights drastically altered the distribution of spatial access 2SFCA scores (Fig. 3.3d, 3.3e, 3.3f).

Across each measure that varied catchment sizes—slow, fast, and continuous decay weightings—the Pittsburg area went from having among the highest scores to either mediocre or poor scores. For reference, there were roughly 2,900 primary care providers within a 1 hour drive of the urban core of Pittsburgh. When variable catchments were introduced, those urban populations had their possible number of primary care services capped at 100, resulting in the significant drop in scores. The block groups that saw comparatively higher scores, conversely, were largely in the smaller town areas of central PA, southern OH, eastern KY, and western NC. In the southern OH town of Portsmouth, for example, there were approximately 190 primary care physicians within a 1 hour drive. Consequently, capping the number of providers had a much smaller effect than in Pittsburg. Yet, because the population density is lower (approximately

20,000 compared to 300,000 in Pittsburgh), the 2SFCA step one ratio scores for Portsmouth and similar small towns were higher, resulting in higher overall 2SFCA scores.

The difference between slow and fast-decay variable catchment 2SFCA scores was similar to the difference when only decay weights and no variable catchment function were used. Fast-decay weights resulted in tighter groupings of higher access block groups at population centers, rather than the slow-decay weight scenario of high access block groups sprawled across population centers and surrounding suburbs (Fig. 3.3d, 3.3e). When a continuous decay weighting scheme was used with varying catchment sizes, the result was nearly identical to the slow-decay, variable catchment size approach (Fig. 3.3d, 3.3f). The slow and fast-decay, variable catchment SPAR scores demonstrated the same differences as their corresponding 2SFCA scores (Fig. 3.3g, 3.3h). Figure 3.4 compares the change in actual 2SFCA scores across rural and urban areas when using slow and fast decay weights. Urban areas were responsible for the highest scores, regardless of the 2SFCA iteration used. After using only the distance decay technique, scores generally did not uniformly increase or decrease according to the weight used (Fig. 3.4a, 3.4b). The exception was for the highest scoring urban block groups, where the fast-decay weights universally increased scores (Fig. 3.4b). After adding variable catchment sizes and decay weights, however, a clear pattern emerged where fast-decay weights decreased scores for most block groups (Fig. 3.4c, 3.4d). The results when using the decay-weighted and variable catchment SPAR scores were mixed (Fig. 3.4e, 3.4f). For block groups classified as urban (7,266 of the 8,721 block groups were classified as urban), the SPAR technique achieved it intended effect of preventing any noticeable increase or decrease in scores based on decay weighting (Fig. 3.4f). The SPAR technique was less effective for rural block groups, where fast-decay weights uniformly decreased scores (Fig. 3.4e).

Table 3.2 compares the rural and urban mean spatial access scores across each state of the study region. Urban block groups in PA, OH, and NC had higher provider to population ratios, while higher ratios came from rural block groups in KY. Travel time to the closest primary care provider was lower for urban block groups across each state. The mean scores for the entire region demonstrate that the addition of a variable catchment size decreased scores considerably, regardless of the decay weighting scheme chosen. In PA and NC, urban areas had larger 2SFCA scores except for the slow and continuous decay, variable catchment scores, where rural areas performed better. Ohio's urban block groups performed even better, scoring higher than rural block groups across each 2SFCA iteration. Scores in KY were largely reversed, where rural block groups performed better across each 2SFCA method except the fast-decay and variable catchment size approach.

3.6 Discussion

This research presented the first comparison in Appalachia of the most recently developed spatial access to care methods. Based on nonparametric correlations, the 2SFCA measures that included variable catchment sizes appeared distinct from the measures of provider to population ratios, closest travel time, and the 2SFCA method with and without decay weights. When comparing provider to population ratios across the study region, the most populated urban areas (e.g., Pittsburgh, Greensboro) had the highest ratios. Several rural areas had high ratios as well, but abrupt county boundaries made interpretation difficult because several adjacent counties had ratios in the highest and lowest quintile groups. The majority of census block groups were within 10 minutes of their closest primary care provider. Rural block groups in each state faced longer travel times than urban block groups. The urban-rural travel time was most similar in KY, where times were higher than average for both groups. Urban block groups in KY, for instance, had longer travel times (9.54 min.) than the rural block groups of PA (7.50 min.) and OH (8.26 min.). Mapping the different 2SFCA scores across the study region revealed differences—similar to the correlation analysis—between the 2SFCA approaches that included variable catchment sizes and those that did not. The variable catchment size method resulted in several urban areas, including the greater Pittsburgh area, having among the worst access scores. When catchment sizes remained static, and corresponding populations of these urban areas had no cap on the number of physicians they could access, these same urban areas had among the highest 2SFCA access scores.

Traditionally, when the original 2SFCA method was applied to primary care physicians, urban areas often received the highest scores (Luo & Wang, 2003; Wang & Luo, 2005). This also proved true with our results in Appalachia. Pittsburgh was the largest urban area in the study region, and it had the highest original 2SFCA scores. When distance-decay weights were added, Pittsburgh's core maintained comparatively high access scores, but the surrounding suburbs were reduced to more mediocre scores. Those surrounding populations, especially those with a 30-60 minute drive from the center of Pittsburgh, were no longer able to count the abundance of central Pittsburgh physicians as accessible. This pattern matches the intended effect behind distance-decay functions, which was partly to limit accessibility at the boundaries of physician catchments (Luo & Qi, 2009). After adding distance-decay weights, the same effect occurred across other areas where original 2SFCA scores were high across a contiguous area, such as near Youngstown, OH and along the Williamsport to Scranton corridor in central and northeastern PA.

Adding variable catchment sizes to the slow and fast decay 2SFCA approach produced dramatic differences in the distribution of access scores. Pittsburgh's scores ranked among the lowest of the study region. Smaller rural towns in eastern KY, southern OH, western NC, and central PA—who without variable catchments only had mediocre scores—had among the highest scores in the study region. The variable catchment approach we used capped the number of physicians a population could access at 100 (McGrail & Humphreys, 2009). Most towns of 15,000 people or greater were within a one hour drive of 100 primary care physicians. But, the population demand on those physicians was less in rural areas, resulting in the increased scores relative to the major urban area of Pittsburgh.

Although theoretically similar, the distance-decay weights chosen in this study did impact the distribution of the highest and lowest access scores. Fast decay weights, when used independently and when used in conjunction with variable catchment sizes, resulted in a more concentrated pattern of high scores in town centers compared to slow and continuous weighting schemes. The SPAR method was created in part to mitigate those differences (Wan et al., 2012). For the 83% of census block groups classified as urban (this included suburban areas as well, for ease of interpretation), changing from the slow decay, variable catchment method to the fast decay, variable catchment method did not result in any uniform increase or decrease in mean access scores. For rural block groups, however, SPAR did not have its intended effect as nearly all scores decreased when changing from the slow decay, variable catchment method to the fast decay, variable catchment method. Wan et al. (2012) did not include variable catchment sizes in their SPAR models, and it appears that the relative smoothing effect of SPAR was not enough to counteract the effect of the fast decay, variable catchment approach of rural block groups' scores.

Indeed, when only including decay weights and not variable catchment sizes, the SPAR technique did stabilize slow and fast decay scores in rural areas (data not shown).

This research had several strengths. It represents the first application of the 2SFCA method to study access to primary care in Appalachia. Public health research in Appalachia is increasingly utilizing GIS techniques (Anderson et al., 2014), which makes it imperative that the most current spatial access methods are evaluated. A methodological strength in this study was the increased geographic specificity as a result of using census block groups as the geographic population area, rather than the larger census tracts or zip codes that other research applying the 2SFCA method has employed (Diad, 2010; Lian et al., 2012; Wan et al., 2012). Another methodological strength was the inclusion of variable catchment sizes, which allowed the comparison between identical distance decay 2SFCA measures either with or without varying catchments, instead of only evaluating decay weights (Mao & Nekorchuk, 2013; Tao et al., 2014).

There were also several limitations to this research. From a methodological perspective, not including neighboring primary care physicians likely created edge effects along the border of our study region. Another limitation, common in much research examining potential spatial access, was the lack of actual patient healthcare utilization behaviors. The application of distance-decay weights and variable catchment sizes is theoretically sound. Grounding distance decay and variable catchment functions in empirical data is more difficult. The specific parameters of these 2SFCA iterations need to be evaluated against patient healthcare behavior, preferably at the local level of a study region.

In summary, we recommend the SPAR technique for use when patients' healthcare travelling data is unavailable. Including variable catchment sizes is a theoretical improvement to the 2SFCA method, but choosing optimal parameters to vary catchment sizes by is difficult. We recommend that researchers compare several parameters (e.g., capping the number of physicians a population can access at higher and lower values than the standard approach of 100) and contrast the effects with 2SFCA methods that only employ distance decay weights.

3.7 References

Anderson RT, Yang TC, Matthews SA, Camacho F et al. (2014). Breast Cancer Screening, Area Deprivation, and Later-Stage Breast Cancer in Appalachia: Does Geography Matter?. Health Services Research, 49: 546–567.

Appalachia Regional Commision (ARC) The Appalachia region. [Accessed October 10, 2014]. Available at: http://www.arc.gov/appalachian_region/TheAppalachianRegion.asp

Dai D (2010). Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. *Health Place*. 16(5), 1038–1052.

Camacho F, Hwang W, Kern T, & Anderson RT (2014). Receipt of Regular Primary Care and Early Cancer Detection in Appalachia. *The Journal of Rural Health*. doi: 10.1111/jrh.12097.

Department of Health and Human Services (2008). Shortage Designation: Health Professional Shortage Areas & Medically Underserved Areas/Populations [Accessed October 20, 2014]. Available at: http://www.hrsa.gov/shortage/

Guagliardo MF (2004). Spatial accessibility of primary care: concepts, methods and challenges. *Int J Health Geogr.* 3:3.

Lengerich EJ, Tucker TC, Powell RK et al. (2005). Cancer Incidence in Kentucky, Pennsylvania, and West Virginia: Disparities in Appalachia. *Journal of Rural Health.* 21(1): 39–47.

Lian M, Struthers J, Schootman M (2012). Comparing GIS-Based Measures in Access to Mammography and Their Validity in Predicting Neighborhood Risk of Late-Stage Breast Cancer. PLoS ONE 7(8): e43000. doi:10.1371/journal.pone.0043000.

Luo W & Qi Y (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health Place.* 15:1100–1107.

Luo, W & Wang F (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B.* 30(6): 865–884.

Mao, L. & Nekorchuk, D. (2013). Measuring spatial accessibility to healthcare for populations with multiple transportation modes. Health Place. 24:115-122.

McGrail MR (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *Int J Health Geogr*. 16:11-50.

McGrail MR & Humphreys JS (2009). The index of rural access: an innovative integrated approach for measuring primary care access. *BMC Health Serv Res*. 9:124.

McGrail MR & Humphreys JS (2009b). A new index of access to primary care services in rural areas. *Aust NZ J Pub Health.* 33:418–423.

Pollard K & Jacobson LA (2014) The Appalachian Region: A Data Overview from the 2008-2012 American Community Survey. *Appalachia Regional Commission.* Retrieved from http://www.arc.gov/research/researchreportdetails.asp?REPORT_ID=109.

Shi L (2012). The Impact of Primary Care: A Focused Review. *Scientifica (Cairo)*. 432892. doi:10.6064/2012/432892

Smith & Hollomon, 2011. Health Status and Access to Health Care Services: A comparison between Ohio's rural Non-Appalachian and Appalachian families. *Fam Community Health.* 34(2): 102-110.

Starfield B, Shi L, & Macinko J (2005). Contribution of Primary Care to Health Systems and Health. *The Milbank Quarterly*. 83(3):457-502.

Tao Z, Cheng Y, Dai T, & Rosenberg MW. (2014). Spatial optimization of residential care facility locations in Beijing, China: maximum equity in accessibility. *International Journal of Health Geographics.* 14:13-33.

U.S. Census Bureau (2014). Centers of Population [accessed June 5, 2014]. Available at: https://www.census.gov/geo/reference/centersofpop.html

U.S. Census Bureau (2014b). Geographic Terms and Concepts – Place [Accessed June 5, 2014]. Available at: https://www.census.gov/geo/reference/gtc/gtc_place.html

Wan N, Zhan FB, Zou B, & Chow E (2012). A relative spatial access assessment approach for analyzing potential spatial access to colorectal cancer services in Texas. *App Geography*. 12:291–299.

Wang F (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Ann Assoc Am Geogr.* 102(5): 1104–1112.

Wang F & Luo W (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health Place*. 11:131–146.

Weeks WB, Kazis LE, Shen Y, et al. (2004). Differences in Health-Related Quality of Life in Rural and Urban Veterans. *American Journal of Public Health.* 94(10):1762-1767.

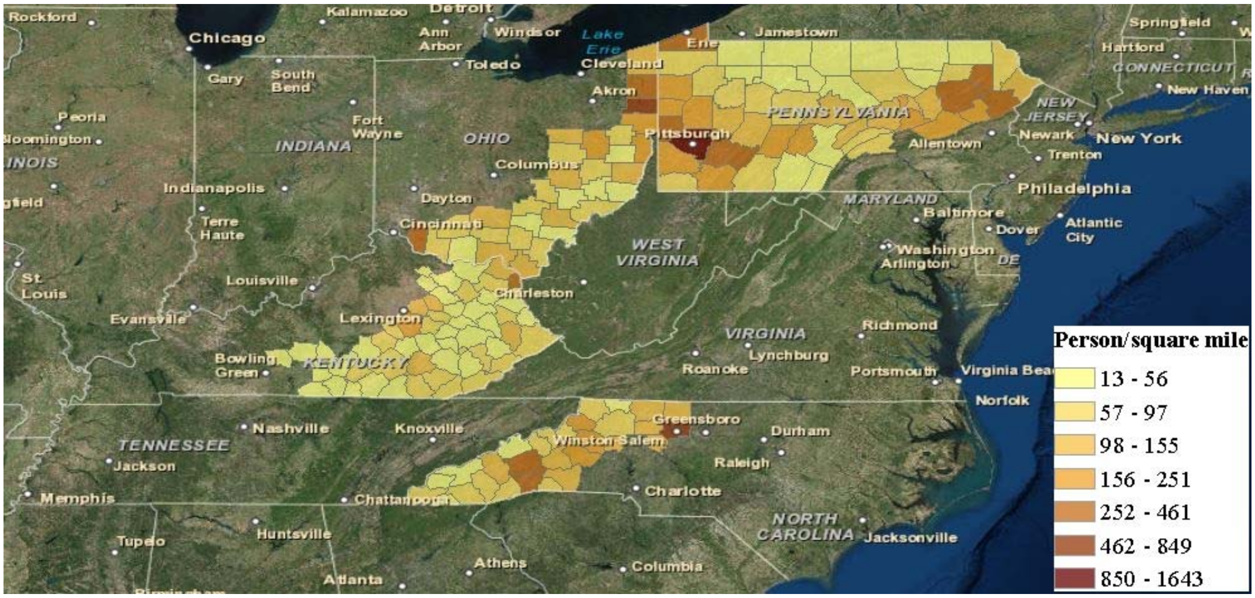

Figure 3.1 Population densities from the 2010 U.S. Census in Appalachia counties of PA, OH, KY, and NC.

Table 3.1 Spearman correlations* between 10 spatial access to prmiary care measures across Appalachia PA, OH, KY, and NC

	Prov. - Pop. Ratio	Travel Time	Original 2SFCA	Slow 2SFCA	Fast 2SFCA	Slow - Var. Catch. 2SFCA	Fast - Var. Catch. 2SFCA	Cont. - Var. Catch. 2SFCA	Slow - Var. Catch. SPAR	Fast - Var. Catch. SPAR
Prov. - Pop. Ratio	1									
Travel Time	-0.381	1								
Original 2SFCA	0.564	-0.426	1							
Slow 2SFCA	0.765	-0.566	0.833	1						
Fast 2SFCA	0.760	-0.626	0.719	0.970	1					
Slow - Var. Catch. 2SFCA [a]	-0.046	-0.045	-0.115	0.056	0.124	1				
Fast - Var. Catch. 2SFCA [b]	0.198	-0.342	0.150	0.374	0.463	0.874	1			
Cont. - Var. Catch. 2SFCA [c]	-0.117	0.017[+]	-0.174	-0.043	0.023	0.972	0.805	1		
Slow - Var. Catch. SPAR	-0.046	-0.045	-0.115	0.056	0.124	1.000	0.874	0.972	1	
Fast - Var. Catch. SPAR	0.198	-0.342	0.150	0.374	0.463	0.874	1.000	0.805	0.874	1

2SFCA, Two-Step Floating Catchment Area; *SPAR*, Spatial Access Ratio

* All correlations significant at the 0.05 level, unless otherwise noted

[+] Not significant

[a] Slow - Var. Catch. 2SFCA: Both slow-decay weightings and varying catchment size rules

[b] Fast - Var. Catch. 2SFCA: Both fast-decay weightings and varying catchment size rules

[c] Cont. - Var. Catch. 2SFCA: Both continuous-decay weightings and varying catchment size rules

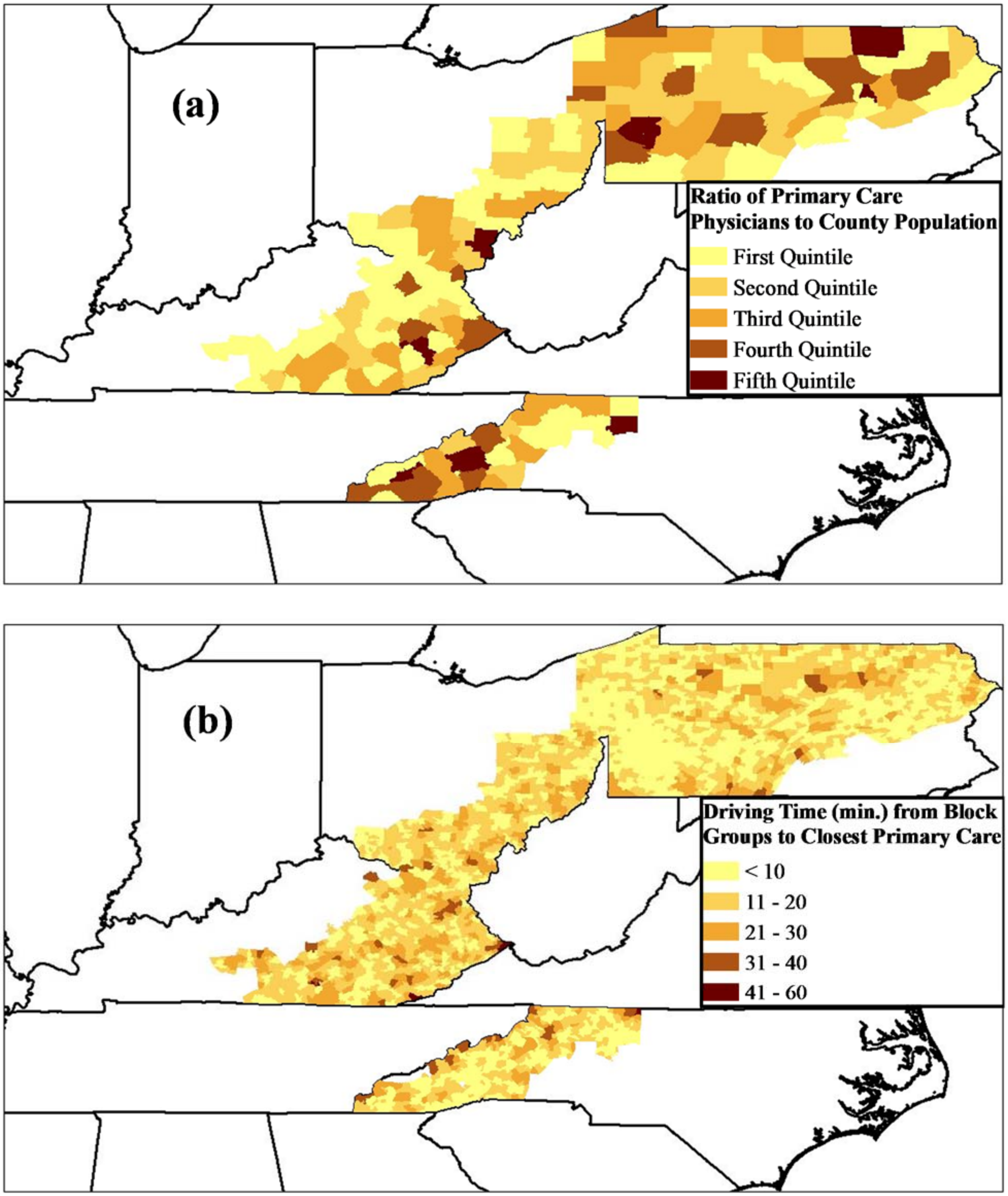

Figure 3.2. Spatial access to primary care providers by (a) county provider to population ratio, broken into quintiles, and (b) closest driving time from census block groups

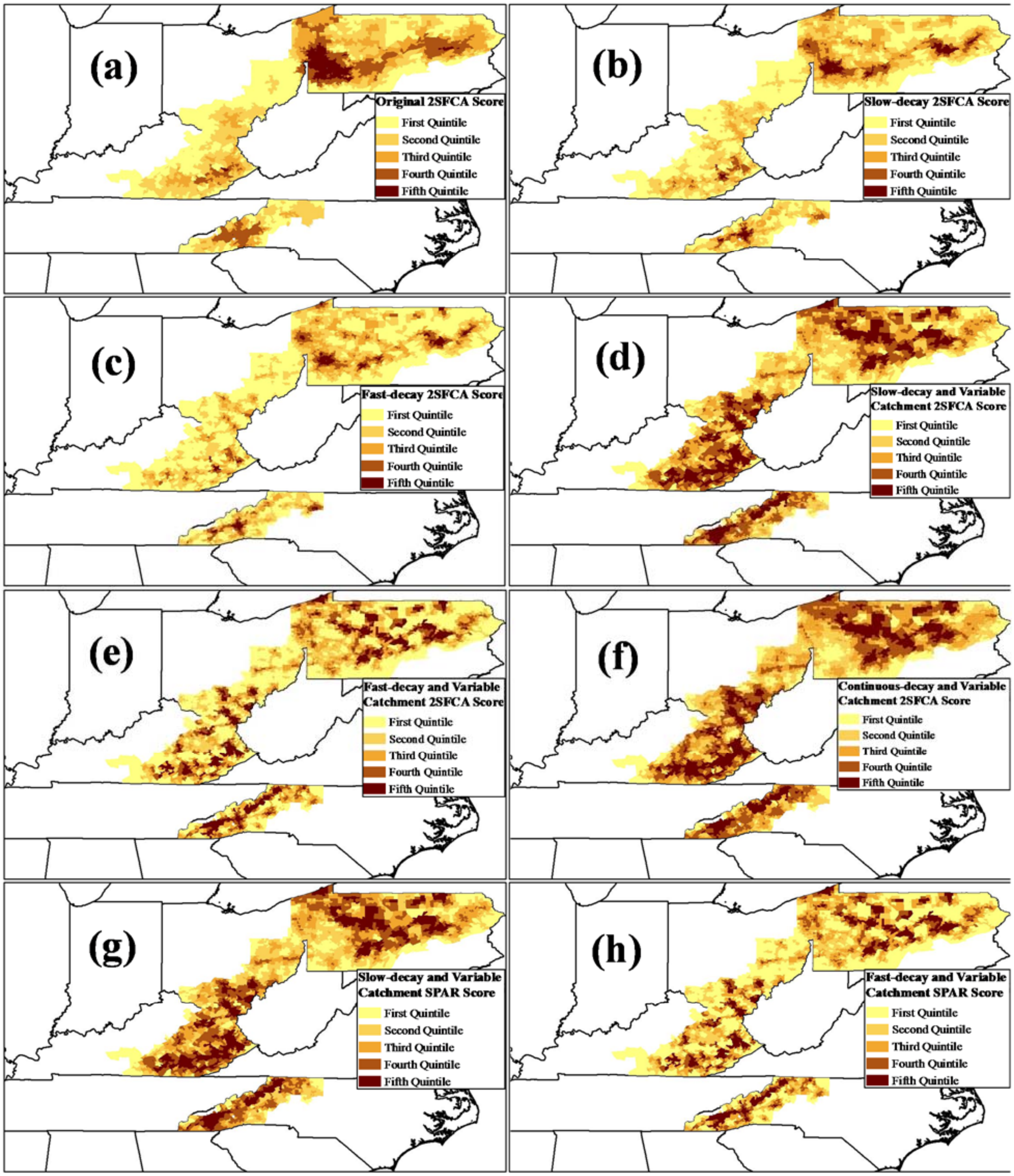

Figure 3.3. Spatial access of census block groups in Appalachia PA, OH, KY, and NC toprimary care providers by (a) original 2SFCA scores; (b) slow-decay 2SFCA scores; (c) fast-decay 2SFCA scores; (d) slow-decay and variable catchment 2SFCA scores; (e) fast-decay and variable catchment 2SFCA scores; (f) continuous-decay and variable catchment 2SFCA scores; (g) slow-decay and variable catchment SPAR scores; (h) fast-decay and variable catchment SPAR scores. Scores are broken into quintiles, with larger quintiles representing greater spatial access.

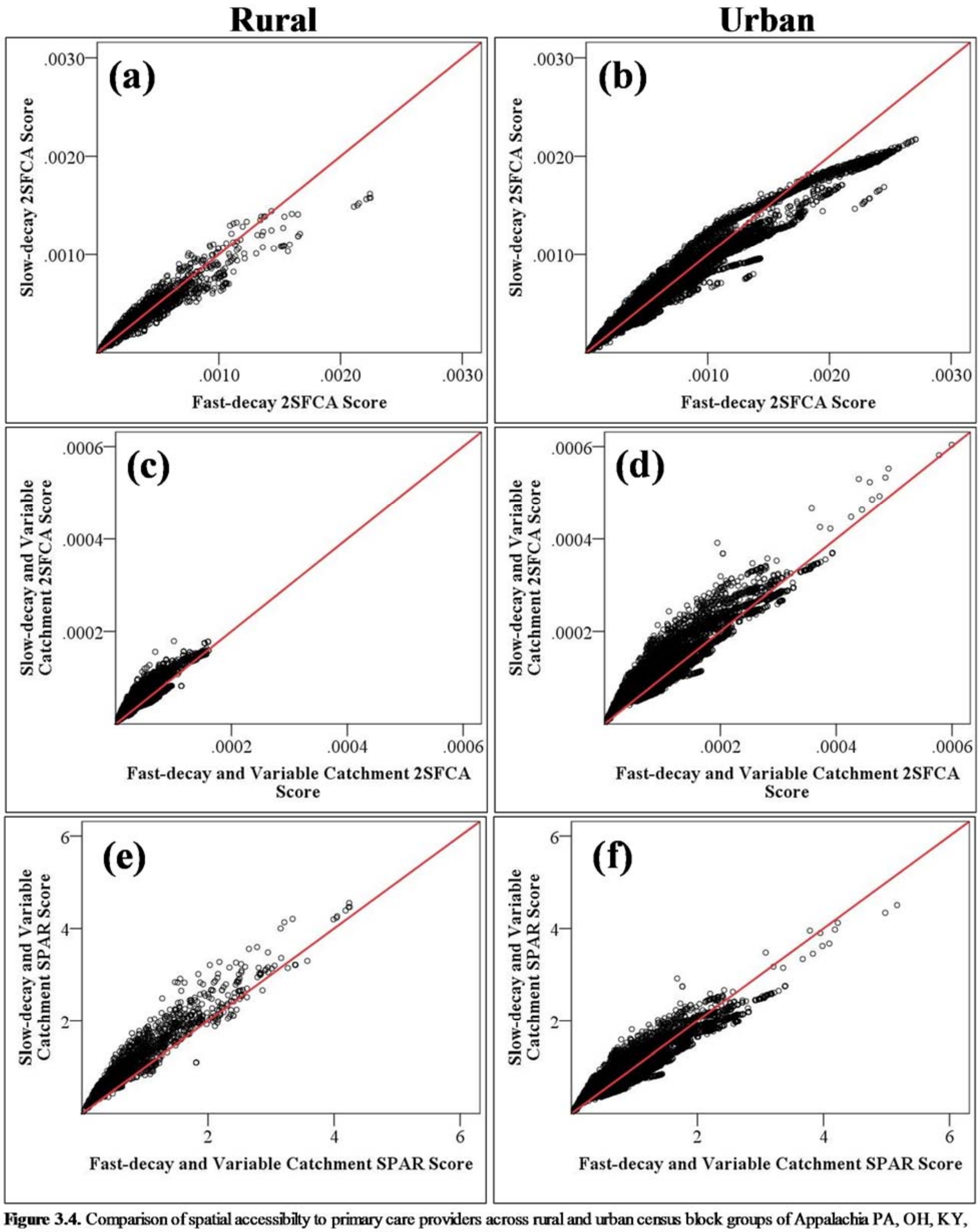

Figure 3.4. Comparison of spatial accessibilty to primary care providers across rural and urban census block groups of Appalachia PA, OH, KY, and NC between (a-b) slow-decay and fast-decay 2SFCA scores; (c-d) Slow-decay, variable catchmentand and fast-decay, variable catchment 2SFCA scores; (e-f) Slow-decay, variable catchment and fast-decay, variable catchment SPAR scores.

Table 3.2. Descriptive statistics of primary care spatial access scores between rural and urban census block groups of Applachia PA, OH, KY, and NC.

Spatial Access Measures	All States	Pennsylvania		Ohio		Kentucky		North Carolina	
	Mean (SD)	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Provider to Population Ratio	0.000880 (0.000542)	0.000710 (0.000347)	0.001031 (0.000637)	0.000481 (0.000209)	0.000708 (0.000301)	0.000636 (0.000304)	0.000553 (0.000257)	0.000747 (0.000307)	0.000970 (0.000467)
Travel Time to Closest Mammograpy Center (min.)	5.81 (6.31)	7.50 (6.92)	4.03 (4.77)	8.26 (7.29)	5.88 (5.98)	11.89 (8.81)	9.54 (7.79)	10.62 (7.64)	6.122 (5.79)
Origianl 2SFCA	0.000706 (0.000382)	0.000507 (0.000238)	0.000936 (0.000351)	0.000321 (0.000195)	0.000465 (0.000272)	0.000470 (0.000232)	0.000410 (0.000174)	0.000370 (0.000160)	0.000583 (0.000265)
Slow-decay 2SFCA	0.000768 (0.000490)	0.000477 (0.000245)	0.001002 (0.000518)	0.000323 (0.000161)	0.000529 (0.000320)	0.000501 (0.000306)	0.000435 (0.000223)	0.000416 (0.000190)	0.000724 (0.000378)
Fast-decay 2SFCA	0.000804 (0.000573)	0.000472 (0.000260)	0.001032 (0.000629)	0.000342 (0.000186)	0.000579 (0.000389)	0.000528 (0.000398)	0.000451 (0.000265)	0.000455 (0.000254)	0.000807 (0.000480)
Slow-decay, variable catchment 2SFCA	0.000134 (0.000078)	0.000154 (0.000075)	0.000122 (0.000074)	0.000114 (0.000056)	0.000125 (0.000055)	0.000191 (0.000115)	0.000177 (0.000083)	0.000183 (0.000107)	0.000136 (0.000069)
Fast-decay, Variable Catchment 2SFCA	0.000116 (0.000071)	0.000107 (0.000062)	0.000116 (0.000072)	0.000076 (0.000045)	0.000108 (0.000054)	0.000119 (0.000090)	0.000132 (0.000082)	0.000133* (0.000095)	0.000126 (0.000069)
Continuous-decay, Variable Catchment 2SFCA	0.000147 (0.000086)	0.000177 (0.000082)	0.000130 (0.000080)	0.000134* (0.000068)	0.000142 (0.000061)	0.000206 (0000125)	0.000195 (0.000092)	0.000212 (0.000123)	0.000149 (0.000079)
Slow-decay, Variable Catchment SPAR	1.00 (0.58)	1.15 (0.56)	0.91 (0.55)	0.85 (0.42)	0.94 (0.41)	1.43 (0.86)	1.32 (0.62)	1.39 (0.79)	1.01 (0.52)
Fast-decay, Variable Catchment SPAR	1.00 (0.61)	0.92 (0.53)	1.00 (0.62)	0.66 (0.39)	0.93 (0.47)	1.03 (0.77)	1.14 (0.71)	1.14* (0.82)	1.08 (0.59)

2SFCA, Two-Step Floating Catchment Area; *SPAR*, Spatial Access Ratio

* Not siginificant at $P < 0.05$; all other Rural and Urban comparisons were significant at $P < 0.05$.

CHAPTER 4

Predicting Late-Stage Breast Cancer Diagnosis and Receipt of Adjuvant Therapy: Applying Current Spatial Access to Care Methods in Appalachia

4.1 Abstract

Purpose

The two-step floating catchment area (2SFCA) method of measuring access to care has never been used to study cancer disparities in Appalachia, despite its improvements over traditional measures of access to care. We examine the impact of access to mammography centers and primary care on late stage breast cancer diagnosis and receipt of adjuvant hormonal therapy.

Methods

Cancer registries from Pennsylvania, Ohio, Kentucky, and North Carolina were linked with Medicare data to identify the stage of breast cancer diagnosis for Appalachia women diagnosed between 2006-2008. Women eligible for adjuvant therapy included those with stage I, II, or III diagnosis; mastectomy or breast conserving surgery; and hormone-receptor positive breast cancers. Geographically weighted regression (GWR) was used to explore non-stationarity in the demographic and spatial access predictor variables.

Results

Over 21% of 15,299 women diagnosed with breast cancer received a late-stage (stages III-IV) diagnosis. Predictors included age at diagnosis (odds ratio [OR], 0.86; $P < 0.001$), insurance status (OR, 1.32; $P < 0.001$), county primary care to population ratio (OR, 0.95; $P < 0.001$), and

primary care 2SFCA score (OR, 0.96; *P* = 0.006). Only 46.9% of eligible women received adjuvant hormonal therapy, and significant predictors included comorbidity status (OR, 1.18; *P* = 0.047), county economic status (OR, 1.32; *P* = 0.006), and mammography center 2SFCA scores (OR, 1.12; *P* = 0.021).

Conclusion

The 2SFCA method offered the greatest predictive validity of the access measures. Rates of late stage breast cancer diagnosis and adjuvant hormonal therapy are substandard in Appalachia.

4.2 Introduction

The Appalachia region of the United States has reduced health outcomes and treatment patterns across a number of diseases, including breast cancer (Blackley et al., 2012; Reiter et al., 2013; Sergeev, 2013). Because many areas of Appalachia have lower socioeconomic status (Pollard & Jacobson, 2012) and occupy rural, mountainous terrain, reduced access to care is often implicated in the region's cancer disparities (Paskett et al., 2011).

Spatial access to care is traditionally measured using either provider to population ratios or by computing the travel time between patient and provider (Wang, 2012). Both methods have limitations, however. Provider to population ratios use fixed geographic boundaries (e.g., counties) that do not reflect actual patient behaviors, while travel time fails to account for supply and demand factors (Luo & Wang 2003). More recently, the two-step floating catchment area (2SFCA) method was developed to overcome these limitations (Wang & Luo, 2005). Despite its improvement over traditional measures of spatial access to care, the 2SFCA method has never been used to study cancer outcomes or treatment patterns in Appalachia.

We recently evaluated the impact of different 2SFCA parameter options when measuring access to mammography centers (chapter 2) and primary care physicians (chapter 3) in Appalachia. Here, we used a linked central cancer registry and Medicare dataset across four Appalachian states to evaluate the relationship between spatial access to care and two important clinical indicators for breast cancer—late stage diagnosis and receipt of adjuvant hormonal therapy. Late stage breast cancer diagnosis leads to fewer treatment options and increased mortality (Henry et al., 2011) and is more prevalent in lower socioeconomic, rural, and black populations (Amey et al., 1997; Clegg et al., 2009; Lantz et al., 2006). Adjuvant hormonal therapy is recommended for hormone receptor positive patients after either breast conserving surgery or mastectomy (Howell et al., 2005; Winer et al., 2005). Lower socioeconomic status is also associated with reduced rates of adjuvant hormonal therapy (Wu et al., 2012).

First, we evaluated the predictive ability of the 2SFCA method compared to traditional spatial access approaches. We then used geographically weighted regression (GWR) to examine whether the influence of demographic or spatial access predictor variables differed throughout the study region.

4.3 Methods

This research was approved by the institutional review board at the University of Michigan.

STUDY POPULATION

We used PA, OH, KY, and NC Central Cancer Registry (CCR) datasets to identify 15,299 women diagnosed with breast cancer between 2006 and 2008 who lived in Appalachia counties, defined by the Appalachia Regional Commission (ARC). To examine receipt of hormonal therapy, CCR datasets were linked with Medicare claims to further limit the sample to patients with Medicare Part D enrollment; diagnosis during year 2007; stage I, II, or III diagnosis (Johnson & Ademo, 2007); confirmed mastectomy or breast conserving surgery; and hormone-receptor positive breast cancers; resulting in 834 women eligible for adjuvant hormonal therapy (EBCTCG, 2005). The methods used to link CCR and Medicare datasets have been previously described (Anderson et al., 2005).

DEFINITION OF VARIABLES

Dependent Variables. Breast cancer stage at diagnosis and receipt of adjuvant hormonal therapy were the two dependent variables. Similar to previous research (Markossian & Hines, 2012), early stage diagnosis was defined as stages 0, I, or II, while late stage was defined as stages III and IV. Patients with prescription codes for tamoxifen, anastrozole, letrozole, or exemestane were defined as having received adjuvant hormonal therapy (Howell et al., 2005).

Spatial Access Measures. Three spatial access methods were evaluated: 1) county provider to population ratios, 2) travel time to closest provider, and 3) the 2SFCA method.

Provider to population ratios were calculated for primary care physicians and mammography centers. The American Medical Association (AMA) Physician Masterfile from 2008 was used to identify primary care physicians, which we defined as having specialties of Family Practice, General Practice, Internal Medicine, or General Pediatrics (Camacho et al., 2014). The number of primary care physicians in each county was divided by that county's population, using counties and populations from the 2010 Census, to yield a ratio for each county. All 2008 U.S. Food and Drug Administration (FDA) accredited mammography centers in Appalachia counties of the study region were identified. The number of mammography centers in each county was divided by that counties population of women age 45 and older.

Although 2008 guidelines (cite) called for mammography screening beginning at age 40, the 2010 Census used age groups 25-34 and 35-44. We chose to use women age 45 and older. Each patient was assigned a primary care ratio and a mammography center ratio based on their county of residence.

For travel time calculations, the physician and mammography centers were geocoded using ArcGIS 10.1. Exact office addresses were available for 8,039 of the 9,483 physicians. For the remaining 1,444 physicians, the population weighted centroid of the census tract associated with the physician's office was geocoded. All 1,181 mammography centers had geocodable addresses. We used ArcGIS Network Analyst extension to calculate the driving time between the closest physician and mammography center and the population weighted centroid of every census block group in the study region. For reference, census block groups are generally composed of between 600 and 3,000 people. Each patient was assigned a travel time to their closest primary care physician and mammography center based on the census block group of their residence.

More detailed methodological explanations of the 2SFCA method are described elsewhere (McGrail, 2012). To complete the first step of the 2SFCA, we created service provider to population ratios using each provider and the sum of all the census block group populations (or populations of females 45 and older for mammography centers) within that provider's catchment area. Catchment areas were set to 60 minutes to reflect the rural nature of Appalachia (McGrail & Humphreys, 2009). Step two then moved to each block group population center and identified all service providers within the designated 60 minute catchment around that block group. The step one ratios within each population's catchment were summed, resulting in each block group having a primary care access score and a mammography center access score.

Two critiques of the original 2SFCA method are its failure to vary catchment sizes based on population needs and its failure to account for distance decay within catchments (Luo & Whippo, 2012). We added a distance decay function by breaking catchments into four time zones: 0-10 minutes, 11-20 minutes, 21-30 minutes, and 31-60 minutes (Luo & Whippo, 2012). Two sets of weights were chosen for the distance decay function, corresponding to fast decay (weights 1, 0.60, 0.25, 0.05) and slow decay (weights 1, 0.80, 0.55, 0.15; McGrail, 2012). Distance decay was applied at both steps.

We also varied the catchment size at each step of the 2SFCA, using McGrail's (2012) approach. At step one, the technique uses a set of rules to determine whether the distance decay weights are applied (thereby expanding or reducing a service provider's coverage area). No step one decay weight was applied if the travel time between a provider and population was less than 10 minutes, if the provider was one of the 25 closest (5 closest for mammography centers) to that population center, or if the population center had a population less than 5,000 and less than half the population of the service provider's town (McGrail, 2012). For step two, catchments were varied by only including the closest 100 primary care physicians and only the closest 20 mammography centers (McGrail, 2012). Thus, each census block group in the study region was given an original 2SFCA score without any distance decay or varied catchments, a slow decay 2SFCA score with the variable catchment rules, and a fast decay 2SFCA score with the variable catchment rules. Patients were matched to 2SFCA scores by the census block group of their address.

Demographic and Additional Independent Variables. We included age, insurance status, race/ethnicity, and county economic status along with the spatial access measures for the late stage breast cancer diagnosis model. Age was categorized as younger than 50, 50-64, or 65 or older. Insurance status was split into five categories: private, Medicare, Medicaid, uninsured, and unknown. Race/ethnicity was defined as white, black, or Hispanic/other. We used the Appalachia Regional Commission's (ARC) designation for labeling counties as either attainment, competitive, transitional, at-risk, or distressed (ARC, 2012). The ARC uses the three-year average unemployment rate, the per capita market income, and the poverty rate to calculate an average for every county in the nation, and then ranks each county into quintiles to yield the final economic status descriptor.

When creating the model for receipt of adjuvant hormonal therapy, we included age, race/ethnicity, county economic status, cancer stage, and comorbidity index along with the spatial access measures. Age was defined as 65-69, 70-74, 75-79, and older than 80.. , Race/ethnicity and county economic status were defined as described in the above model. Cancer stage was categorized as stage I, II, or III, reflecting those patients eligible for adjuvant hormonal therapy. The Charlson Comorbidity Index, which creates a weighed score of comorbidity, was constructed from Medicare claims data (D'Hoore et al., 1996).

DATA ANALYSIS

Initial exploratory analysis and goodness-of-fit statistics were used to determine which spatial access and demographic variables to include in the logistic regression models predicting late stage breast cancer diagnosis and receipt of adjuvant hormonal therapy (Fotheringham et al., 2002; Windle et al., 2010). Goodness-of-fit statistics included the corrected Akaike's Information Criterion (AIC_c) and the area under the receiver operating characteristic curve (AUC). AIC_c values lower than 3 indicate better model fit (Windle et al., 2010). An AUC value of 1 indicates that a model perfectly predicts the dependent variable, while an AUC value of 0.5 indicates random chance (Zou et al., 2007). Adjusted R squared, a measure of variation explained by the chosen predictor variables, was also used to compare models (Fotheringham et al., 2002).

Once an appropriate model was specified for the traditional logistic regression and the logistic GWR, we looked for evidence of spatial non-stationarity of the predictor variables. Briefly, GWR includes geographic coordinates at each study observation to construct a series of local regressions, resulting in unique predictor coefficients at each study observation. Traditional global logistic regression creates one coefficient for each predictor variable that is assumed to be stationary across the entire study region. We used a fixed Gaussian kernel type and the golden selection search to minimize the AIC_c when selecting bandwidth size for the logistic GWR. GWR 4.0 software was used for the logistic GWR (available at https://geodacenter.asu.edu/gwr_software). Fotheringham et al. (2002) provide a more detailed methodological overview of the GWR approach.

We also examined descriptive statistics of the local coefficients in the logistic GWR to identify large minimum to maximum ranges, existence of both positive and negative coefficient values, skewness, or any other indicators of spatial non-stationarity. When spatial non-stationarity was found, we mapped the coefficients onto the geographic study area to examine varying geographic effects of each predictor variable.

4.4 Results

There were 15,299 women living in Appalachia counties of PA, OH, KY, and NC diagnosed with breast cancer between years 2006 and 2008. Table 4.1 provides descriptive and chi-square statistics of the study variables characterized by early and late stage breast cancer diagnosis. The majority of women had early stage diagnosis (78.8%), were older than 65 (mean age = 71.1, SD = 10.4), and were insured by Medicare (65.5%). Although race/ ethnicity was not included in Table 1, most women were white (95.3%). Race/ethnicity was not associated with stage at diagnosis ($X^2 = 4.26$, $p = 0.119$). The fast decay 2SFCA scores for primary care and mammography centers showed a stronger relationship to stage of diagnosis than slow decay and original 2SFCA scores in both chi-square analysis and subsequent regression modeling, and for brevity are the only 2SFCA scores shown in Table 4.1. 2SFCA scores are broken into quintiles for ease of interpretation, with larger scores (and larger quintiles) indicating greater spatial access to care.

Table 4.2 provides descriptive and chi-square statistics characterized by receipt of adjuvant hormonal therapy for the 834 eligible women living in Appalachia during 2007. The majority of women (53.1%) did not receive adjuvant hormonal therapy. Most women were white (97.1%), and race/ethnicity was again omitted from the table for brevity. Unlike the stage at diagnosis analysis, the slow decay 2SFCA scores showed a stronger relationship to receipt of adjuvant hormonal therapy compared to fast decay and original 2SFCA scores, and are the only 2SFCA scores shown in Table 4.2.

Goodness-of-fit statistics revealed that the model with variables age, insurance status, fast decay 2SFCA primary care score, and county primary care to population ratio was the best predictor of late stage breast cancer diagnosis. Table 4.3 shows the resulting parameter estimates for the global logistic regression using these variables to predict stage at diagnosis. Travel time to the closest primary care provider was not a significant predictor in the global logistic regression (OR, 1.051; $p = 0.241$) and did not improve overall model fit. The mammography center spatial access scores did not fit the model as well as the primary care spatial access scores, as measured by AIC_c, AUC, and adjusted r^2.

Descriptive statistics for the identical logistic GWR model predicting stage at diagnosis are presented in Table 4.4. The range of coefficient estimates, including both positive and negative values for each predictor variable, suggests spatial non-stationarity across the study region. Goodness-of-fit statistics showed that the logistic GWR model performed better than the global logistic regression model when predicting late stage breast cancer diagnosis (Table 4.5). The logistic GWR model had a lower AIC_c, higher AUC, and higher adjusted r^2, demonstrating better model performance because of spatial non-stationarity in the final predictor variables. This non-stationarity is displayed in Figure 4.1, where the effects of the final predictor variables clearly vary throughout the four state study region.

The model including Charlson comorbidity index, county economic status, and slow decay 2SFCA mammography score was the best predictor of adjuvant hormonal therapy receipt. The parameter estimates for the resulting global logistic regression model are shown in Table 4.3. Travel time to the nearest mammography center (OR, 0.954; $p = 0.621$) and the county mammography center to women age 45 and older ratio (OR, 0.985; $p = 0.792$) did not improve overall model fit. Unlike the stage at diagnosis model, the mammography access scores resulted in a better fitting model than using primary care access scores. The receipt of adjuvant therapy model also used the slow decay 2SFCA score rather than the fast decay 2SFCA score, based on improved AIC_c, AUC, and adjusted r^2 values.

Descriptive statistics for the identical logistic GWR model predicting receipt of adjuvant hormonal therapy are presented in Table 4.4. The smaller coefficient ranges and lack of positive and negative values indicate that the effect of the predictor variables were more stationary across the study region. Goodness-of-fit statistics demonstrated that the logistic GWR model was no better than the global logistic regression model in predicting receipt of adjuvant hormonal therapy (Table 4.5). There were no meaningful differences between the AIC_c, AUC, and adjusted r^2 values of the two approaches. As a result, we did not map the geographic variability of the logistic GWR parameter coefficients because we found no evidence of spatial non-stationarity in the predictor variables.

4.5 Discussion

Our analysis of breast cancer patients in Appalachia PA, OH, KY, and NC found that over 20% of women were diagnosed with late stage breast cancer. Age and insurance status predicted late stage diagnosis, as well as the spatial access to care measures of fast-decay 2SFCA primary care score and primary care to county population ratio. In our sample, only 46.9% of eligible women received adjuvant hormonal therapy following breast cancer surgery. Patients' comorbidity and county economic status predicted receipt of adjuvant therapy. The slow-decay 2SFCA mammography center score also predicted receipt of adjuvant therapy. Methodologically, our analysis showed the superior predictive validity of the 2SFCA method compared with the traditional spatial access measure of driving time from patient to provider. We also demonstrated the necessity of considering spatial non-stationarity across study areas. A GWR approach was more appropriate for our late stage diagnosis model than a traditional global regression.

Previous research using similar early (0-II) and late (III-IV) stage classification found late stage breast cancer diagnosis in anywhere from 13 percent of cases in Kentucky (Huang et al., 2009), 16.7 percent of cases in Georgia (Markossian & Hines, 2012), and 17.3 percent of cases in Appalachia counties of Pennsylvania, Ohio, and Kentucky (Anderson et al., 2014). The national average from Surveillance, Epidemiology, and End Result data is 16 percent (SEER, 2012). Our 21.2 percent finding continues the trend of increased late stage diagnosis in Appalachia.

Prior research was inconclusive as to whether spatial access to care increased risk of late stage diagnosis, with some research (Huang et al., 2009) finding that greater travel time to mammography centers did increase risk, while other research (Henry et al., 2011; Lian et al., 2012) found no increase due to longer travel times. Our work adds clarity by using a more advanced spatial access method than travel time, the 2SFCA method, and by including access to primary care instead of only mammography centers. For our study population, a fast-decay 2SFCA primary care score was a better predictor of late stage diagnosis than 2SFCA mammography scores, travel time to mammography centers, or travel time to primary care providers. This result supports the finding that primary care providers are often a crucial gateway to breast cancer care (Roetzheim et al., 2012). We also found that lower primary care to county population ratios predicted late stage diagnosis. This ratio may reflect the increased risk of late stage diagnosis among lower socioeconomic groups (Ruffin et al., 2000), as it was strongly correlated with county economic status ($p < 0.001$).

Only 46.9 percent of eligible women in our sample received adjuvant hormonal therapy of tamoxifen, anastrozole, letrozole, or exemestane. This is considerably lower than the 67 percent reported for an analysis of patients in California, Florida, Illinois and New York (Yen et al., 2007), and the 64 percent reported in an analysis of low-income patients throughout North Carolina (Kimmick et al., 2009). Again, the 2SFCA method was a better predictor of adjuvant hormonal therapy than travel time. In this model, slow decay 2SFCA mammography scores were a better predictor of treatment than primary care 2SFCA scores. We hypothesize that it is access to these more specialized cancer care services, rather than primary care providers, that impact later stages of care, such as our measure of adjuvant hormonal therapy. Previous research has found that access to specialized cancer care services improves cancer outcomes and reduces treatment complications (Hill et al., 2000).

There are both methodological and clinical strengths to this study. To the best of our knowledge, this is the first study that uses the more advanced 2SFCA method to measure spatial access to care in Appalachia. Because Appalachia is largely rural and mountainous, and many areas experience physician shortages (Halverson et al., 2004), accurately measuring spatial access to care is essential. Another strength was our GWR approach. Logistic GWR allowed us to visualize the varying effects of our included predictor variables, such as the finding that having Medicaid, being uninsured, or having an unknown insurance status are more important predictors of late stage diagnosis in northeastern PA than 2SFCA scores. Clinically, our study benefits from the inclusion of both access to primary care services and mammography centers. Access to care is multifaceted, with variables contributing differentially depending on the outcome or treatment pattern studied (Wang & Luo, 2005).

There are also important limitations to the study. We were not able to obtain provider data for states bordering our study area. We did include bordering populations when calculating the 2SFCA scores, but additional provider data would have further reduced any edge effects in our spatial access calculations. Additionally, we used the closest provider and mammography centers to patients, not the actual service providers. For patients with more restrictive or no insurance, this could represent a significant discrepancy. We also did not include non-automobile transportation, which is possible to account for in the 2SFCA method (Mao & Nekorchuk, 2013.

Overall, we found disparities in diagnosis and treatment patters in our sample of breast cancer patients in Appalachia. The 2SFCA method offered the best predictive ability of the spatial access measures we used and should be included whenever measuring access to care in Appalachia. Geographic weighted regression was useful in identifying spatial non-stationarity in our model predicting late stage breast cancer diagnosis. Further qualitative research is needed to understand the range of factors that reduce access to care. Additional research is also needed to identify the clinical decisions that lead to reduced rates of adjuvant hormonal therapy among women in Appalachia.

4.6 References

Amey CH, Miller MK, & Albrecht SL (1997). The role of race and residence in determining stage at diagnosis of breast cancer. *Journal of Rural Health.* 13(2), 99–108.

Anderson RT, Camacho FT, Balkrishnan R, et al. (2005). Use of cancer registry data for research on patterns of breast cancer care of individuals with Medicaid insurance. *J Clin Oncol.* 24:533s (suppl; abstr 6021).

Appalachian Regional Commission (ARC; 2012). Economic Assessment of Appalachia. County Economic Status in Appalachia, FY 2012. Retrieved from http://www.arc.gov/research/MapsofAppalachia.asp?MAP_ID=55.

Blackley D, Behringer B, & Zheng S (2012). Cancer mortality rates in Appalachia: descriptive epidemiology and an approach to explaining differences in outcomes. *J Community Health*. 37(4): 804-813.

Camacho F, Hwang W, Kern T, & Anderson RT (2014). Receipt of Regular Primary Care and Early Cancer Detection in Appalachia. *The Journal of Rural Health.* doi: 10.1111/jrh.12097.

Clegg LX, Reichman ME, Miller BA, et al. (2009). Impact of socioeconomic status on cancer incidence and stage at diagnosis: Selected findings from the surveillance, epidemiology, and end results: National Longitudinal Mortality Study. *Cancer Causes Control.* 20(4), 417–435.

D'Hoore W, Bouckaert A, & Telquin C (1996) Practical considerations on the use of the Charlson Comorbidity Index with administrative data bases. *J Clin Epidemiol.* 49:1429-1433.

Early Breast Cancer Trialists' Collaborative Group (EBCTCG; 2005). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomized trials. *Lancet*. 365:1687-1717.

Fotheringham AS, Brunsdon C, & Charlton ME (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, Wiley, Chichester, UK. 269 pp.

Halverson J, Lin M, & Harner E (2004). An analysis of disparities in health status and access to health care in the Appalachian region. Washington, DC: Appalachian Regional Commission, 2004.

Henry KA, Boscoe FP, Johnson CJ, Goldberg DW, Sherman R, & Cockburn M (2011). Breast cancer stage at diagnosis: Is travel time important? *J. Community Health*. 1(36):933–942.

Hillner BE, Smith TJ, Desch CE (2000). Hospital and physician volume or specialization and outcomes in cancer treatment: Importance in quality of cancer care. *Journal of Clinical Oncology*. 18: 2327–2340.

Howell A, Cuzick J, Baum M, et al. (2005). Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet* 365:60-62.

Johnson CH & Adamo M (eds; 2007). SEER Program Coding and Staging Manual 2007. Bethesda, MD, National Cancer Institute, NIH Publication 07-5581.

Lantz PM, Mujahid M, Schwartz K, et al. (2006). The influence of race, ethnicity, and individual socioeconomic factors on breast cancer stage at diagnosis. *American Journal of Public Health.* 96(12), 2173–2178.

Luo W & Wang F (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B*. 30(6): 865–884.

Luo W & Whippo TL (2012). Variable catchment sizes for the two-step floating catchment area (2SFCA) method. *Health Place.* 18:789–795.

Mao L & Nekorchuk D (2013). Measuring spatial accessibility to healthcare for populations with multiple transportation modes. *Health Place*. 24:115-122.

Markossian TW, Hines RB (2012). Disparities in Late Stage Diagnosis, Treatment, and Breast Cancer-Related Death by Race, Age, and Rural Residence Among Women in Georgia, *Women & Health*. 52(4):317-33.

McGrail MR (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *Int J Health Geogr*. 16:11-50.

McGrail MR & Humphreys JS (2009). A new index of access to primary care services in rural areas. *Aust NZ J Pub Health.* 33:418–423.

Onega T, Duell EJ, Shi X, Wang D, Demidenko E, et al. (2008) Geographic access to cancer care in the US. *Cancer*. 112:909–918.

Paskett ED, Fisher JL, Lengerich EJ et al. (2011). Disparities in underserved white populations: the case of cancer-related disparities in Appalachia. *Oncologist*. 16(8): 1072-81.

Pollard K & Jacobson LA (2012). The Appalachian Region: A Data Overview from the 2006-2010 American Community Survey. *Appalachia Regional Commission.* Retrieved from http://www.arc.gov/research/researchreportdetails.asp?REPORT_ID=96.

Reiter PL, Katz ML, Paskett ED (2013). Correlates of HPV vaccination among adolescent females from Appalachia and reasons why their parents do not intend to vaccinate. *Vaccine.* 31: 3121–3125.

Roetzheim RG, Ferrante JM, Lee JH, et al. (2012). Influence of primary care on breast cancer outcomes among Medicare beneficiaries. *Ann Fam Med*. 10(5):401-411.

Ruffin M, Gorenflo F, & Woodman B (2000). Predictors of screening for breast, cervical, colorectal, and prostatic cancer among community-based primary care practices. *J Am Board FamMed*. 13(1):1-10.

Sergeev AV (2013). Stroke mortality disparities in the population of the Appalachian Mountain region. *Ethn Dis.* 23(3):286-91.

Wang F (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Ann Assoc Am Geogr*. 102(5): 1104–1112.

Wang F & Luo W (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health Place.* 11:131–146.

Windle MJS, Rose GA, Devillers R, & Fortin, M-J (2010). Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. *ICES Journal of Marine Science*, 67:145–154.

Winer EP, Hudis C, Burstein HJ, et al. (2005). American Society of Clinical Oncology technology assessment on the use of aromatase inhibitors as adjuvant therapy for postmenopausal women with hormone receptor-positive breast cancer: Status report 2004. *J Clin Oncol*. 23:619-629.

Wu XC, Lund MJ, Kimmick GG, Richardson LC, Sabatino SA, et al. (2012). Influence of race, insurance, socioeconomic status, and hospital type on receipt of guideline-concordant adjuvant systemic therapy for locoregional breast cancers. *J Clin Oncol* 30: 142–150.

Zou KH, O'Malley AJ, & Mauri L (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation.* 115: 654–657.

Table 4.1. Characteristics of study patients by breast cancer stage at diagnosis.

Factors	Total cases	%	% Early stage	% Late Stage	*p* -value
	15,299		78.8	21.2	
Age					
<50	501	3.3	64.5	35.5	<0.001
50 - 64	3,242	21.2	80.9	19.1	
65+	11,556	75.5	78.8	21.2	
Insurance Status					
Private	4,309	28.2	81.8	18.2	<0.001
Medicare	10,027	65.5	78.6	21.4	
Medicaid	544	3.6	71.1	28.9	
Uninsured	112	0.7	68.8	31.3	
Unknown	307	2	59.3	40.7	
County Economic Status					
Attainment	0	0	0	0	<0.001
Competitive	3,351	21.9	80.6	19.4	
Transitional	9,519	62.2	78.7	21.3	
At-risk	1,482	9.7	78.3	21.7	
Distressed	947	6.2	73.5	26.5	
Driving Time to Closest Primary Care (minutes)					
<10	13,010	85.0	79.2	20.8	0.031
10 - <20	1,863	12.2	76.8	23.2	
20 - <30	383	2.5	74.2	25.8	
30 - <40	41	0.3	85.4	14.6	
40 - <50	1	<0.001	100	0	
50+	1	<0.001	100	0	
Driving Time to Closest Mammography Center (minutes)					
<10	9,332	61.0	79.3	20.7	0.337
10 - <20	4,008	26.2	78.3	21.7	
20 - <30	1,417	9.3	76.7	23.2	
30 - <40	422	2.8	79.4	20.6	
40 - <50	97	0.6	78.4	21.6	
50+	23	0.2	82.6	17.4	
Ratio of Primary Care Physicians / County Population					
1st Quintile	3,104	20.3	75.9	24.1	<0.001
2nd Quintile	3,055	20.0	77.5	22.5	
3rd Quintile	3,065	20.0	80.0	20.0	
4th Quintile	3,364	22.0	79.3	20.7	
5th Quintile	2,711	17.7	81.4	18.6	
Ratio of Mammography Centers / County Population of Females Age 45 and older					
1st Quintile	3,185	20.8	79.3	20.7	<0.001
2nd Quintile	3,050	19.9	81.9	18.1	
3rd Quintile	3,352	21.9	78.6	21.4	
4th Quintile	2,735	17.9	77.4	22.6	
5th Quintile	2,977	19.5	76.4	23.6	
2SFCA Fast Decay Scores to Primary Care					
1st Quintile	3,059	20.0	76.4	23.6	<0.001
2nd Quintile	3,060	20.0	79.0	21.0	
3rd Quintile	3,063	20.0	80.5	19.5	
4th Quintile	3,057	20.0	78.1	21.9	
5th Quintile	3,060	20.0	79.9	20.1	
2SFCA Fast Decay Scores to Mammography Centers					
1st Quintile	3,060	20.0	79.0	21.0	0.130
2nd Quintile	3,061	20.0	79.6	20.4	
3rd Quintile	3,060	20.0	79.7	20.3	
4th Quintile	3,061	20.0	77.4	22.6	
5th Quintile	3,057	20.0	78.2	21.8	

2SFCA, Two-Step Floating Catchment Area

Table 4.2. Characteristics of study patients by receipt of adjuvant hormonal therapy (AT).

Factors	Total cases	%	% Had AT	% No AT	*p*-value
	834		46.9	53.1	
Age					
65 - 69	172	20.6	48.8	51.2	0.159
70 - 74	195	23.4	40.5	59.5	
75 - 79	208	24.9	51.4	48.6	
80+	259	31.1	46.7	53.3	
Charleson Comorbidity Index					
0	535	64.1	43.9	56.1	0.110
1	199	23.9	52.8	47.2	
2	57	6.8	47.4	52.6	
3+	43	5.2	55.8	44.2	
County Economic Status					
Attainment	0.0	0.0	0.0	0.0	0.025
Competitive	110	13.2	40.0	60.0	
Transitional	571	68.5	45.9	54.1	
At-risk	100	12.0	51.0	49.0	
Distressed	53	6.4	64.2	35.8	
Driving Time to Closest Primary Care (minutes)					
<10	706	84.7	47.3	52.7	0.386
10 - <20	99	11.9	44.4	55.6	
20 - <30	26	3.1	50.0	50.0	
30 - <40	3	0.4	0.0	100.0	
40 - <50	0	0.0	0.0	0.0	
50+	0	0.0	0.0	0.0	
Driving Time to Closest Mammography Center (minutes)					
<10	522	62.6	46.9	53.1	0.576
10 - <20	201	24.1	48.8	51.2	
20 - <30	81	9.7	40.7	59.3	
30 - <40	22	2.6	54.5	45.5	
40 - <50	6	0.7	50.0	50.0	
50+	2	0.2	0.0	100.0	
Ratio of Primary Care Physicians / County Population					
1st Quintile	174	20.9	49.4	50.6	0.609
2nd Quintile	157	18.8	48.4	51.6	
3rd Quintile	174	20.9	46.0	54.0	
4th Quintile	168	20.1	48.8	51.2	
5th Quintile	161	19.3	41.6	58.4	
Ratio of Mammography Centers / County Population of Females Age 45+					
1st Quintile	170	20.4	40.0	60.0	0.339
2nd Quintile	176	21.1	50.6	49.4	
3rd Quintile	183	21.9	47.0	53.0	
4th Quintile	141	16.9	48.2	51.8	
5th Quintile	164	19.7	48.8	51.2	
2SFCA Slow Decay Scores to Primary Care					
1st Quintile	165	19.8	40.6	59.4	0.259
2nd Quintile	172	20.6	45.9	54.1	
3rd Quintile	162	19.4	53.1	46.9	
4th Quintile	166	19.9	48.2	51.8	
5th Quintile	169	20.3	46.7	53.3	
2SFCA Slow Decay Scores to Mammography Centers					
1st Quintile	168	20.1	41.1	58.9	0.037
2nd Quintile	172	20.6	42.4	57.6	
3rd Quintile	162	19.4	51.9	48.1	
4th Quintile	167	20.0	44.3	55.7	
5th Quintile	165	19.8	55.2	44.8	

2SFCA, Two-Step Floating Catchment Area

Table 4.3. Parameter estimates for the global logistic regression model predicting late stage breast cancer diagnosis* and receipt of adjuvant hormonal therapy⁺.

Variable*	β	s.e.	z-statistic	Sig.	Exp(β)	95% C.I. for Exp(β)
Intercept	-1.151	0.123	88.267	<0.001	0.316	
Age	-0.146	0.038	15.103	<0.001	0.864	0.803 - 0.930
Insurance Status	0.276	0.026	110.749	<0.001	1.318	1.252 - 1.388
2SFCA score- Primary Care, Fast Decay	-0.036	0.014	7.509	0.006	0.962	0.935 - 0.989
County Primary Care / Population Ratio	-0.056	0.014	14.837	<0.001	0.946	0.919 - 0.973

Variable⁺	β	s.e.	z-statistic	Sig.	Exp(β)	95% C.I. for Exp(β)
Intercept	-1.431	0.353	16.464	<0.001	0.239	
Charlson Comorbidity Index	0.168	0.084	3.948	0.047	1.183	1.002 - 1.395
County Economic Status	0.280	0.101	7.620	0.006	1.323	1.085 - 1.615
2SFCA score - Mammography Center, Slow Decay	0.115	0.050	5.340	0.021	1.122	1.018 - 1.236

2SFCA, Two-Step Floating Catchment Area

Table 4.4. Summary statistics of the parameter estimates for the logistic GWR models for late stage breast cancer diagnosis*[a] and receipt of adjuvant hormonal therapy+[b].

Variable*	Minimum	1st Quartile	Median	3rd Quartile	Maximum
Intercept	-4.532	-1.860	-1.276	-0.709	1.725
Age	-1.090	-0.229	-0.126	-0.078	0.389
Insurance Status	-0.104	0.231	0.299	0.370	0.883
2SFCA score - Primary Care, Fast Decay	-0.522	-0.083	-0.020	0.023	0.408
County Primary Care / Population Ratio	-1.806	-0.078	-0.022	0.057	0.465

Variable+	Minimum	1st Quartile	Median	3rd Quartile	Maximum
Intercept	-1.555	-1.437	-1.318	-1.287	-1.261
Charlson Comorbidity Index	0.131	0.155	0.181	0.186	0.193
County Economic Status	0.217	0.237	0.252	0.291	0.324
2SFCA score - Mammography Center, Slow Decay	0.100	0.102	0.105	0.108	0.121

2SFCA, Two-Step Floating Catchment Area

[a] 36.9 km bandwidth

[b] 541.5 km bandwidth

Table 4.5. Comparison of fit between global logistic regression and GWR models for late stage breast cancer diagnosis*[a] and receipt of adjuvant hormonal therapy+[b].

Model*	*n*	k_e	-2 log likelihood	AIC_c	AUC ± s.e.	r^2 (adj.)
Global Logistic	15,299	5.0	15,670.6	15,680.6	0.562 ± 0.006	0.009
GWR	15,299	144.1	15,299.6	15,590.5	0.623 ± 0.006	0.033
Model+	*n*	k_e	-2 log likelihood	AIC_c	AUC ± s.e.	r^2 (adj.)
Global Logistic	834	4.0	1,135.0	1,143.0	0.583 ± 0.20	0.016
GWR	834	5.2	1,133.4	1,143.9	0.587 ± 0.20	0.017

n, number of patients; ke, effective number of parameters; AICc, corrected Akaike's Information Criterion; AUC, area under the receiver operating characteristic curve.

[a] 36.9 km bandwidth

[b] 541.5 km bandwidth

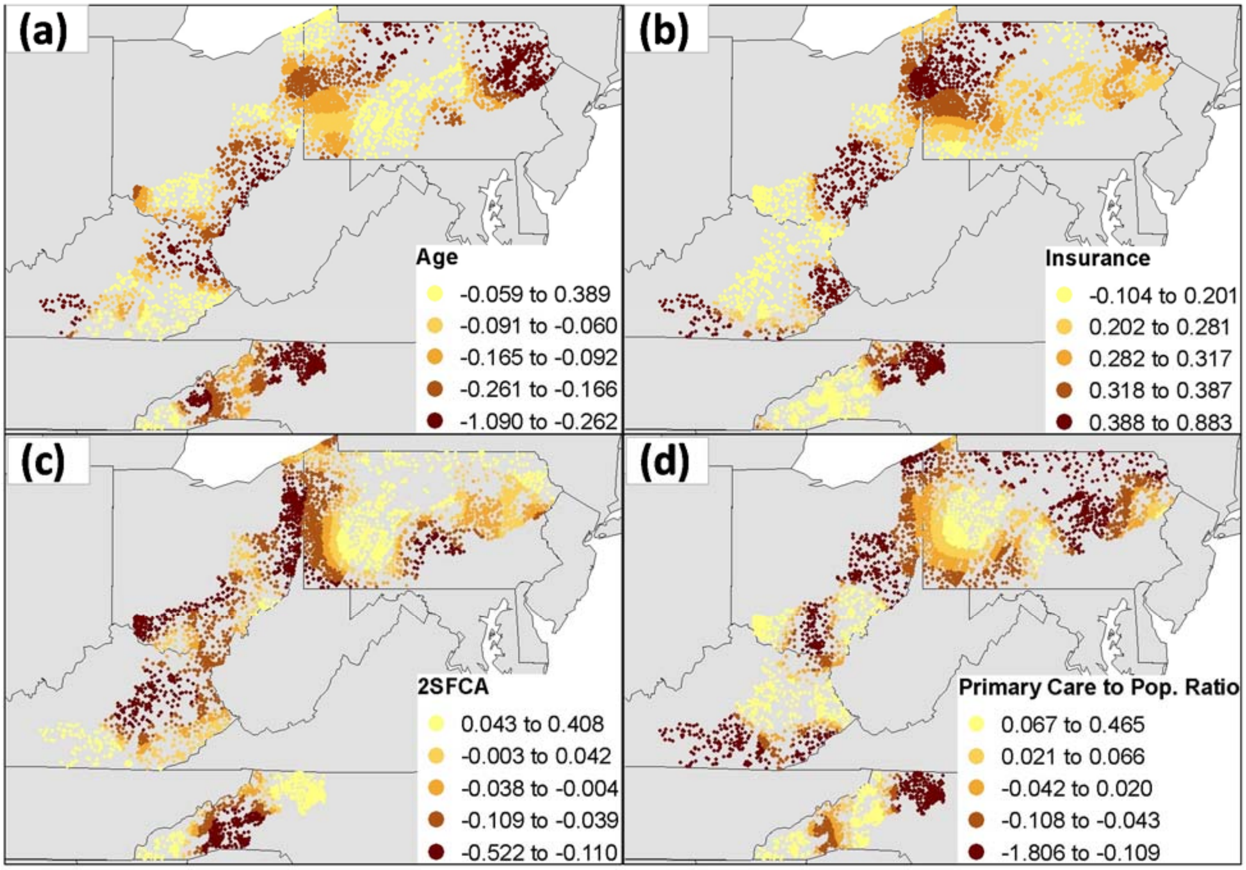

Figure 4.1. Local coefficient estimates for the logistic GWR model predicting late stage breast cancer diagnosis across variables (a) age, (b) insurance status, (c) 2SFCA score, and (d) county primary care to population ratio.

CHAPTER 5

Overall Conclusion

5.1 Research Question

The Appalachia region of the U.S. suffers from numerous socioeconomic disparities (Pollard & Jacobsen 2013), increased incidence and mortality for many cancers (Blackley, Behringer, & Zheng, 2012; Winger et al., 2003), and substandard cancer screening and treatment patterns (Freeman, Huang, & Dragun, 2012; Hall et al., 2002). Cumulatively, these disparities suggest the region has reduced access to health care (Paskett et al., 2011). Geographic information systems (GIS) technologies and the methodology used to measure spatial access to care, meanwhile, have grown from relatively simple techniques to more sophisticated approaches that provide increased geographic specificity and more realistic assumptions about patient healthcare utilization (McGrail, 2012; Wang, 2012). Nonetheless, despite Appalachia's need, these technologies have not been evaluated in the region.

5.2 Research Goals

This dissertation focused on the methodology used to measure spatial access to healthcare. The study region included four states in Appalachia: Pennsylvania, Ohio, Kentucky, and North Carolina. Traditional methods—including provider to population ratios and travel time between populations and providers—were evaluated in conjunction with the more advanced techniques of the two-step floating catchment area (2SFCA) method and the spatial access ratio

(SPAR) method. Spatial access was computed in reference to primary care physicians as well as to mammography centers. The goal was to compare and contrast what each method said about the resulting spatial access to care in the study region. Also, within the 2SFCA and SPAR methods, I evaluated the different parameters of each method among which researchers need to choose. Finally, the dissertation used the resulting spatial access scores to predict cases of late stage breast cancer diagnosis and

5.3 Study Results

In chapter two, spatial access to mammography centers was mapped across the entire states of PA, OH, KY, and NC. The larger urban areas were all within 10 minutes of their closest mammography centers, while the most rural areas of north central PA, southeastern OH, and eastern KY consistently traveled over 40 minutes to their closest mammography centers. Provider to population ratios—found by dividing the number of mammography centers in a county by the total population of county women age 45 and older—varied widely throughout the study area. The county boundaries proved problematic, as many adjacent counties had drastically different scores. Each of the 2SFCA measures were strongly correlated with each other. The original 2SFCA method showed the highest spatial access scores in many of the major urban areas of the study region. However, adding distance decay functions and varying the catchment sizes of the 2SFCA decreased access scores around many of those same urban areas and increased scores in many smaller, rural areas. Eastern Kentucky, for example, had some of the highest modified 2SFCA scores, likely due to their small population numbers and even distribution of mammography centers. Spatial access was also compared within rural and urban census block groups. Rural areas in each state had longer travel times to their closest mammography center, but also had larger provider to population ratios because of their reduced population. The 2SFCA scores were moderately higher for urban census block groups in PA, OH, and NC, while the rural block groups in KY generally had higher 2SFCA scores. When comparing spatial access between Appalachia and non-Appalachia areas of the same states, a similar pattern emerged to the urban-rural comparison. Appalachia regions had longer travel times to their closest mammography center but also had larger provider to population ratios. The 2SFCA access scores did not show a clear divide between Appalachia and non-Appalachia areas, though, with Appalachia areas having higher scores in PA and KY and non-Appalachia areas having higher scores in NC and OH. Overall, the SPAR technique was recommended because it minimized the difference in scores when deciding between which distance decay weight or variable catchment method to choose.

In Chapter three, spatial access to primary care physicians was mapped across the Appalachia regions of PA, OH, KY, and NC. Unlike travel times to closest mammography centers, most census block groups throughout the study region were less than 10 minutes from their closest primary care physicians. There was a clear pattern of relationships between each of the spatial access measures when measuring access to primary care physicians. Closest travel time and provider to population ratios were both more strongly correlated with the 2SFCA measures that did not include variable catchment sizes: the original 2SFCA, the slow decay weight 2SFCA, and the fast decay weight 2SFCA. These 2SFCA measures were also more strongly correlated with each other. Comparatively, the 2SFCA measures that included variable catchment sizes—slow decay, variable catchment size 2SFCA; fast decay, variable catchment size 2SFCA; and continuous decay, variable catchment size 2SFCA—were less strongly correlated with closest travel times, provider to population ratios, and the other 2SFCA measures, but were more strongly correlated with each other. When mapping 2SFCA scores across the study region, the same distinction occurred based on whether variable catchment sizes were included. Without variable catchments, and thus without any limit on the theoretical number of physicians a population could access, several of the largest urban areas in the study region had the highest 2SFCA scores, including Pittsburgh, Youngstown, and Greensboro. After adding rules to vary catchment sizes and cap the number of physicians a population could access, these same areas saw drops in their 2SFCA scores, while more rural areas of eastern KY and central PA had increased scores. Similar to access to mammography centers, rural census block groups had a longer average travel time to their closest primary care physician than urban census block groups. Of the four study states, KY had the longest travel times. Even Kentucky's urban block groups had longer travel times (9.54 min.) than the rural block groups of PA (7.50 min.) and OH (8.26 min.). There were also similarities to the rural-urban 2SFCA score divide in the mammography research when comparing the 2SFCA access scores to primary care physicians between rural and urban block groups. The 2SFCA scores were again moderately higher for urban areas in PA, OH, and NC, while the rural block groups in KY generally had higher 2SFCA scores. It is important to note that the urban classification also included suburban areas, for ease of interpretation. Chapter three also recommended using the SPAR score because it had a generally normalizing effect of 2SFCA scores when deciding between which decay weight and which variable catchment parameters to choose.

In Chapter four, spatial access scores to mammography centers and primary care physicians were used to predict two important clinical outcomes in breast cancer treatment: stage at diagnosis and receipt of guideline concordant adjuvant hormonal therapy. Similar to Chapter three, this research focused only on Appalachia areas of PA, OH, KY, and NC. For the 20% of women who were diagnosed with late stage breast cancer (stages III-IV), the demographic variables that predicted late stage diagnosis were a younger age and having an insurance status of Medicaid, uninsured, or unknown insurance. The spatial access to care measures of fast decay 2SFCA primary care score and primary care to county population ratio also predicted late stage diagnosis. Those access to care measures were better predictors of late stage diagnosis than travel time to a patient's closest primary care provider or any of the spatial access to mammography center measures. Geographically weighted regression demonstrated that the effect of these predictor variables on late stage diagnosis varied throughout the study region, highlighting the importance of considering spatial non-stationarity when employing regression models across large geographic areas. The demographic variables of comorbidity status and county economic status predicted whether patients received the recommended adjuvant hormonal therapy after breast cancer surgery. The spatial access to care score from the slow decay 2SFCA method in reference to mammography centers also predicted receipt of adjuvant therapy. The mammography 2SFCA measure was a better predictor than the other mammography center spatial access scores of closest travel time or provider to population ratio, and it was also a better predictor than any of the primary care spatial access measures. According to geographically weighted regression analysis, the effect of the predictor variables on the dependent variable of receipt of adjuvant therapy did not vary throughout the study region, unlike in the late stage diagnosis model. Overall, Chapter four demonstrated the effectiveness of the 2SFCA method in predicting important clinical indicators. The chapter also demonstrated the necessity of considering spatial non-stationarity when predicting clinical outcomes and treatment patterns across large geographic areas.

5.4 Limitations of the Study

There were several limitations to this dissertation research. Methodologically, the research would have benefited from including mammography centers and primary care physicians that bordered the study area. Neighboring population data for use when calculating 2SFCA scores was included, but several populations in the study area are likely served by health providers outside of the study area. For example, rural areas along the Pennsylvania-New York border in northeastern PA likely utilize healthcare providers in the small city of Binghamton, New York, less than 20 minutes from the state border. Because of data permissions, however, we were not able to include Binghamton's mammography centers and primary care physicians in the analyses. Another methodological limitation was the lack of non-automobile transportation options. In several of the urban areas of the study region, particularly the greater Philadelphia area, walking, riding bicycle, and using public transportation likely constitute a large portion of travel for certain population groups. Therefore, the most accurate 2SFCA catchment size for these populations is likely less than the maximum of 60 minutes used throughout the dissertation. Another limitation, most relevant to Chapter four, was the use of the closest mammography center or primary care physician rather than the healthcare provider that each patient actually uses. Patients with fewer provider choices because of more restrictive or no insurance may travel considerably farther for healthcare than to their closest geographic provider.

5.5 Future Research Possibilities

This dissertation work suggests numerous possibilities for further research. When comparing between different 2SFCA distance decay weights and methods to vary catchment sizes, it would be informative to include patient healthcare utilization behavior to guide the choice of 2SFCA parameters. There would likely be interaction effects with socioeconomic status and automobile ownership as well. For example, urban, economically distressed populations that do not own automobiles might have faster decay weights and smaller catchment sizes (i.e., their travel impediment is greater and total length of distance they are able to travel for healthcare is shorter). A mixture of population-level demographic data (e.g., U.S. Census) could be used to estimate car ownership. Health insurance claims data could be used to identify how far patients are actually traveling for different healthcare providers. Qualitative data would also be helpful to identify patient preferences and the socioeconomic factors that interact with healthcare travel decisions. In reference to the clinical indicators studied in Chapter four, more research is needed to understand the range of factors that influence late stage breast cancer diagnosis. This dissertation further refined how spatial access to care impacts the likelihood of late stage diagnosis, but the interactions between spatial access and socioeconomic variables needs further exploration. Similarly, the research in Chapter four found that only 46.9 percent of eligible women in our sample received adjuvant hormonal therapy. Qualitative research and detailed patient clinical history reviews would be helpful to understand the low rate of recommended adjuvant therapy.

5.6 Overall Study Impact

In conclusion, this dissertation provided a template for future public health research examining spatial access to care in Appalachia. Researchers in Appalachia can now make more informed decisions about which spatial access measures to employ, how to apply the various parameters of each measure, and how to interpret the resulting map of spatial access scores.

Additionally, this dissertation provided an example of how spatial access can be incorporated into models predicting cancer outcomes and treatment patterns of interest.

5.7 References

Blackley D, Behringer B, & Zheng S (2012). Cancer mortality rates in Appalachia: descriptive epidemiology and an approach to explaining differences in outcomes. *J Community Health*. 37(4): 804-813.

Freeman, A.B., Huang, B., & Dragun, A.E. (2012). Patterns of Care With Regard to Surgical Choice and Application of Adjuvant Radiation Therapy for Preinvasive and Early Stage Breast Cancer in Rural Appalachia. *Am J Clin Oncol*. 35(4): 358-63.

Hall, H.I., Uhler, R.J., Coughlin, S.S., & Miller, D.S. (2002). Breast and cervical cancer screening among Appalachian women. Cancer *Epidemiol Biomarkers Prev.* 11:137–142.

McGrail MR (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *Int J Health Geogr*. 16:11-50.

Paskett, E.D., Fisher, J.L., Lengerich, E.J. et al. (2011). Disparities in underserved white populations: the case of cancer-related disparities in Appalachia. Oncologist. 16(8): 1072-81.

Pollard K & Jacobsen LA (2013). The Appalachian Region: A Data Overview from the 2007-2011 American Community Survey Chartbook. Retrieved from http://www.prb.org/pdf13/appalachia-census-chartbook-2012.pdf

Wang F (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Ann Assoc Am Geogr.* 102(5): 1104–1112.

Winger, P.A., Tucker, T.A., Jamison PM, et al. (2003). Cancer in Appalachia, 2001-2003. *Cancer*. 112:181–192.

APPENDICES

Appendix 1: The University of Michigan IRB Approval

UNIVERSITY OF MICHIGAN
eResearch.umich.edu

Health Sciences and Behavioral Sciences Institutional Review Board (IRB-HSBS) • 2800 Plymouth Rd., Building 520, Room 1170, Ann Arbor, MI 48109-2800 • phone (734) 936-0933 • fax (734) 998-9171 • irbhsbs@umich.edu

---

To: Rajesh Balkrishnan

From:

Thad Polk

Cc:

Rajesh Balkrishnan
Jennifer Griggs

Subject: Scheduled Continuing Review [CR00043878] Approved for [HUM00062883]

SUBMISSION INFORMATION:
Study Title: Pharmacotherapy Evaluation Tools for Improving Breast Cancer Outcomes in Rural Appalachia
Full Study Title (if applicable): Pharmacotherapy Evaluation Tools for Improving Breast Cancer Outcomes in Rural Appalachia
Study eResearch ID: HUM00062883

SCR eResearch ID: CR00043878
SCR Title: HUM00062883_Continuing Review - Thu Oct 2 11:27:03 EDT 2014
Date of this Notification from IRB:10/6/2014
Review: Expedited
Date Approval for this SCR: 10/3/2014
Current IRB Approval Period: 10/3/2014 - 10/2/2015
Expiration Date: Approval for this expires **at 11:59 p.m. on 10/2/2015**
UM Federalwide Assurance:FWA00004969 (For the current FWA expiration date, please visit the UM HRPP Webpage)
OHRP IRB Registration Number(s): IRB00000245

Approved Risk Level(s) as of this Continuing Report:

Name	**Risk Level**
HUM00062883	**No more than minimal risk**

NOTICE OF IRB APPROVAL AND CONDITIONS:
The IRB HSBS has reviewed and approved the scheduled continuing review (SCR) submitted for the study referenced above. The IRB determined that the proposed research continues to conform with applicable guidelines, State and federal regulations, and the University of Michigan's Federalwide Assurance (FWA) with the Department of Health and Human Services (HHS). You must conduct this study in accordance with the description and information provided in the approved application and associated documents.

APPROVAL PERIOD AND EXPIRATION DATE:
The updated approval period for this study is listed above. Please note the expiration date. If the approval lapses, you may not conduct work on this study until appropriate approval has been re-established, except as necessary to eliminate apparent immediate hazards to research subjects or others. Should the latter occur, you must notify the IRB Office as soon as possible.

IMPORTANT REMINDERS AND ADDITIONAL INFORMATION FOR INVESTIGATORS

APPROVED STUDY DOCUMENTS:
You must use any date-stamped versions of recruitment materials and informed consent documents available in the eResearch workspace (referenced above). Date-stamped materials are available in the “Currently Approved Documents” section on the “Documents” tab.

In accordance with 45 CFR 46.111 and IRB practice, consent document(s) and process are considered as part of Continuing Review to ensure accuracy and completeness. The dates on the consent documents, if applicable, have been updated to reflect the date of Continuing Review approval.

RENEWAL/TERMINATION:
At least two months prior to the expiration date, you should submit a continuing review application either to renew or terminate the study. Failure to allow sufficient time for IRB review may result in a lapse of approval that may also affect any funding associated with the study.

AMENDMENTS:
All proposed changes to the study (e.g., personnel, procedures, or documents), must be approved in advance by the IRB through the amendment process, except as necessary to eliminate apparent immediate hazards to research subjects or others. Should the latter occur, you must notify the IRB Office as soon as possible.

AEs/ORIOs:
You must continue to inform the IRB of all unanticipated events, adverse events (AEs), and other reportable information and occurrences (ORIOs). These include but are not limited to events and/or information that may have physical, psychological, social, legal, or economic impact on the research subjects or others.

Investigators and research staff are responsible for reporting information concerning the approved research to the IRB in a timely fashion, understanding and adhering to the reporting guidance (http://medicine.umich.edu/medschool/research/office-research/institutional-review-boards/guidance/adverse-events-aes-other-reportable-information-and-occurrences-orios-and-other-required-reporting), and not implementing any changes to the research without IRB approval of the change via an amendment submission. When changes are necessary to eliminate apparent immediate hazards to the subject, implement the change and report via an ORIO and/or amendment submission within 7 days after the action is taken. This includes all information with the potential to impact the risk or benefit assessments of the research.

SUBMITTING VIA eRESEARCH:
You can access the online forms for continuing review, amendments, and AE/ORIO reporting in the eResearch workspace for this approved study, referenced above.

MORE INFORMATION:
You can find additional information about UM's Human Research Protection Program (HRPP) in the Operations Manual and other documents available at:http://hrpp.umich.edu.

Thad A. Polk

Thad Polk
Chair, IRB HSBS

Appendix 2: Data User Agreement

Agreement to use CMS data for the "Breast Cancer Research" Project:

I ____Joseph Donohoe________________ unique name __joedonoh___ agree to the following:

1. Use is limited to your specific (your IRB approved proposal) project in collaboration with your mentor.
2. These data cannot be placed or stored on a portable device or personal computer. (see 3 and 8)
3. Data must be stored and used on a departmental centralized server that sits behind a firewall intended to protect sensitive personal information. This folder is on the Research Drive and is called *Breast Cancer Project.* None of the data may be stored on the C drive or any other network locations.
4. Data should not be stored on removable media without permission from Dr. Anderson.
5. You will not permit others to use or view the raw data (except your faculty mentor).
6. There can be no emailing of the data.
7. All computers will be protected with an automatic screen lock and password re-entry.
8. The use of a personal laptop as a remote desktop will only be acceptable if you do not transfer files from your remote session onto your local C drive in the laptop.
9. You must take the HIPAA and Security Awareness Training per Data Use Agreement 25507. https://maislinc.umich.edu/maislinc/app/taxonomy/learnerSearch/LearnerSearch.aspx?RootNodeID=-1&NodeID=195&UserMode=0 Print out the certificate of competition and attach.
10. If you suspect any unauthorized access or use of the data -- you will report to your supervisor and Pharmacy ITS.
11. The user of the data has read the Data Use Agreement 25507 and agrees to those terms.
12. This document will be resubmitted each year.

________________________ ________09/10/2014________

Signed Date