2014

Effects of Locality and Risk of Late Stage Breast Cancer Diagnosis in Kentucky Females, 2001-2011

Michael Jacob Sither
University of Kentucky

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Michael Jacob Sither, Student

Stephen Wyatt, DMD, MPH, Major Professor

Dr. William Pfeifle, Director of Graduate Studies

Effects of Locality and Risk of Late Stage Breast Cancer Diagnosis in Kentucky Females, 2001-2011

Capstone Project Paper

This paper is submitted as a portion of the requirements to complete a degree of Master of Public Health with a concentration in Epidemiology from the University of Kentucky

By

Michael Jacob Sither, OMS-IV

Lexington, KY

Dean Stephen Wyatt, DMD, MPH Committee Chair

Dr. Tom Tucker, PhD, MPH

Dr. Douglas F. Scutchfield, MD

Table of Contents

I. Abstract

Objective:

Breast cancer survival is heavily dependent on stage at diagnosis. Recent evidence supports rural women are more likely than urban women to be diagnosed with advanced stage disease but historically these findings have been mixed when investigated at the state level. Cancer surveillance in the Appalachian region suggests Kentucky females are at significantly greater risk of premature breast cancer mortality due to consistently lower prevalence of screening mammography utilization. In the literature, disparities in breast cancer outcomes have been largely explained by population demographics, area-based measures of socioeconomic deprivation, shortages of referring providers as well as differences in spatial access to mammography based on a women's place of residence. Results from community-based participatory research in Appalachian-designated counties of KY have uncovered concordant cultural beliefs, knowledge and attitudes towards breast cancer prevention, which may be contributory to the level of risk present in these communities. This study investigates factors associated with late stage breast cancer diagnosis among KY women aged 40 years and older to determine if county level factors may impact breast cancer outcomes among women, even when controlling for known risk factors of advanced disease. Since the high level of socioeconomic distress in KY may serve to mask the effects of a woman's place of residence on the risk of late stage presentation, we will assess the effects of available census tract estimates to better understand how socioeconomic context modulates risk in the rural and Appalachian communities of KY. Secondary objectives include assessing for the presence of effect modification based on available measures of locality and monitoring for dose response

effects which may be present with increasing poverty, declining educational attainment and other area-based indicators of the economic hardship present in these communities.

Methods:

This study utilizes all incident cases of *in situ* and invasive breast cancer diagnosed from 2001-2011 in the Kentucky Cancer Registry (KCR). Two separate staging criteria were used to create outcome variables of interest in an attempt to compare measures of association with independent predictors and further support the validity of findings. Utilizing the American Joint Committee on Cancer's (AJCC) TNM staging system and the National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) summary staging criteria, a total of 23,100 and 23,769 women respectively with no previous history of cancer, aged 40 and older at diagnosis with known cancer stage were included in the study samples. Estimates for socioeconomic context were obtained from the State Data Center and linked to the health record using geocodes for the county of residence at time of breast cancer diagnosis. Hierarchical multivariate logistic regression models were used to analyze the relationship between measures of locality and advanced stage diagnosis. Using a step-wise backward elimination strategy, a final mixed effects random intercept model was fit to evaluate the effects of county-based socioeconomic indicators on the odds of advanced breast cancer among rural and Appalachian women of Kentucky.

Results:

KY females with Appalachian residence were most likely to reside in counties with the highest levels of socioeconomic deprivation. Clear dose response relationships were established between increasing levels of socioeconomic distress in the county of residence and the risk of advanced breast cancer. After adjusting for age, race, insurance and marital status, the odds of a late stage breast cancer diagnosis were significantly greater in KY females with rural residence compared to urban residence *(adjusted OR = 1.13, 95% CI = 1.06 – 1.21)* and Appalachian residence compared to non-Appalachian residence *(adjusted OR = 1.22, 95% CI = 1.13 – 1.31)*. The validity of these findings was corroborated by results from the SEER-based analysis. Additive effect modification was present using an interaction term for residence type in the SEER-based analysis, suggesting the odds of advanced stage breast cancer was greatest in rural Appalachian women when compared to urban, non-Appalachian women *(adjusted OR = 1.25, 95% CI = 1.17 – 1.34)*. Results from the mixed model analysis indicated differences in county income levels largely explained the effect of rural residence on the odds of advanced diagnosis, but was not a significant confounder for Appalachian residence. When controlling separately for county level poverty and education in the mixed models, the effect of rural and Appalachian residence on the odds of a late stage diagnosis was no longer significant. However, even when controlling for poverty and education, Appalachian residence still increased the odds of late stage diagnosis compared to non-Appalachian residence, *(adjusted OR = 1.046, 95% CI= 0.92 – 1.18)* though the result was not statistically significant.

Conclusion:

Although age, race, insurance type and marital status may help explain advanced breast cancer diagnoses, the influence of high levels of socioeconomic distress present in rural and Appalachian communities of KY largely account for the effects of residence on the risk of advanced breast cancer. Since individual socioeconomic status is unavailable in the individual health record, these measures of locality will serve as an important tool in the surveillance of vulnerable populations and gauging progress of strategies to rectify disparities in breast cancer outcomes in our state. Community-based interventions designed to increase mammography utilization should continue to target impoverished and highly uneducated women, especially in the most rural and isolated counties of KY Appalachia. Finally, innovative health policies, which fosters both educational opportunity and economic development in distressed counties of rural, and Appalachian KY will likely have the greatest impact on breast cancer outcomes long term.

II. Introduction

As the second most common cause of cancer death among women in the United States, breast cancer continues to remain a significant source of human suffering and financial burden to our health system.[1] Excluding cancers of the skin, cancer of the breast is the most frequently occurring primary site among US and Kentucky females[1,2] and the American Cancer Society (ACS) estimates that 1 in every 8 women will be diagnosed at some point in their lifetime.

The purpose of this capstone is to use primary cases of breast cancer diagnosed in KY women 40 years and older during the period 2001-2011 to assess the association

between socioeconomic, demographic and geographic risk factors with stage of breast cancer at the time of diagnoses. The primary aim of this case control study is to determine if KY females with a rural or Appalachian county of residence are at greater risk of a late stage diagnosis compared to those from an urban or non-Appalachian county respectively, even when controlling for known risk factors of breast cancer; which include age, race, insurance type, family history, parity, smoking status and quantified history by pack years, marital status as well as county-based measures of socioeconomic deprivation, such as the percent of adults 18 years and older living in poverty, average median household income, average per capita income, percent rate of unemployment, and percent obtaining a high school level of education or higher.

My hypothesis is that women from either a rural or Appalachian-designated county of residence will be at greater risk of a late stage diagnosis compared to women from urban or non-Appalachian counties of residence respectively, even when controlling for possible confounders not present in the health record, such as the socioeconomic context of a women's county of residence which will serve in this study as a proxy measure of socioeconomic status. We intend to demonstrate that measures of locality are important determinants of breast cancer outcomes in KY females and the effects of locality may be associated with excess risk not explained by the social determinants of health. Furthermore, I believe women from the rural, more isolated counties of Appalachia will experience a magnitude of risk beyond either factor alone, suggesting the presence of effect modification by these measures of locality. As supported by previous literature, I also believe that age, race, insurance type, marital status, and smoking will be significant predictors of late stage disease and that increasing levels of socioeconomic depravity

found in a woman's county of residence will have a dose response effect on the magnitude of risk for advanced breast cancer.

Identifying subpopulations of KY females who are more vulnerable to premature breast cancer mortality based on geographic and demographic characteristics of residence will support the continued argument for expanded access to screening resources as well as more targeted surveillance of high risk groups. Identifying and rectifying health disparities represents a critical function of cancer prevention and control. The findings of this study may enhance the service delivery efforts of the KY Women's Cancer Screening Program (KY-WCSP), lend further support to previous findings of the Appalachian Cancer Control Network as well as inform policy decisions regarding key cancer control initiatives in the Commonwealth of KY.

III. Literature Review

The Burden of Breast Cancer in KY & US Women by Measures of Locality

Consistent with the previous five year estimate from 2001-2005, the average age-adjusted incidence of invasive breast cancer among KY females from 2006-2010 remained slightly lower than the United States as a whole, at 121.3 and 121.9 respectively.[1] According to the KY-WCSP, the overall incidence of invasive breast cancer during this period has been declining annually and data from the KY Cancer Registry (KCR) suggests the decline has been occurring in both the rural and urban segments of the female population since 2000,[4,5] suggesting overall, women in KY appear to be collectively benefiting from current levels of screening. However, when KCR data has been compared with estimates from the national surveillance

epidemiology and end results (SEER), a collection of population based registries comprising 25% of the US population, the ratio of early stage versus late stage breast cancer diagnoses appear to be consistently lower in Kentucky women compared to the rest of the nation[10], suggesting certain subgroups of KY females are yet to fully benefit from clinically appropriate surveillance. In a large study encompassing data from 29 population-based cancer registries participating in the North American Association of Central Cancer Registries, researchers found despite unprecedented declines in invasive breast cancer incidence among US women from 2001-2004, these trends were not equally shared by white women living in rural, middle or lower income counties when compared to urban and affluent counties respectively.[35] The disparity in this rate decline led researchers to conclude that despite recent changes to the clinical application of hormone replacement therapy, a known risk factor for hormonally-driven cancers, inadequate screening surveillance in rural and lower income counties still likely accounts for higher rates of invasive breast disease present in these populations.[35]

The incidence of invasive breast cancer has previously been reported as lower in rural areas than urban areas[30, 45] and in the Central Appalachia region from 2001-2003, researchers found the rate of breast cancer incidence to even be lower than the US rate overall, though at the time, availability of high quality cancer data was still in the infancy of collection.[47] However, more recent findings suggest a more ominous effect of Appalachian residence on breast cancer survival. In study based on SEER data, breast cancer mortality was consistently higher among Appalachian Kentuckians when compared to the US population overall, despite lower incidence, suggesting inadequate

screening for early detection is likely responsible for the discrepancy in rates of invasive breast cancer among these women.[15, 33]

Women from Appalachian counties of KY may be at greater risk of late stage breast cancer for a variety of reasons, including both disparities in healthcare-related access, demographic, socioeconomic and cultural factors.[2-3, 6, 12, 14-15] Historically, Appalachia has been underserved by the health system and knowledge of cancer screening and its benefits tend to be much lower in the Appalachian community, which are uniformly predicted by both age and educational attainment.[6, 14-15] Furthermore, possessing insurance was an important predictor of an individual's likelihood to obtain screening, while family history of cancer was not, suggesting higher proportions of uninsured women likely contributes to lower rates of mammography utilization present in Eastern KY.[12] In a population-based study of patients with newly diagnosed breast cancer, rates of mammography and clinical breast examinations in Appalachian women were significantly lower than national averages and the targets set by Healthy People 2010 even after controlling for other predictors.[13]

Demographic and Socioeconomic Effects of a Rural and Appalachian Residence in KY

Kentucky is a geographically distinct state, considering it is both heavily rural and a significant portion of the state, 54 of its 122 counties, are designated as part of the Appalachian region by the Appalachian Regional Commission (ARC). Nestled in heart of the Central Appalachia, these counties of Eastern KY are known to suffer from disproportionately high levels of poverty and unemployment, lower levels of educational attainment, poor rates of health insurance coverage and significant shortages of health

care providers.[45-46] In 2012, the majority of Appalachian counties of KY, 46 out of 54, were designated as health professional shortage areas for primary medical care suggesting provider availability may still contribute to lower rates of screening present in these areas. According to DuBard et al., physician shortages lead to inadequate levels of patient counseling and ultimately, insufficient health promotion, risk awareness and preventive medical service utilization within affected communities.[14-15]

The Appalachian region itself is a large, heterogeneous region of the Eastern United States, marked by disproportionately higher poverty levels, shortages of health professionals, and overall cancer incidence and mortality.[3] Forty-two percent of the Appalachian region's population is rural, more than twice the national average.[3] Though much progress has been made over the later part of the 20th century, many parts of Appalachia remain geographically isolated, economically distressed and lacking in basic infrastructure, such as water and sewer systems.[3, 12] According to the ARC, as of the 2010 fiscal year, 116 of the 420 Appalachian-designated counties were considered high poverty (those with rates 150% higher than national averages), and the vast majority of these distressed counties fall within Kentucky and the Central Appalachian Region. During the same period, unemployment in the Appalachian region has exceeded both state and national averages.[7] Per capita personal income, average wages and salary were found to be 20% lower in Appalachia compared the entire U.S. population in 2007, highlighting the harsh economic conditions present in this region of the country.[7]

Despite tremendous improvements in educational attainment over the past several decades, the ARC still estimates the proportion of adults with a college degree in Appalachia to be only two-thirds of the national average and in Central Appalachia, the

figure is well below the national average at an estimated 25% compared to 50% for the US population overall. Due in part to continuing outmigration of the college and working-age populations, roughly 80% of Appalachian counties had lower population growth than the nation as a whole.[7] Moreover, as the baby boomer generation continues to age, more individuals continue to seek retirement outside of metropolitan areas and the growth of the population over age of 65 in Appalachian counties continues to exceed the rest of the nation. This shift in demographics may potentially contribute to a greater need for medical services, especially screening and chronic disease management care in rural and Appalachian counties.

Although the greater Appalachian region is culturally and ethnically diverse, Appalachian Kentucky is predominately white and poor. Consistent cultural traditions include conservative religious beliefs, a focus on family, and the importance of knowing family history.[12] As of 2010, the majority of Kentucky's Appalachian counties were designated as economically distressed by the ARC and most of which are considered rural by the U.S. Department of Agriculture's Rural-Urban Commuting Area Codes (RUCA) and non-metropolitan by current census track designations. In general, women living in rural areas of the United States are typically older, more likely to be White, less likely to be single, less educated, more likely to report fair or poor general health status as compared with good or excellent general health status, less likely to have health insurance, and more likely to have a lower household income, than women in urban and metropolitan areas of the country.[29, 30]

Effects of Locality on the Utilization of Screening Mammography

Several factors in the literature have been reported to influence a woman's likelihood of maintaining a regular schedule of mammogram screening for breast cancer: rural or urban residence[21]; convenience of accessing a mammogram including transportation[23;] distance from the nearest mammography facility[10, 23] age[23]; race[38]; median county income and levels of education.[39] In a study pooling Behavioral Risk Factor Surveillance System (BRFSS) sample estimates for the entire Appalachian region, researchers found Appalachian women tended to have significantly lower rates of mammograms, clinical breast exams, and Pap screening even after adjustment for other predictors.[6] Qualitative research investigating the disparity in mammography usage, have identified attitudes, beliefs and qualities specific to the contexts of rural Appalachian communities, which may contribute to the underutilization of screening even when highly accessible.[20] These include the misconception that breast cancer has noticeable symptoms before diagnosis, the fear of finding cancer, fatalistic beliefs regarding the association of cancer with death,[18-19] as well as concerns about maintenance of privacy and confidentiality within the confines of a close-knit community.[15]

Studies of breast and cervical cancer screening in the United States consistently demonstrate that women with better access to health care, including those with health insurance coverage or a higher family income, are more likely to have recent screening tests.[6] When using Beale codes to capture effects of locality based on the rural-urban continuum**,** researchers noted women from non-rural counties were more likely to have had a recent mammogram than women living in rural, more geographically isolated counties of residence.[21]

More recent analysis of KY BRFSS data suggests rates of screening mammography in Appalachian females are consistently lower than national levels.[5] In Kentucky, 2006 BRFSS estimates indicate the disparity in mammography utilization among women residing in Appalachian counties of KY was still fairly significant, with an estimated 37.8% of Appalachian women 50 and older reporting not having an updated biennial mammogram, compared to only 26.5% of Non-Appalachian women.[5] In 2008, although the prevalence of screening mammography in Kentucky females closely mirrored the national average (76%), when stratifying by level of income or education, clear gradients emerge in the prevalence of mammography utilization, with only 61.8% earning less than $15,000 per year and 62.2% with less than a high school education reporting a mammogram in the past two years.[4] According to Lyttle et al., the greatest concerns identified by Appalachian women regarding mammography were health care costs and lack of health insurance. Interestingly enough, fear and embarrassment were the most consistent barriers towards breast and cervical cancer screening in this population, suggesting attitudes and cultural norms may also play a contributory role in the etiology of a late stage diagnoses.[22] The body of evidence strongly supports Appalachian females in KY are at significant disadvantage to maintaining appropriate levels of screening based on a combination of both the influence of socioeconomic position, factors related to health care delivery and cultural differences in use of available resources. In addition, more recent studies employing geographic information technology (GIT) lend credence to effects of geographic isolation, lack of transportation and lower rates of health insurance found in these areas, and these limitations may continue to thwart efforts to improve breast cancer screening for early detection in both Appalachian[14-15] and rural

counties of Kentucky.[10]

Effects of Locality on Stage at Breast Cancer Diagnosis

Breast cancer survival is highly dependent on the stage at diagnosis with 5-year survival estimates ranging from 100% for in situ (Stage 0) to as low as 15% for metastatic disease (Stage 4). [8, 9] Likewise, stage at diagnosis has been associated with many factors including race, obesity, history of mammography, method of tumor detection, insurance status, distance from residence to nearest hospital, residence in rural areas, nursing homes and areas of high socioeconomic depravity.[25-28] In a 2003 analysis comparing KCR and SEER data, researchers previously reported that the burden of late stage breast cancer was higher in rural than urban women and highest of all in rural counties of the Appalachian region, suggesting that disparities in breast cancer outcomes in KY may be partly attributable to geographic characteristics of a women's place of residence.[2,16]

The rural disadvantage is well documented in the literature but historically the association between stage at diagnosis and measures of locality has been mixed.[30]In a recent study of New Hampshire cancer registry data, researchers did not find an association between rural residence and late stage at diagnosis based on the RUCA classification.[23] Similarly, Liff, Chow and Greenberg (1991) concluded that rural residents in Georgia were neither more or less likely to be diagnosed at later stage breast cancer than urban residents, while Farley and Flannery et al., (1989) even noted a beneficial effect of rural residence.[30] A 2012 systematic review and meta-analysis supports the more consistent effect of residence type on stage at diagnosis, with rural women being more likely than their urban counterparts to be diagnosed with advanced

breast cancer.[29] This finding emphasizes the importance of improving early detection of breast cancer in rural populations since previous trends in healthcare delivery have consistently demonstrated inequalities in the overall quality, availability and accessibility of medical services for rural women.[30]

Recent changes to clinical practice guidelines may harbor potential to alter trends in mammography utilization relative to a woman's age since the United States Preventive Services Task Force (USPSTF) opted in 2009 to defer initiating biennial breast cancer screening until age 50 compared to starting at 40 years of age as recommended by the ACS and the American College of Obstetrics and Gynecology.[38, 55-56]Although screening mammography has remained the gold standard for the early detection of breast cancer for over 25 years[55], it is now being complemented with other emerging modalities when clinically necessary, considering certain histologic subtypes of breast cancer, such as invasive lobular carcinoma, are more likely to be missed due to preferential development within the confines of normal tissue architecture. Despite a less than desirable sensitivity, especially in younger women with dense breast tissue, a recent Cochrane Review still reports mammography to be associated with a 15% relative risk reduction in breast cancer mortality and an absolute risk reduction of 0.05%. [50] Considering inadequate screening for early detection is implicated in the elevated breast cancer mortality rates found in Appalachian Kentucky[15], improving mammography utilization in this vulnerable population remains a grave public health concern since a mammogram performed every 1-2 years has been shown to reduce breast cancer mortality by approximately 20%-25% over a 10-year period.[53]

Ensuring all Kentucky women share equitable access to and optimal use of screening mammography through sufficient exposure to cancer awareness and health promotion activities are key components of current community-based strategies. Considering previous failures with achievement of Healthy People 2010 objectives, it has become more evident that in order to be successful in routing out cancer disparities, these efforts must be complemented with larger scale improvements in health system delivery and performance, especially in problematic areas.[37] Continuing to optimize policy decisions through research and evidence-based practice are indelibly the most crucial elements in the war on cancer.

Known Individual Level Predictors of Late Stage Breast Cancer

Previous research demonstrates the incidence of late-stage breast cancer is elevated in older women[11]; those who have never been married[23]; those who reside in low-income environments as well as racial and ethnic minorities.[25-26, 32] Disparities in advanced breast cancer are consistently noted for African Americans and Hispanics[28], women suffering from obesity[25, 28]; those with a rural type of residence or among women with a significant geographic disadvantage to mammography access, measured by the distance to the nearest screening center.[10] However, when controlling for poverty, insurance status, history of mammography, method of tumor detection and obesity, these racial and ethnic disparities have been largely explained.[42-43, 59]

Although marital status was previously used in an attempt to control for the confounding effects of a woman's parity on her risk of reproductive cancers, since nulliparity is an established risk factor for invasive breast and uterine cancers[41]; its

consistent implication with cancer outcomes has led marital status to emerge as an important confounder in cancer research.[51, 54] For example, when accounting for women who are currently or have ever been married at the time of diagnosis, multiple studies have demonstrated a protective effect against advanced breast cancer, and based on a retrospective cohort of Medicare patients, researchers estimate unmarried women were 24% more likely not to receive definitive therapy and 25% more likely to die from breast cancer when compared to married women, suggesting that marital status is an important predictor of survival, likely due to the health benefits derived from increased social support and better social networks.[51]

Childbearing, particularly at a younger age, and the practice of breast-feeding has been consistently shown to reduce a woman's risk of invasive breast cancer in the later years of life.[41] Since having children later in a woman's reproductive years has become a defining characteristic of contemporary fertility in more affluent societies, we will also attempt to control for a women's parity, considering potential differences that may exist in the socio-demographics between residence types. Since the rate of divorce, separation and spouse bereavement may differ by locality, we will also attempt to control for marital status with a previously used method of ever married versus never married[10, 23] to improve estimates of association with the outcome variable of interest.

Although the International Agency for Research on Cancer concluded in 2009 there was limited evidence that breast cancer is directly attributable to tobacco use, being a current or former smoker has been implicated in the etiology of a late stage diagnosis[26] and heavy cumulative smoking history has also been associated with fatal breast cancer outcomes.[28, 44] For example, Saquid et al found that heavy lifetime smoking exposure,

indicated by former smokers with 20 or more pack years, were found to be at 77% greater risk of breast cancer-specific mortality than non-smokers, leading to the recommendation that quantifiable measurements of smoking be used preferentially when evaluating associations with breast cancer mortality.[44] Considering smoking prevalence in rural and Appalachian communities of KY may be higher than urban and non-Appalachian counties, we will attempt to account for differences in tobacco use between these populations when modeling other known predictors for advanced breast cancer.

Racial Disparities in Breast Cancer

Although breast cancer incidence is somewhat lower among African American women than among White women in the United States, mortality is consistently higher among African Americans.[26] In this sense, being both African American and a resident of a rural county could present a form of 'double jeopardy' regarding the risk of poor outcomes. From 2005-2009, five-year averages demonstrate that the incidence of breast cancer was actually higher among KY African Americans compared to KY Whites (131.2 vs. 120.5), a finding inconsistent with current national trends.[5] The disparity likely reflects improved surveillance during this period among a population known to be at greater risk of late stage presentation. Though overall mortality rates for breast cancer have been steadily declining over the past decade, a recent report from the KWCSP implies a persistent racial disparity in breast cancer survival in KY. From 2005-2009, mortality rates actually increased for African American females (from 23.3 to 40.3) compared to a slight decline in mortality for Whites (23.8 to 22.3). [5] Considering 43% of cases among African Americans in this period were diagnosed at late stage compared to

36% of Whites cases, health inequalities in early detection and delayed treatment among racial and ethnic minorities is likely still problematic in Kentucky.

White women typically have higher age-adjusted incidence of breast cancer than do racial and ethnic minorities.[13, 25] However, when compared to Whites, the literature consistently notes disparities in breast cancer outcomes among African American and Hispanic women, though much of these racial differences can be explained when controlling for insurance status, poverty, history of mammography, and obesity.[28] Despite lower overall incidence, African American women are more likely than their White and Hispanic counterparts to die from breast cancer.[13] A recent review noted that although it appears certain there is a role for socioeconomic deprivation as a factor contributing to racial differences in breast cancer prognosis, a strong biological argument exits for the importance of more intrinsically aggressive genetically or epigenetically determined nature of tumors in African American women.[28] Considering women in Appalachian counties of KY are mostly White and differences in tumor biology may exist across racial and ethnic strata, controlling for race is necessary to assess if risk of late stage diagnosis is due to an underlying effect of locality or simply differences in population demographics.

Type of Health Insurance and Risk of Advanced Breast Cancer

Approximately 13% of Appalachians are considered to be medically indigent.[7] In Kentucky, this figure may be much higher considering poverty rates in Appalachian counties of residence are both higher and more heavily concentrated in the Central Appalachian region.[27] Substantial evidence exists in the literature that women are more

likely to be diagnosed with late stage breast cancer if they lacked health insurance[27, 36] or if they were on a public health insurance option, such as Medicaid or Medicare.[36] Kuzmiak et al., 2008, demonstrated that uninsured patients had a 66% higher likelihood of presenting with late stage disease and a larger tumor size compared to patients with insurance. Other studies note that health insurance type may modify risk of disease severity at time of diagnosis, suggesting potential differences in the quality or timeliness of care provided.[55] Compared to women with private insurance, uninsured women and those on Medicaid had a greater likelihood of regional and distant stages (2-4) compared to local stage at diagnosis (1). [27] In the past, health insurance has been used as a proxy measure for both socioeconomic status and as a direct estimate of access to care.[26] The literature also finds possessing health insurance and payer type affects survival following a breast cancer diagnosis in Kentucky.[39-40] McDavid et al., 2003, found 3-year relative survival for breast, lung/bronchus and colon cancer was highest for private and the lowest for unknown insurance, with survival decreasing 33.1% for breast cancer outcomes between primary payers. In this study, determining if possessing insurance or insurance type modifies the risk of a late stage diagnosis is essential for characterizing the underlying risk for Appalachian and rural women since differences are likely to exist in payer type based on locality.

Social Determinants and Risk of Late Stage Breast Cancer

In Appalachia Kentucky, both the socioeconomic environment and socioeconomic position of women may play a role in access to and proper utilize of cancer screening. Few studies have operationalized socioeconomic status (SES) at the individual level, though Lantz et al., 2006, determined that even after controlling for

individual level SES, age and study site, that the risk of late stage disease was still greater for Hispanics and African Americans in urban areas, though individual measures of income and education were not significantly associated with late stage diagnosis in the multivariate analysis. Since individual SES is not available in the health record, typically researchers have controlled for the influence of the area-based measures of socioeconomic environment as a proxy for a woman's SES.[58-60] Several large-scale studies evaluating predictors of stage at diagnosis have reported a significant association of census-tract derived measures of household median-income, poverty and education with risk of advanced diagnoses.[52, 58-59]Since women with lower levels of income and education typically have worse compliance with age appropriate screening practices for breast and cervical cancer, it is indeed plausible that women in areas with lower levels of educational attainment and higher levels of socioeconomic distress would collectively be at greater risk of late presentation, and evidence from large scale studies supports this claim.[42-43] Since estimating the effect of residence type on risk of advanced breast cancer is our primary objective, we will attempt to control for the collective effects of county level poverty, unemployment, economic distress as indicated by average per capita or median household income, as well as decrements in educational attainment based on previously used methods.[10, 25-26]

IV. Research Questions

The current study intends to assess if KY females from a rural or Appalachian designated county of residence are at increased risk of an advanced breast cancer diagnosis compared to their urban and non-Appalachian counterparts, even when controlling for known predictors of late stage disease. The study also intends to quantify

levels of risk associated with known risk factors of invasive breast cancer including age, race, insurance type, parity, marital status, smoking and family history. We will assess if effect modification is present in the magnitude of risk by residence type using an interaction term. Finally we will also assess for dose response effects of county-based estimates of socioeconomic deprivation on the risk of advanced breast cancer and evaluate for changes in the magnitude of associations by residence type when controlling for contextual effects of poverty, income, unemployment and education in the multilevel analysis.

V. Methods

Data Sources, Study Sample and Exclusion Criteria

The primary data source will be the Kentucky Cancer Registry. An estimated 32,800 cases of invasive breast carcinoma were diagnosed in Kentucky females from 2001-2011. Inclusion criteria were all primary cases of breast cancer diagnosed in KY women age 40 years and older with known stage at time of diagnosis. Women who were diagnosed at autopsy or from a death certificate, women without diagnostic confirmation or unknown TNM or SEER summary stage were excluded from the study. Census tract socioeconomic covariates were obtained from the 2007-2011 American Community Survey (ACS) and represent five year estimates reflective of county level percent poverty rate in adults 18 years and older, average median household income, average per capita income, percent rate of unemployment and percent of adults 25 years and older obtaining a high school degree equivalent or higher. These county-based socioeconomic

indicators were linked to KCR data using geocodes for county of residence at time of breast cancer diagnosis.

Stage at Diagnosis and Independent Covariates Included for Analysis

Tumors were staged using the American Joint Committee on Cancer's (AJCC) Manual for Staging of Cancer. The fifth edition of the manual was used for years 1999-2002 and the sixth edition was used for year 2003 and beyond.[21, 22] Tumors that were *in situ* (TNM stage 0) or designated as local (TNM stage I) were considered "early stage," while tumors considered locally advanced, regional or distant (TNM stage II-IV) were considered "late stage" similar to the dichotomy used by Montella et al., 1995, and Celaya et al., 2010. Our decision to modify the primary outcome variable of interest in this manner was based on the reasoning age-appropriate mammography should be able to identify earlier stage tumors and that *in situ* and localized tumors were more likely to be asymptomatic as compared to more advanced stages. Regardless of an incomplete rate of progression to invasive disease, our decision to include *in situ* cases in the analysis was based on reasoning that surgical or medical management is clinically indicated for all histologic forms of *in situ* disease, despite the fact though these tumors are yet to demonstrate the hallmark feature of cancer behavior: invasion of the basement membrane. For comparison purposes, a secondary analysis was conducted with a dichotomous dependent variable created from the SEER summary staging criteria of 2000, with *in situ* and localized tumors treated as "early stage" and locally advanced, regional, regionally extended and metastatic tumors considered as "late stage." Furthermore, in both binary logistic regression models, an interaction term for locality was tested to assess for the presence of effect modification by residence type.

Race was categorized as White, African American, Other Minority or unknown. Since changes recently occurred to age-dependent guidelines for initiating screening mammography, two separate age covariates were created to model the effects of increasing age on stage at diagnosis. The first age covariate was previously used by Huang et al., 2009, to account for increasing effects of age on the risk of advanced breast cancer based on age-appropriate screening intervals and Medicare qualification: (1) 40-49, (2) 50-64, (3) 65-75 and (4) 75 and older. The second age covariate was previously used by both Amey et al., 1997, and Reynolds et al., 2005, to functionally trichotomize age based on menopausal status, roughly grouping together pre-, peri- and postmenopausal groups respectively.[30, 34] These age categories consist of: (1) Less than 45, (2) 45 to 54 and (3) 55 and older. Similarly, there were two covariates used to control for the influence of tobacco use: smoking status at the time of diagnosis and smoking history quantified by number of pack years. Smoking status was dichotomized as "ever smoked," which includes current and former users of cigarettes, pipes or cigars versus "never smoked." Number of pack years was trichotomized to account for the cumulative effect of heavy smoking on risk of late stage diagnosis, and includes: (1) never smoked, (2) 0-20 pack years and (3) 20 or more pack years. Health insurance was divided into 5 categories—insured, uninsured, Medicare, Medicaid and unknown. The "insured" category includes a composite of those with private insurance, such as managed care, PPO or HMO. Due to low numbers, military payers such as Veterans Affairs (VA) and CHAMPUS were collapsed into the "insured" category. To account for the effects of increasing parity on risk of late stage diagnosis, parity was categorized into high (1) 3 or more live births, low (2) 1-2 live births and nulliparity (3) no live births.

Marital status was dichotomized into a covariate previously used by Lannin et al., 1998, which combined divorced, separated and widowed women with currently married women as "ever married" and are compared to women who were "never married." To account for the effects of locality on stage at diagnosis, residence in urban versus rural counties was categorized as metropolitan or nonmetropolitan based on the rural-urban continuum code classification, also known as Beale codes previously used by Friedell et al., 2003. This scheme distinguishes metropolitan counties by the population size of their metro area and non-metropolitan counties by their degree of urbanization and adjacency to a metro area. Within this classification system, urban counties are designated by Beale codes ≤3, and rural counties, by Beale codes ≥6. Finally the ARC designation was used to determine whether a county of residence was officially an Appalachian county or not.

A second set of variables was examined to control for the influences of county composition, which may serve to mask the influence of rurality or being considered Appalachian on the risk of an advanced diagnosis. Census tract estimates for poverty rate, median household income, per capita income, rate of unemployment, and educational attainment from the 2007-2011 ACS were used for temporal consistency. For counties below 20,000 in population, these figures represent 5-year estimates to account for the larger sampling error occurring with smaller survey sample sizes. To account for the high levels of poverty, economic distress and low educational attainment in KY, these covariates were stratified into categories to reflect gradients of increasing socioeconomic depravity. County-level poverty was characterized as very low (0-0.07), low (0.07-0.13), high (0.13-0.22) or very high (>0.22) similar to categories used by Huang et al. A covariate for average median household income based on 2011 inflation

adjusted dollars was created by splitting the range of estimates into quartiles and then collapsing the two highest income quartiles for better dispersion to create a high-income county (1) $50,962 or higher, medium-income county (2) $35,153-$50,961, and low-income county (3) $19,344 - $35,152 categories. Per the recommendation of the ARC, a covariate for per capita income based on 2011 inflation adjusted dollars was created by stratifying county estimates into quartiles, which includes very high per capita income (1) $27,996 - $33,366, high per capita income (2) $22,625 - $27,995, low per capita income (3) $17,254 - $22, 624 and very low per capita income county (4) $11,883 - $17,253 categories. Unemployment rate was initially stratified by quartiles and then the two highest unemployment quartiles were combined to create 3 categories for severity of unemployment: low-unemployment (1) 5.3% - 10.5% high-unemployment (2) 10.6% - 15.7% and a very high-unemployment (3) 15.8% - 26.1%. Finally, the percent high school graduation rate or higher among adults 25 years and older was stratified by quartiles as very low (1) 56.1 % - 64.8%, low (2) 64.9% - 73.6%, high (3) 73.7% - 82.3%, and very high (4) 82.4% or higher similar to previous strategy employed by Huang et al.

Statistical Analysis

This case control study will include both descriptive statistics for the entire study sample as well bivariate descriptive statistics with respect to the two primary independent covariates of interest: Appalachian and rural residence. Since most of the covariates will be constructed to model a dose-response relationship, Pearson Correlation Coefficients were calculated to assess the degree of linear dependence between stage of diagnosis and each independent variable of interest. Chi-square tests

were also calculated to measure the likelihood that the observed difference between the dependent variable and each independent variable of interest was due to chance. All significantly correlated covariates were included in the initial model with a significance level of $p \leq 0.1$ required for retention in the model. Since our dataset included both individual and county-level data, a hierarchical random intercept logistic regression model was used to model the effects of census tract estimates for socioeconomic context with all covariates meeting retention criteria in the fixed effects model. The final mixed models were identified using a step-wise backward elimination strategy since spatial autocorrelation was likely to exist when modeling census-derived covariates for the socioeconomic context of a woman's county of residence. Goodness of fit was also tested. The univariate and bivariate analyses were performed with SPSS Statistics Version 22 and the final mixed effects binary logistic regression models were fit using R analytic software. All statistical tests were 2-sided with a P-value$\leq$0.05 used to identify statistical significance. The Institutional Review Board of the University of Kentucky approved this study (Protocol No. 14-0145-X3B).

VI. Results

Table 1.1 presents descriptive statistics of the sample population of women with known stage of breast cancer according to the AJCC's TNM Staging Criteria, stratified by type of residence. The sample included in the main analysis is (N=23,100). Based on our method of dichotomy, the total frequency of cases diagnosed "early" was 56.7% versus 43.3% of cases considered "late." Among rural women, 52.9% were diagnosed early and 47.1% were diagnosed late, compared to 58.8% and 41.2% respectively in urban women. Similarly, Appalachian women had a higher proportion of late stage

diagnoses than non-Appalachian (49.9% vs. 41.5%) and thus, a lower proportion of breast cancers diagnosed early (51.1% vs. 58.5%). The age dispersion by screening decade was relatively similar by locality, however when age was categorized by menopausal status, the rural and Appalachian strata had higher proportions of women considered as postmenopausal (69.3% and 68.2%) compared to urban and non-Appalachian women (65.3% and 66.2%). The majority of women in the sample were white (92.6%), and among rural and Appalachian subgroups, this racial homogeneity was more significant, at 97.3% and 98.3% respectively. Compared to their counterparts, rural and Appalachian women had lower frequencies of private insurance (42.1% and 40.7%) and higher proportions of uninsured (4.1% and 4.3%), Medicare (44.4% and 44%) and Medicaid beneficiaries (9.4% and 11.1%). In the Urban and non-Appalachian subgroups, there were higher proportions of women who were never married (10.1% and 9.6%) compared to rural and Appalachian women (6.4% and 6.2%). Surprisingly, rural and Appalachian women had lower proportions of reported tobacco use, however, the variable for smoking status was associated with a fair amount of missing data (N=203) and unknowns, at roughly 20% of the sample. The variable for number of pack years had an even more substantial proportion of unknowns, with almost 93% of the rural and 51% of the Appalachian women without quantified smoking history. Of note, the frequency of women smoking 20 pack years or more demonstrated an anticipated trend by residence type, with more rural than urban women (12.9% vs. 12.6%) and more Appalachian than non-Appalachian women (13.3% vs. 12.5%) reporting heavy smoking history. The variable for parity was also limited due to a significant proportion of unknowns in the dataset, with as much as 50% of rural and Appalachian women without

quantified live birth history. Urban and non-Appalachian women had higher frequencies of reporting family history of breast cancer, though roughly 30% of the Appalachian and rural subgroups were unknown.

Table 1.2 provides the frequency of census tract socioeconomic characteristics based on a woman's county of residence at the time of her diagnosis, drawing comparison of residential contexts by our measures of locality. The results suggest that, compared to urban women, there was a greater proportion of rural women residing in high poverty (58.4% vs. 7%) or very high poverty counties (20.5% vs. 0%). No rural women were considered residents of counties with very low poverty by our classification scheme. Rural residences had greater proportions than urban residences of women in counties with low (42.1% vs. 0.3%) or very low educational attainment (5.9% vs. 0%). Although the urban stratum had a higher proportion of women from counties classified in the middle tertile of median household income (77% vs. 45%), only rural counties of residence met classification into the lowest tertile of median household income among KY counties (53% vs. 0%). Similarly, when categorized into quartiles of per capita income, the rural-urban difference in the proportion of women with low and very low-income counties of residence was even more pronounced. Compared to 16% of urban women, 53% of rural women lived in low per capita income counties and roughly 37.6% of rural women were also considered to be residents of very low per capita income counties. No urban women were considered residents of very low per capita income counties and similarly, no rural women were considered residents of very high per capita income counties. As expected, rural women in our sample were also more likely to

reside in counties with high (25% vs. 3%) or very high unemployment (5.2% vs. 0%) compared to those with an urban residence.

Compared to non-Appalachian women, the Appalachian subset has higher proportions of residence in high poverty (56% vs. 15%) or very high poverty counties (30% vs. 0.2%), and the proportion of women residing in the highest quartile of county-level poverty was greater in Appalachian women than rural women (30% vs. 20.5%). No Appalachian women were considered residents among counties with very low poverty. The Appalachian sample had the greatest proportion of residents in low (58%) and very low educational attainment (8.6%) counties and this subgroup also had the lowest proportion of residence in counties with the highest level of educational attainment (5.6%). When categorized by county level income, none of the Appalachian residents were also considered to reside in counties found in the highest tertile of median household income or the highest quartile of per capita income. Compared to non-Appalachian women, the proportion of Appalachian women living in low (36% vs. 27%) and very low per capita income (54% vs. 0.3%) counties was more significant. As expected, the proportion of residences with high unemployment among Appalachian women was greater than among non-Appalachian women (27% vs. 5.7%) and similar to rural-urban differences, 7.6% of Appalachian women lived in very high unemployment counties compared to none of the non-Appalachian residents. These differences in socioeconomic context by our measures of locality suggest that the rural women of our sample were indeed more likely than their urban counterparts to reside in socioeconomically deprived counties, and that these levels of distress appear even more significant with the Appalachian residence.

According to Table 1.3, all individual level covariates met our criteria for inclusion into the initial logistic regression model based on the chi square test for independence at level of $p \leq 0.1$. The ordinal variable for age based on menopausal status was chosen preferentially over the age covariate by screening decade since Pearson correlation tests suggested there was a more significant linear relationship with the dependent variable as shown in Table 1.4 and findings in subsequent analysis that the association between age categorized by screening decade and stage at diagnosis were not statistically significant. In the preliminary model containing the dependent variable based on TNM staging, covariates for smoking status, pack years, family history and parity were subsequently removed based on our criteria for retention, revealing the associations of the final (fixed-effects) multivariate logistic regression model depicted in Table 1.5. The results suggest from 2001-2011, Appalachian women were 22% more likely than non-Appalachian to be diagnosed with late stage breast cancer, even when controlling for the effects of age, race, marital status, type of insurance and residence, and the association was statistically significant ($p \leq 0.001$). Furthermore, compared to urban women, rural women were 13% more likely to be diagnosed with late stage breast cancer and the result was significant ($p \leq 0.001$). No effect modification appeared to be present in the associations of locality with stage at diagnosis when an interaction term for rural Appalachian residence was tested separately in the model. Increasing age by our categories of menopausal status actually demonstrated a protective effect against advanced breast cancer, with perimenopasual and postmenopausal females 14% and 23% less likely, respectively, than premenopausal females to be diagnosed late and the results were significant ($p=0.001$). This result can likely be explained by the effect of increasing screening

prevalence occurring in women as they cross the 50-year old threshold for initiating mammography. Furthermore, even though the incidence of breast cancer escalates with increasing age, changes in breast tissue with age invariably allows for easier detection of tumors when present.

Even when controlling for age, marital status, insurance and residence type, African American females were 31% more likely than Whites to have a late stage diagnosis, and the result was significant (p=0.001). Compared to women who have ever been married, women who were never married were 15% more likely to be diagnosed late and the result was significant (p=0.003). Finally, insurance type remained an important predictor of an advanced diagnosis. Compared to women with private or military insurance, women who were uninsured (adjusted OR=1.93, p=0.001) or had public insurance, such as Medicare (adjusted OR= 1.15, p=0.001) or Medicaid (adjusted OR=1.65, p=0.001), were at increased odds of an advanced diagnosis and the results were statistically significant.

For comparative purposes, a secondary analysis was conducted using a primary outcome variable created from the SEER Summary Staging Criteria manual of 2000. Descriptive statistics of the sample are shown in Table 2.1. Following the bivariate analysis, all covariates were included in the initial model based on the chi square test for independence results shown in Table 2.2. The decision to remove family history, parity and smoking status from the model was based on retention criteria of $p \leq 0.1$. Despite a significant amount of missing data, the decision was made to leave the quantified smoking history in the model, since accounting for differences in the frequency of heavy smoking by residence type may provide better predictive value to the final model. As

shown in Table 2.4, when modeling the odds of late stage diagnosis, the directions of association were consistent with results from the primary analysis. Compared to non-Appalachian women, Appalachian women were 16.2% more likely to be diagnosed at late stage even when controlling for the effects of age, race, marital status, smoking, insurance and residence type and the result was significant (p=0.001). Compared to their urban counterparts, rural women were 6.9% more likely to be diagnosed at late stage, though the result was borderline significant (p=0.067). However, when an interaction term for residence was added separately to the model, the result suggested the presence of effect modification by residence type, since the magnitude of association was greater than either effect of residence type alone, and was statistically significant (adjusted OR= 1.249, p=0.001). This finding suggests that compared to women in urban, non-Appalachian counties, residents of the rural counties of Appalachian KY were 25% more likely to be diagnosed with late stage breast cancer even when controlling for the effects of age, race, marital status, smoking and insurance. Furthermore, the Appalachian and rural residence designation appeared to have an additive effect on the odds of advanced breast among KY females from 2001-2011.

In the secondary analysis, age categorized by menopausal status still appeared to exert a protective effect on odds of late stage diagnosis, with perimenopausal and postmenopausal women 15% and 21% less likely than premenopausal women respectively to be diagnosed with late stage breast cancer and the associations were statistically significant. African Americans were estimated to be 20.8% more likely than Whites to be diagnosed late and the result was significant (p=0.001). Women who were never married had a greater odds of advanced diagnosis compared to women who were

ever married (adjusted OR=1.108, p=0.041) and notably, women who smoked greater than 20 pack years had greater odds of a late stage diagnosis compared to women who never smoked (adjusted OR=1.137, p=0.005). Consistent with previous results, insurance type also remained predictive of advanced breast cancer, with uninsured women (adjusted OR=1.88, p=0.001), those with Medicare (adjusted OR=1.127, p=0.001) or Medicaid plan (adjusted OR=1.357, p=0.001) all at greater odds of advanced breast cancer compared to women with private or military insurance.

Results from the Multilevel Analysis

Table 1.5 shows the crude associations between county socioeconomic measures and advanced breast cancer among all women in our sample. Clear dose response gradients were present in measures of association between these contextual risk factors and odds of late stage diagnosis, suggesting the effects of these socioeconomic indicators were important contributors to levels of risk present in a woman's county of residence. As expected, increasing poverty and unemployment rates as well as decreasing educational achievement and income levels were highly associated with late stage diagnosis, prior to controlling for the effects of individual level covariates contained in the fixed effects model. In Kentucky, residence in very high poverty counties appeared to have the greatest magnitude of association with late stage breast cancer (Crude Odds=2.034, p=0.001), followed by residence in counties with very low educational attainment (Crude Odds=1.937, p=0.001), very low per capita income (Crude Odds=1.762, p=0.001) and very high unemployment (Crude Odds=1.642, p=0.001).

Since county-based measures for socioeconomic deprivation were likely to exhibit a high degree of autocorrelation, the mixed models were fitted using a sequential approach to control for each characteristic and depict the impact of each contextual risk factor on the magnitude of association of residence type and stage at diagnosis. The results of each sequential random intercept mixed logistic regression model are displayed separately for Appalachian and rural residence in Figures 1.1 and 1.2, since inclusion of both covariates for residence type appeared to cause multicollinearity in the initial model and thus diminished predictive value. Table 1.6 shows the cumulative adjusted associations for each contextual risk factor when controlling for age, race, marital status, type of insurance and residence. Based on the magnitude of these associations, poverty appeared to be the most significant county level risk factor for late stage breast cancer during the study period (adjusted Odds=1.163) followed by education (adjusted Odds=1.139) and per capita income (adjusted Odds=1.125).

Figure 1.1 demonstrates when controlling for age, race, marital status and insurance type, Appalachian women were 28% more likely than non-Appalachian women to be diagnosed with advanced breast cancer and the result was significant. When county-level poverty was added to the model, the magnitude of association dropped almost three-fold and was no longer significant (adjusted Odds=1.099, 95% Confidence Interval= 0.99 – 1.22). Similarly, when controlling for educational attainment, the magnitude of association decreased even more dramatically (adjusted Odds= 1.073, 95% CI = 0.94 – 1.22), suggesting levels of education may be a more influential confounder in the etiology of a late stage diagnosis in Appalachian women of KY. When controlling for per capita income, Appalachian women remained at greater risk of advanced diagnosis

and the result was significant (adjusted Odds= 1.129, 95% CI =1.01 – 1.25). When controlling for median household income, Appalachian women remained at greater risk of advanced diagnosis, though the result was marginally significant (adjusted Odds= 1.125, 95% CI =0.99 – 1.27). Unemployment had the lowest degree of impact on the association of Appalachian residence with advanced breast cancer. When controlling for the effects of both county level poverty and education, Appalachian women were 4.6% more likely to be diagnosed at late stage than non-Appalachian women, though the association was no longer significant (adjusted Odds= 1.046, 95% CI= 0.92 – 1.18).

Figure 1.2 demonstrates when controlling for age, race, marital status and insurance type, rural women were 21% more likely than urban women to be diagnosed at late stage. When poverty, education, per capita income and median household income were added sequentially to the model, the magnitude of association of rural residence with late stage breast cancer was heavily accounted for and no longer significant. When controlling for the effects of unemployment, rural residence still increased the risk of a late stage diagnosis compared to urban women and the result was significant (Adjusted Odds= 1.157, 95% CI= 1.05 – 1.27) suggesting this factor was not a strong contextual risk factor for late stage breast cancer in either rural or Appalachian women.

VII. Discussion

The results of this study are consistent with previous research findings characterizing disparities in the timeliness of breast cancer diagnoses based on residence type. Previous studies have shown associations between rural residence and advanced stage breast cancer and this study lends further support to the body of evidence suggesting residence

based on the rural-urban continuum has significant value to cancer surveillance efforts. This study is novel in that a multilevel analysis was conducted to further clarify the effects of locality on a woman's risk of advanced breast cancer. Based on our results, the effects of both rural and Appalachian residence on the risk of a late stage diagnosis were almost entirely explained by differences in levels of poverty and educational attainment between these geographically distinct parts of Kentucky. Educational attainment appears to be the most robust confounder in the association of late stage breast cancer and Appalachian residence, and consistent with previous contextual level analyses[58-59] poverty also appears to be a highly predictive of late stage presentation among Appalachian women. Although measures of economic distress, such as county-based income levels and unemployment, appear to be associated with advanced breast cancer, these socioeconomic indicators did not account for differences between Appalachian and non-Appalachian residence as strongly. On the other hand, the effects of rural residence were better explained by differences county-based income compared to Appalachian residence, though levels of poverty and education still appear to be the most significant contributors to risk of a late diagnosis in both geographic localities of Kentucky. These findings support that highly impoverished women with low educational attainment in Appalachian counties of KY remain key interventional targets in light of lower prevalence estimates of reported mammography, increased odds of an advanced breast cancer at time of diagnosis and findings of increased breast cancer mortality in this population. Characterizing rural-urban differences in mammography utilization should be recognized as an important strategy in breast cancer prevention and control, especially considering evidence spatial clustering of late stage breast cancer has also

occurred in rural counties of western Kentucky.[61] Furthermore, our findings also suggest that particularly high-risk areas of Kentucky include the rural, isolated counties of the Appalachian region, more specifically, those with a Beale code designation ≥6. Taking into account the presence of clear dose response gradients of advanced diagnosis by socioeconomic measures, these findings reinforce the perception that socioeconomic deprivation plays an important role in health status especially since cancer stage is a strong determinant of individual patient survival. Public health planning in KY should continue to allocate preventive medical resources toward highly vulnerable women living in poverty and consider our rural and Appalachian women as priority populations for surveillance in order to improve breast cancer outcomes in Kentucky.

Limitations

The limitations of the study include the inability to approximate the true relative risk of late stage breast cancer in Kentucky females based on our case control design. Also, considering breast cancer incidence is a fairly common occurrence in the population of study, using odds ratios may potentially overestimate the true measures of association. Furthermore, making inferences about Kentucky women individually based on aggregate census data is weakened by the argument of ecological fallacy. However, given that individual socioeconomic status is not available in the health record and that one of our primary goals was to clarify the effects of residence on a woman's risk of late stage diagnosis, a multilevel analysis was supported. Other limitations include an inability to control for other potential confounders. Considering that rural and Appalachian residences are well known to be deficient in primary health care providers, not controlling for differences in the level of access to care by residence type may

potentially affect the predictive value of multilevel modeling. However, the decision was made not to include readily available county estimates of primary care providers from the Health Resources and Services Administration (HRSA) because these measures are plagued with inaccuracy since they are often temporally inconsistent due to provider outmigration and often do not account for the entirety of referring providers in the community, including physician assistants and nurse practitioners. Obesity was another potential confounder not accounted for in the analysis, since BMI is only in its infancy of collection in the health record, and estimates for county level obesity prevalence were only readily available through BRFSS, which is prone to significant sampling and non-response bias. Finally, the substantial amount of missing and unknown data inherent to the cancer registry data file may have affected the validity of certain variables being considered in the analysis, however among covariates selected in the primary analysis, we are fairly confident in the predictive ability of our final models.

VIII. Conclusions

This study demonstrated an association between residence type and odds of late stage breast cancer diagnosis among Kentucky females aged 40 years and older. Residence in either rural and Appalachian counties appears to increase a woman's risk of late stage breast cancer compared to urban and non-Appalachian counties respectively, although the effects of locality were largely explained by the influence of high levels of poverty, economic distress and lower levels of educational attainment present in these geographically distinct areas of KY. Based on our analysis, women living in the rural counties of Appalachia and highly socioeconomically distressed counties of rural KY appear to be at greatest risk of an advanced diagnosis. Implications of these findings

include the need for developing health policy that simultaneously integrates initiatives fostering social and economic growth in problematic counties. Likewise, interventions enhancing social capital and class mobility among impoverished and geographically isolated communities will likely have the greatest impact on breast cancer outcomes long term.

IX. Acknowledgements

I would like to extend my deepest thanks to Dean Stephen Wyatt, Dr. Douglas Scutchfield and Dr. Tom Tucker for allowing me the opportunity to submit this research paper as a portion of the requirements to complete a degree of Master of Public Health with a concentration in Epidemiology. I would also like to express my sincerest appreciation to Dominique Zephyr in the Applied Statistics Lab for his technical support merging datasets and for his assistance with the R software programming language. A special thanks to Jaclyn K. Nee, MPH at the University of Kentucky Markey Cancer Control Program Kentucky Cancer Registry as well as Sarah Ehresman at the KY State Data Center for providing the data sets and Bin Huang, DrPH for technical guidance with KCR data coding. Finally, I would also like to thank my friend and colleague, Stephen Jurich at the University of Mississippi College of Finance for his help with SAS programming language and obtaining SPSS for the analysis.

Dean Stephen Wyatt, University of Kentucky College of Public Health

Dr. Douglas F. Scutchfield, University of Kentucky College of Public Health

Dr. Tom Tucker, University of Kentucky College of Public Health

Dominique Zephyr, University of Kentucky College of Public Health

Bin Huang, The Kentucky Cancer Registry

Jaclyn K. Nee, The Kentucky Cancer Registry

Sarah Ehresman, Kentucky State Data Center

Stephen Jurich, University of Mississippi College of Finance

X. Tables and Figures

Table 1.1: Frequency of Characteristics by Type of Residence Among KY Females, Age 40 Years and Older with Known Stage at Time of Breast Cancer Diagnosis 2001-2011, using the AJCC TNM Staging Criteria.

Characteristic		All N=23,100 (100%)	Rural N=8,233 (35.6%)	Urban N=14,867 (64.4%)	Appalachian N=5,610 (24.3%)	Non-Appalachian N=17,490 (75.7%)
Mean Age (SD)		61.4 (±12.5)	61.9 (±12.4)	61.1 (±12.5)	61.5 (±12.2)	61.4 (±12.6)
Age at Diagnosis	40-49	4,558 (19.7%)	1,512 (18.4%)	3,046 (20.7%)	1,056 (18.8%)	3,502 (20%)
	50-64	9,625 (41.7%)	3,389 (41.1%)	6,236 (41.9%)	2,359 (42%)	7,266 (41.5%)
Missing (N=0)	65-75	4,858 (21%)	1,866 (22.7%)	2,992 (20.1%)	1,266 (22.6%)	3,592 (20.5%)
	75 and older	4,059 (17.6%)	1,466 (17.8%)	2,593 (17.4%)	929 (16.6%)	3,130 (17.9%)
Age at Diagnosis (By menopausal status)	Premenopausal (40-44)	1,886 (8.2%)	636 (7.7%)	1,250 (8.4%)	430 (7.7%)	1,456 (8.3%)
	Perimenopausal (45-54)	5,802 (25.1%)	1,890 (23%)	3,912 (26.3%)	1,347 (24%)	4,455 (25.5%)
Missing (N=0)	Postmenopausal (≥55)	15,412 (66.7%)	5,707 (69.3%)	9,705 (65.3%)	3,833 (68.3%)	11,579 (66.2%)
Race	White	21,392 (92.6%)	8,011 (97.3%)	13,381 (90%)	5,515 (98.3%)	15,877 (90.8%)
	Black	1,536 (6.6%)	182 (2.2%)	1,354 (9.1%)	68 (1.2%)	1,468 (8.4%)
Missing (N=0)	Other Minority	101 (0.4%)	7 (0.08%)	94 (0.6%)	6 (0.1%)	95 (0.5%)
	Unknown	71 (0.3%)	33 (0.4%)	38 (0.3%)	21 (0.4%)	50 (0.3%)
Type of Insurance	Private Insurance (including Military)	11,673 (50.5%)	3,464 (42.1%)	8,209 (55.2%)	2,283 (40.7%)	9,390 (53.7%)
Missing (N=0)	Uninsured	689 (3%)	341 (4.1%)	348 (2.3%)	240 (4.3%)	449 (2.6%)
	Medicare	9,327 (40.4%)	3,657 (44.4%)	5,670 (38.1%)	2,466 (44%)	6,861 (39.2%)
	Medicaid	1,411 (6.1%)	771 (9.4%)	640 (4.3%)	621 (11.1%)	790 (4.5%)
Marital Status	Ever Married	20,568 (89%)	7,456 (90.6%)	13,112 (88.2%)	5,010 (89.3%)	15,558 (89%)
	Never Married	2,026 (8.8%)	523 (6.4%)	1,503 (10.1%)	349 (6.2%)	1,677 (9.6%)
Missing (N=10)	Unknown	496 (2.1%)	250 (3%)	246 (1.7%)	251 (4.5%)	245 (1.4%)
Smoking Status	Never Smoked	11,422 (49.9%)	4,113 (50.4%)	7,309 (49.6%)	2719 (49.1%)	8703 (50.1%)
	Ever Smoked	7,457 (32.6%)	2,449 (30.1%)	5,008 (34%)	1613 (29.1%)	5826 (33.5%)

Missing (N=203)	Unknown	4,018 (17.5%)	1,592 (19.5%)	2,426 (16.4%)	1186 (21.4%)	2832(16.3%)
Pack History	Never Smoked	8,625 (37.3%)	2,822 (34.3%)	5,803 (39.1%)	1,773 (31.6%)	6,852 (39.3%)
	0-20 Pack Years	974 (4.2%)	313 (3.8%)	661 (4.5%)	218 (3.9%)	756 (4.3%)
	20+ Pack Years	2,928 (12.7%)	1,059 (12.9%)	1,869 (12.6%)	746 (13.3%)	2,182 (12.5%)
Missing (N=50)	Unknown	10,523 (45.6%)	4,031 (92.9%)	6,492 (43.8%)	2,866 (51.2%)	7,657 (43.9%)
Family History	Yes	6751 (29.2%)	2,199 (26.7%)	4,552 (30.6%)	1,549 (27.6%)	5,202(29.7%)
	No	10,977 (47.5%)	3,543 (43%)	7,434 (50%)	2,316 (41.3%)	8,661 (49.5%)
	Unknown	5,371 (23.3%)	2,491 (30.3%)	2,880 (19.4%)	1,745 (31.1%)	3,626 (20.7%)
Parity (Number of Live Births) Missing (N=1)	Nulliparous	1,403 (6.1%)	373 (4.5%)	1,030 (6.9%)	257 (4.6%)	1,146 (6.6%)
	Low (1-2)	6,109 (26.4%)	1,938 (23.5%)	4,171 (28.1%)	1,364 (24.3%)	4,745 (27.1%)
	High (3 or more)	4,072 (17.6%)	1,292 (15.7%)	2,780 (18.7%)	926 (16.5%)	3,146 (18%)
	Unknown	11,515 (49.8%)	4,630 (56.2%)	6,885 (46.3%)	3,063 (54.6%)	8,452 (48.3%)
Stage at Diagnosis	Early (0-1)	13,099 (56.7%)	4,352 (52.9%)	8,747 (58.8%)	2,867 (51.1%)	10,232 (58.5%)
	Late (2-4)	10,001 (43.3%)	3,881 (47.1%)	6,120 (41.2%)	2,743 (49.9%)	7258 (41.5%)

Note: Percentages represent the proportion of women in each column with this characteristic. Denominator is column total

Table 1.2: Proportion of Selected Socioeconomic Characteristics† for County of Residence in Kentucky Females Age 40 Years and Older with Known Stage of Breast Cancer, 2001-2011.

Socioeconomic Characteristic		**All** N= 23,100 (%)	**Rural** N=8,233 (%)	**Urban** N=14,867 (%)	**Appalachian** N=5,610 (%)	**Non-Appalachian** N=17,490 (%)
Percent Living Below the Poverty Line	Very Low (<7%)	1,011 (4.4%)	0 (0%)	1,011 (6.8%)	0 (0%)	1,011 (5.8%)
	Low (7% - 13%)	14,545 (63%)	1,730 (21%)	12,815 (86%)	787 (14%)	13,758 (78.7%)
	High (13% - 22%)	5,850 (25.3%)	4,809 (58.4%)	1,041 (7%)	3,169 (56%)	2,681 (15%)
	Very High (>22%)	1,694 (7.3%)	1,694 (20.5%)	0 (0%)	1,694 (30%)	40 (0.2%)
Educational Attainment: Percent HS Graduate or Higher	Very High (≥82.4%)	14,912 (64.6)	1,257 (15.3%)	13,655 (92%)	315 (5.6%)	14,597 (83%)
	High (73.7% - 82.3%)	4,192 (18.1%)	3,020 (36.7%)	1,172 (7.9%)	1,548 (28%)	2,644 (15%)
	Low (64.9% - 73.6%)	3,511 (15.2%)	3,471 (42.1%)	40 (0.3%)	3,262 (58%)	249 (1.4%)
	Very Low (≤64.8%)	485 (2.1%)	485 (5.9%)	0 (0%)	485 (8.6%)	0 (0%)

Median Household Income	High (≥$50,962)	3,580 (15.5%)	156 (1.9%)	3,424 (23%)	0 (0%)	3,580 (20%)
	Low ($35,153-$50,961)	15,185 (65.7%)	3,742 (45%)	11,443 (77%)	1,727 (31%)	13,458 (77%)
	Very Low (≤ $35,152)	4,335 (18.8%)	4,335 (53%)	0 (0%)	3,883 (69%)	452 (2.6%)
Per Capita Income	Very High (≥ $27,996)	2,834 (12.3%)	0 (0%)	2,834 (19%)	0 (0%)	2,834 (16%)
	High ($22,625 - $27,995)	10,388 (45%)	771 (9.4%)	9,617(65%)	546 (9.7%)	9,842 (56%)
	Low ($17,254 - $22,624)	6,783 (29.4%)	4,367 (53%)	2,416 (16%)	2,026 (36%)	4,757 (27%)
	Very Low (≤ $17,253)	3,095 (13.4%)	3,095 (37.6%)	0 (0%)	3,038 (54%)	57 (0.3%)
Percent Unemployment	Very Low (≤10.5%)	20,149 (87.2%)	5,717 (69%)	14,432 (97%)	3,653 (65%)	16,496 (94.3%)
	High (10.6% - 15.7%)	2,525 (10.9%)	2,090 (25.4%)	435 (3%)	1,531 (27%)	994 (5.7%)
	Very High (15.8% - 26.1%)	426 (1.8%)	426 (5.2%)	0 (0%)	426 (7.6%)	0 (0%)

Note: Percentages represent the proportion of women in each column with this characteristic. Denominator is column total.
[†]Represents county level estimates from the 2007-2011 American Community Survey.

Table 1.3: Association of Risk Factors with Stage at Diagnosis using Chi Square Test for Independence

Individual Risk Factor		Chi Square Statistic ($\chi 2$)	P-value
Age at Diagnosis by Screening Interval		73.316	0.001
Age at Diagnosis by Menopausal Status		25.097	0.001
Race		30.171	0.001
Family History		20.404	0.001
Marital Status		23.179	0.001
Parity		6.466	0.091
Smoking Status (Ever vs. Never Smoked)		19.928	0.001
Smoking History (Number of Pack Years)		16.516	0.001
Insurance Type		192.432	0.001
Type of Residence	Appalachian	94.662	0.001
	Rural	77.041	0.001
Census-tract Risk Factor		Chi Square Statistic ($\chi 2$)	P-value
Percent HS Graduate or Higher		129.271	0.001
Percent Living Below the Poverty Line		141.018	0.001
Median Household Income		110.445	0.001
Per Capita Income		129.530	0.001

Percent Unemployment	60.215	0.001

Table 1.4: Pearson Correlation of Risk Factors with Stage at Diagnosis using TNM criteria

Individual Risk Factors		Pearson Correlation Coefficient (ρ)	P-value (2-tailed)
Age at Diagnosis by Screening Interval		-0.002	0.722
Age at Diagnosis by Menopausal Status		-0.032	0.001
Race		0.017	0.012
Marital Status		0.029	0.001
Smoking Status		0.014	0.037
Smoking History in Pack Years		0.018	0.007
Family History of this Cancer		-0.001	0.846
Parity		0.007	0.319
Type of Residence	Appalachian	0.064	0.001
	Rural	0.058	0.001
	Interaction term	0.068	0.001
Census-tract Risk Factors		Pearson Correlation Coefficient (ρ)	P-value (2-tailed)
Percent HS Graduate or Higher		0.073	0.001
Percent Living Below the Poverty Line		0.077	0.001
Median Household Income		0.059	0.001
Per Capita Income		0.072	0.001
Percent Unemployment		0.051	0.001

Table 1.5: Hierarchal Multivariate Logistic Regression Model of Risk Factors Associated with Late Stage Breast Cancer Diagnosis in Kentucky Females 40 years and older, 2001-2011 (Based on TMN Staging Criteria)

Characteristics		Adjusted Odds Ratio	95% Confidence Interval	P-value
Age	Premenopausal (40-44)	Reference	----	--------
	Perimenopausal (45-54)	0.86	0.771 – 0.952	0.004
	Postmenopausal (55+)	0.77	0.698-0.858	0.001
Race	White	Reference	----	-------
	African American	1.31	1.174 – 1.455	0.001
	Other Minority	0.92	0.611 – 1.372	0.690
	Unknown	0.53	0.317 – 0.886	0.018
Marital Status	Ever Married	Reference	----	----
	Never Married	1.15	1.047 – 1.263	0.003
	Unknown	1.07	0.899 – 1.292	0.411
Insurance Type	Insured*	Reference	----	--------
	Uninsured	1.93	1.646 – 2.256	0.001
	Medicare	1.15	1.077- 1.225	0.001
	Medicaid	1.65	1.476 – 1.853	0.001
Type of Residence	Non-Appalachian	Reference	----	----
	Appalachian	1.22	1.127 – 1.309	0.001
	Urban	Reference	----	----
	Rural	1.13	1.059 – 1.212	0.001
Interaction Term†	Rural*Appalachian	1.167	----	0.001

*Insured category includes private insurance payers and Military payers. †Interaction term Rural*Appalachian was entered separately into model to avoid multicollinearity.

Table 1.5: Crude Associations of County-Based Socioeconomic Characteristics with Late Stage Breast Cancer Diagnosis in Kentucky Females Age 40 years and older, 2001-2011 (Based on TNM Staging Criteria)

Contextual Risk Factor		Crude Odds Ratio	P-value
Percent Living Below Poverty Line	Very Low (<7%)	Reference	----
	Low (7% - 13%)	1.201	0.006
	High (13% - 22%)	1.449	0.001
	Very High (>22%)	2.034	0.001
Percent HS Graduate or Higher	Very High (≥82.4%)	Reference	----
	High (73.7% - 82.3%)	1.116	0.002
	Low (64.9% - 73.6%)	1.420	0.001
	Very Low (≤64.8%)	1.937	0.001
Median Household Income	High (≥$50,962)	Reference	----
	Low ($35,153-$50,961)	1.044	0.253
	Very Low (≤ $35,152)	1.473	0.001
Per Capita Income	Very High (≥ $27,996)	Reference	----
	High ($22,625 - $27,995)	1.216	0.001
	Low ($17,254 - $22,624)	1.332	0.001
	Very Low (≤ $17,253)	1.762	0.001
Unemployment	Very Low (≤10.5%)	Reference	----
	High (10.6% - 15.7%)	1.294	0.001
	Very High (15.8% - 26.1%)	1.642	0.001

Table 1.6: Hierarchal Multivariate Mixed Model Associations of County Socioeconomic Indicators with Late Stage Breast Cancer Diagnosis Among Kentucky Females 40 Years and Older When Entered Sequentially (Staging based on TNM criteria)

County Socioeconomic Indicator	*Adjusted Odds Ratio	95% Confidence Interval (CI)
Percent Living Below Poverty Line	1.163	1.077 – 1.256
Percent HS Graduate or Higher	1.139	1.047 – 1.238

Average County Income	Median Household Income	1.104	1.027 – 1.187
	Per Capita Income	1.125	1.015 – 1.246
Percent Unemployment		1.101	1.005 – 1.207

**Adjusted for Age, Race, Type of Insurance and Residence.*

Figure 1.1: Measure of Association of Appalachian Residence with Late Stage Breast Cancer Controlling for the Effects of County Socioeconomic Indicators Sequentially in the Mixed Model (Based on TNM Staging Criteria)

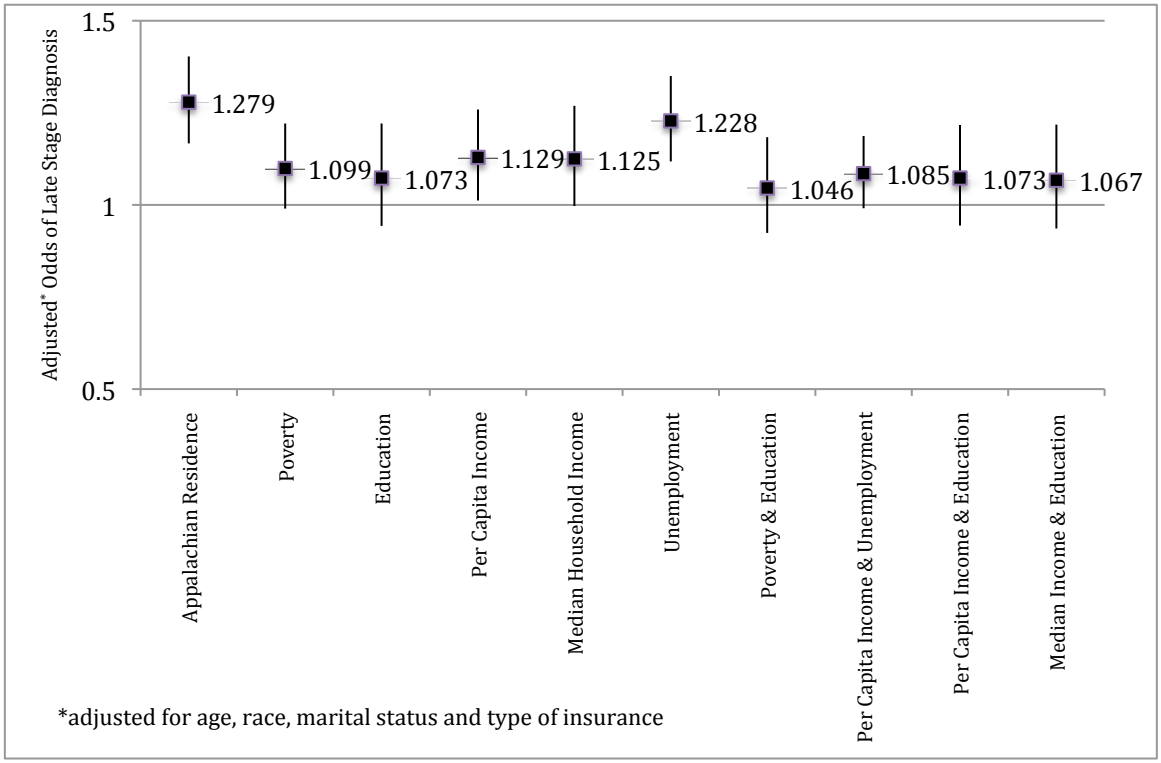

Figure 1.1: Measure of Association of Rural Residence with Late Stage Breast Cancer Controlling for the Effects of County Socioeconomic Indicators Sequentially in the Mixed Model (Based on TNM Staging Criteria)

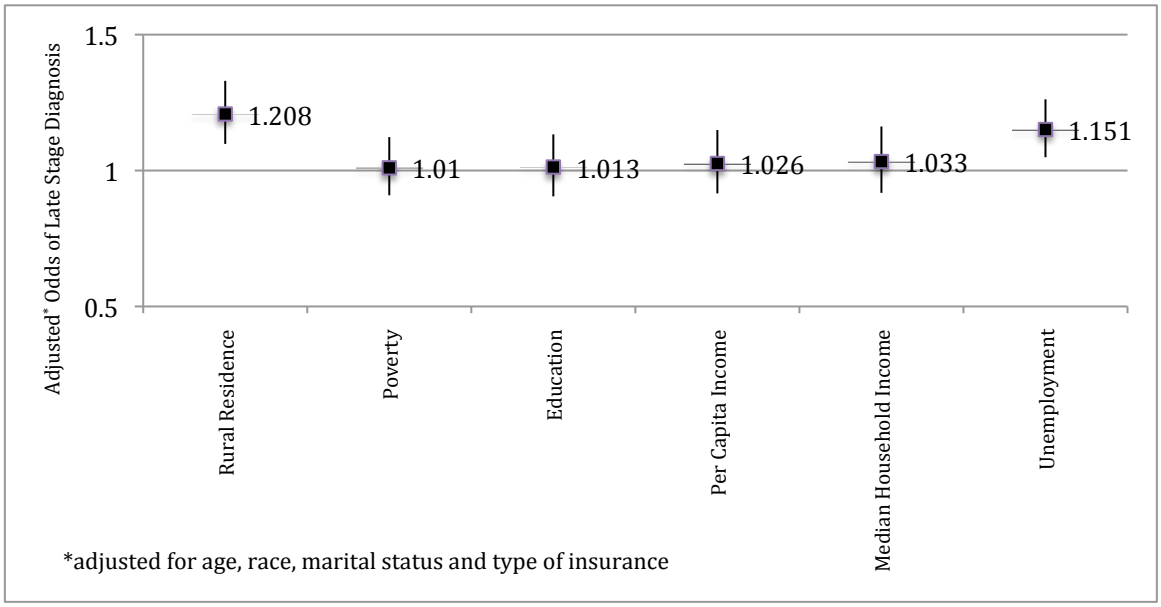

Table 2.1 Frequency of Characteristics by Residence Type among KY females, Age 40 Years and Older with Known Stage at time of Breast Cancer Diagnosis 2001-2011, using SEER 2000 Summary Staging Criteria.

Characteristic		**Overall** N= 23,796 (100%)	**Rural** N=8,592 (36.1%)	**Urban** N=15,204 (63.9%)	**Appalachian** N=5,851 (24.6%)	**Non-Appalachian** N=17,945 (75.4%)
Mean Age (SD)		61.5 (±12.5)	62.0 (±12.4)	61.2 (±12.5)	61.6 (±12.5)	61.5 (±12.6)
Age at Diagnosis Missing (N=0)	40-49	4,679 (19.7%)	1,577 (18.3%)	3,102 (20.4%)	1,104 (18.9%)	3,575 (20%)
	50-64	9,873 (41.5%)	3,524 (41%)	6,349 (41.7%)	2,451 (41.9%)	7,422 (41.3%)
	65-75	4,998 (21%)	1,939 (22.6%)	3,059 (20.1%)	1,310 (22.3%)	3,688 (20.5%)
	75 and older	4,246 (17.8%)	2,553 (29.7%)	2,694 (17.7%)	986 (16.9%)	3,260 (18.2%)
Age at Diagnosis by Menopausal Status Missing (N=0)	Premenopausal (40-44)	1,928 (8.1%)	662 (7.7%)	1,266 (8.3%)	450 (7.7%)	1,478 (8.2%)
	Perimenopausal (45-54)	5,958 (25%)	1,977 (23%)	3,981 (26.2%)	1,400 (23.9%)	4,558 (25.4%)
	Post-menopausal (55)	15,910 (66.9%)	5,953 (69.3%)	9,957 (65.5%)	4,001 (68.4%)	11,909 (66.4%)
Race	White	22,039 (92.6%)	8,359 (97.3%)	13,680 (90%)	5,750 (98.3%)	16,289 (90.8%)
	Black	1,569 (6.6%)	187 (2.2%)	1,382 (9.1%)	70 (1.2%)	1,499 (8.3%)

Missing (N=0)	Other Minority	106 (0.4%)	7 (0.08%)	99 (0.6%)	6 (0.1%)	100 (0.5%)
	Unknown	82 (0.3%)	39 (0.4%)	43 (0.2%)	25 (0.4%)	57 (0.3%)
Type of Insurance	Private Insurance (includes Military)	11,963 (50.3%)	3,608 (42%)	8,355 (55%)	2,370 (40.5%)	9,593 (53.4%)
	Uninsured	714 (3%)	360 (4.2%)	354 (2.3%)	255 (4.4%)	459 (2.5%)
Missing (N=0)	Medicare	9,662 (40.6%)	3,822 (44.4%)	5,840 (38.4%)	2,577 (44%)	7,085 (39.5%)
	Medicaid	1,457 (6.1%)	802 (9.3%)	655 (4.3%)	649 (11.1%)	808 (4.5%)
Marital Status	Ever Married	21,159 (88.9%)	7,764 (90.5%)	13,395 (88.2%)	5,217 (89.1%)	15,492 (86.3%)
	Never Married	2,080 (8.7%)	544 (6.3%)	1,536 (10.1%)	361 (6.2%)	1,719 (9.5%)
Missing (N=11)	Unknown	546 (2.3%)	279 (3.2%)	267 (1.7%)	273 (4.7%)	273 (1.5%)
Smoking Status	Never Smoked	11,691 (49.1%)	4,251 (50.7%)	7,440 (50%)	2,806 (50%)	8,885 (50.1%)
Missing	Ever Smoked	7,610 (32%)	2,539 (30.3%)	5,071 (33.8%)	1,690 (30%)	5,920 (33.4%)
(N=215)	Unknown	4,280 (18%)	1,719 (20.5%)	2,561 (17.1%)	1,276 (22.6%)	3,004 (16.9%)
Pack History	Never Smoked	8,829 (37.1%)	2,924 (34.2%)	5,905 (39%)	1,833 (31.6%)	6,996 (39.1%)
	0-20 Pack Years	979 (4.1%)	316 (3.7%)	663 (4.4%)	220 (3.8%)	759 (4.2%)
	20+ Pack Years	2,988 (12.6%)	1,094 (12.8%)	1,894 (12.5%)	770 (13.3%)	2,218 (12.4%)
Missing (N=51)	Unknown	10,949 (46%)	4,249 (49.7%)	6,700 (44.2%)	3,021 (52.1%)	7,928 (44.3%)
Family History	No	11,247 (47.3%)	3,673 (42.7%)	7,574 (49.8%)	2,396 (41%)	8,851 (49.3%)
of this Cancer	Yes	6,869 (28.9%)	2,258 (26.3%)	4,611 (30.3%)	1,590 (27.2%)	5,279 (29.4%)
Missing (N=1)	Unknown	5,679 (23.9%)	2,661 (31%)	3,018 (19.8%)	1,865 (31.9%)	3,814 (21.2%)
Parity	High (3 or more)	4,153 (17.5%)	1,339 (15.5%)	2,814 (18.5%)	953 (6.3%)	3,200 (17.8%)
(Number of Live	Low (1-2)	6,222 (26.1%)	1,992 (23.2%)	4,230 (27.8%)	1,397 (23.9%)	4,825 (26.9%)
Births)	Nulliparous	1,438 (6%)	389 (4.5%)	1,049 (6.9%)	268 (4.6%)	1,170 (6.5%)
Missing (N=1)	Unknown	11,982 (50.4%)	4,872 (56.7%)	7,110 (46.8%)	3,233 (55.3%)	8,749 (48.7%)
Stage at	Early (0-1)	16,503 (69.4%)	5,760 (67%)	10,743 (70.7%)	3,844 (65.7%)	12,659 (70.5%)
Diagnosis	Late (2-7)	7,293 (30.6%)	2,832 (33%)	4,461 (29.3%)	2,007 (34.3%)	5,286 (29.5%)

Note: Percentages represent the proportion of women in each column with this characteristic. Denominator is column total

Table 2.2: Association of Risk Factors with Stage at Diagnosis using Chi Square Test for Independence

Individual Risk Factor		Chi Square Statistic ($\chi 2$)	P-value
Age at Diagnosis by Screening Interval		35.788	0.001
Age at Diagnosis by Menopausal Status		19.842	0.001
Race		13.887	0.003
Family History		9.500	0.009
Marital Status		12.638	0.002
Parity		11.377	0.010
Smoking Status (Ever vs. Never Smoked)		21.867	0.001
Smoking History (Number of Pack Years)		18.284	0.001
Insurance Type		159.356	0.001
Type of Residence	Appalachian	48.733	0.001
	Rural	33.844	0.001
Census-tract Risk Factor		Chi Square Statistic ($\chi 2$)	P-value
Percent HS Graduate or Higher		78.989	0.001
Percent Living Below the Poverty Line		89.114	0.001
Median Household Income		73.530	0.001
Per Capita Income		82.081	0.001
Percent Unemployment		36.020	0.001

Table 2.3: Pearson Correlation of Risk Factors with Stage of Diagnosis based on SEER Summary 2000 Manual

Individual Risk Factors	Pearson Correlation Coefficient (ρ)	P-value (2-tailed)
Age at Diagnosis by Screening Interval	-0.002	0.722
Age at Diagnosis by Menopausal Status	-0.027	0.001
Race	0.013	0.050
Marital Status	0.020	0.002
Smoking Status	0.007	0.318
Smoking History in Pack Years	0.014	0.035
Family History of this Cancer	-0.003	0.622

Parity		-0.001	0.910
Type of Residence	Appalachian	0.045	0.001
	Rural	0.038	0.001
	Interaction term	0.068	0.001
Census-tract Risk Factors		Pearson Correlation Coefficient (ρ)	P-value (2-tailed)
Percent HS Graduate or Higher		0.052	0.001
Percent Living Below the Poverty Line		0.058	0.001
Median Household Income		0.045	0.001
Per Capita Income		0.055	0.001
Percent Unemployment		0.038	0.001

Table 2.4: Hierarchal Multivariate Logistic Regression Model of Risk Factors Associated with Late Stage Breast Cancer Diagnosis in Kentucky Females 40 years and older, 2001-2011 (Based on SEER Staging Criteria)

Characteristics		Adjusted Odds Ratio	95% Confidence Interval	P-value
Age	Premenopausal (40-44)	Reference	----	0.001
	Perimenopausal (45-54)	0.849	0.761 – 0.948	0.003
	Postmenopausal (55+)	0.794	0.713 – 0.884	0.001
Race	White	Reference	----	0.005
	African American	1.208	1.080 – 1.351	0.001
	Other Minority	0.944	0.616 – 1.446	0.790
	Unknown	0.722	0.437 – 1.192	0.203
Marital Status	Ever Married	Reference	----	0.122
	Never Married	1.108	1.004 – 1.222	0.041
	Unknown	1.023	0.851 – 1.229	0.808
Smoking History by Number of Pack Years	Never Smoked	Reference	----	0.048
	<20 pack years	1.054	0.913 – 1.217	0.476
	≥20 pack years	1.137	1.039 – 1.243	0.005
	Unknown	1.029	0.967 – 1.095	0.364

Insurance Type	Insured*	Reference	----	0.001
	Uninsured	1.881	1.613 – 2.195	0.001
	Medicare	1.127	1.052 – 1.206	0.001
	Medicaid	1.557	1.388 – 1.746	0.001
Type of Residence	Non-Appalachian	Reference	----	----
	Appalachian	1.162	1.074 – 1.257	0.001
	Urban	Reference	----	----
	Rural	1.069	0.995 – 1.148	0.067
Interaction Term	Rural*Appalachian†	1.249	1.166 – 1.337	0.001

*Insured category includes private insurance payers and Military payers. †Rural*Appalachian interaction term entered separately into model to avoid multicollinearity.

Table 2.5: Crude Associations of County-Based Socioeconomic Characteristics with Risk of Late Stage Breast Cancer Diagnosis in Kentucky Females Age 40 years and older, 2001-2011 (Based on SEER Staging Criteria)

Contextual Risk Factor		Crude Odds Ratio	P-value
Percent Living Below Poverty Line	Very Low (<7%)	Reference	----
	Low (7% - 13%)	1.037	0.613
	High (13% - 22%)	1.225	0.006
	Very High (>22%)	1.616	0.001
Percent HS Graduate or Higher	Very High (≥82.4%)	Reference	----
	High (73.7% - 82.3%)	1.021	0.581
	Low (64.9% - 73.6%)	1.319	0.001
	Very Low (≤64.8%)	1.682	0.001
Median Household Income	High (≥$50,962)	Reference	----
	Low ($35,153-$50,961)	0.824	1.009
	Very Low (≤ $35,152)	1.356	0.001
Per Capita Income	Very High (≥ $27,996)	Reference	----
	High ($22,625 - $27,995)	1.192	0.001
	Low ($17,254 - $22,624)	1.250	0.001
	Very Low (≤ $17,253)	1.612	0.001

Unemployment	Very Low (≤10.5%)	Reference	----
	High (10.6% - 15.7%)	1.186	0.001
	Very High (15.8% - 26.1%)	1.588	0.001

Table 2.6: Hierarchal Multivariate Mixed Model Associations of County Socioeconomic Indicators with Late Stage Breast Cancer Diagnosis Among Kentucky Females 40 Years and Older (Based on SEER Staging Criteria)

Socioeconomic Indicators		*Adjusted Odds Ratio	95% Confidence Interval (CI)
Percent Living Below Poverty Line		1.166	1.081 – 1.258
Percent HS Graduate or Higher		1.109	1.023 – 1.203
Average County Income	Median Household Income	1.145	1.033 – 1.269
	Per Capita Income	1.109	1.032 – 1.192
Percent Unemployment		1.081	0.987 – 1.184

**Adjusted for Age, Race, Smoking History, Type of Insurance and Residence.*

XI. References

1. U.S. Cancer Statistics Working Group. *United States Cancer Statistics: 1999–2010 Incidence and Mortality Web-based Report*. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2013. Available at: www.cdc.gov/uscs.
2. Friedell GH, Linville LH, Sorrell CL, Huang B. Kentucky breast cancer report card. J Ky Med Assoc 2003; 101(10):449 –54.
3. Halverson JA. An analysis of disparities in health status and access to heath care in the Appalachian region. November 2004.
4. Tucker, TC. The Cancer Burden in Kentucky. Presentation to the KCC Steering Committee- Louisville, KY July 30th, 2012.
5. Kentucky Women's Cancer Screening Program Annual Report 2012. http://chfs.ky.gov/NR/rdonlyres/7F10ADC3-4DC6-4A66-A008-44BCAAAEA727/0/KWCSPAnnualReport20112.pdf
6. Hall HI, Uhler RJ, Coughlin SS, Miller DS. Breast and cervical cancer screening among Appalachian women. Cancer Epidemiol, Biomarkers Prev. 2002;11:137-142.
7. Appalachian Regional Commission. Economic Overview of Appalachia, 2011. www.arc.gov/images/appregion/Sept2011/EconomicOverviewSept2011.pdf
8. World Health Organization. World Cancer Report 2008. Cedex: International Agency for Research on Cancer; 2008.
9. National Cancer Institute. Cancer Staging. National Institutes of Health.

10. Huang, B., et al., Does Distance Matter? Distance to Mammography Facilities and Stage at Diagnosis of Breast Cancer in Kentucky. Journal of Rural Health, 2009. 25(4): p. 366-371.
11. Wells, B. L. Stage at diagnosis in breast cancer: race and socioeconomic factors. Am J. Public Health, 1992. 82: p. 1383-1385
12. Royse, D. and M. Dignan, Appalachian knowledge of cancer and screening intentions. J Cancer Educ, 2009. 24(4): p. 357-62.
13. Hahn, K.M.E., et al., Factors associated with advanced disease stage at diagnosis in a population-based study of patients with newly diagnosed breast cancer. American Journal of Epidemiology, 2007. 166(9): p. 1035-1044.
14. Schoenberg, N. E. Appalachian Women's Perceptions of Their Community Health Threats. The Journal of Rural Health. 24(1): p. 36-45
15. Schoenberg, N. E. Patterns and Determinants of breast and cervical cancer non-screening among Appalachian women. Women Health, 2013. 53(6): p. 1-8.
16. Fleming, S. T. Diabetes and Cancer Screening Rates among Appalachian and Non-Appalachian Residents of Kentucky. Journal of the American Board of Family Practice, 2011. 24(6): p. 682-686.
17. Kim, S. et al. Sociodemographic Characteristics, Distance to Clinic and Breast Cancer Screening Results. Journal of Health Disparities, 2013. 6(1): p. 70-84.
18. Tessaro, I. et al. Knowledge, barriers, and predictors of colorectal cancer screening in an Appalachian church population. Preventing Chronic Disease, 2006. 3(4): A123.

19. Vanderpool, R. C. Cancer risk perceptions, beliefs and physician avoidance in Appalachia: results from the 2008 HINTS Survey. Journal of Health Communication, 2010. 15(Suppl 3): p. 78 - 91.
20. Kelly, K.M., et al., Information seeking and intentions to have genetic testing for hereditary cancers in rural and appalachian Kentuckians. Journal of Rural Health, 2007. 23(2): p. 166-172.
21. Coughlin, S.S.L., Steven; Richards, Thomas; Sabatino, Susan A., Contextual Analysis of breast and cervical cancer screening and factors associated with health care access among United States women, 2002. Social Science & Medicine, 2008. 66: p. 260-275.
22. Lyttle, N.L. and K. Stadelman, Assessing awareness and knowledge of breast and cervical cancer among Appalachian women. Prev Chronic Dis, 2006. 3(4): p. A125.
23. Celaya, M.O., et al., Breast cancer stage at diagnosis and geographic access to mammography screening (New Hampshire, 1998-2004). Rural Remote Health, 2010. 10(2): p. 1361.
24. Schroen, A. T. and Lohr, M. E. Travel distance to mammography and the early detection of breast cancer. The Breast Journal, 2009. 15(2): p. 216-217
25. Lantz, P.M.e.a., The Influence of Race, Ethnicity, and Individual Socioeconomic Factors on Breast Cacer Stage at Diagnosis. American Journal of public Health, 2006. 96: p. 2173-2178.
26. Lannin, D. R. et al. Influence of Socioeconomic and Cultural Factors on Racial Differences in Late Stage Presentation of Breast Cancer. Journal of the American Medical Association (JAMA), 1998. 279(22): p. 1801-1807.

27. Halpern, M.T., et al., Insurance status and stage of cancer at diagnosis among women with breast cancer. Cancer, 2007. 110(2): p. 403-411.
28. Vona-Davis, L. and D.P. Rose, The Influence of Socioeconomic Disparities on Breast Cancer Tumor Biology and Prognosis: A Review. Journal of Womens Health, 2009. 18(6): p. 883-893.
29. Nguyen, S. et al. Disparities in breast cancer stage at diagnosis in urban and rural adult women: a systematic review and meta-analysis. Annals of Epidemiology, 2013. 24 (2014): p. 228-235.
30. Amey C. H., et al. The role of race and residence in determing stage at diagnosis of breast cancer. Journal of Rural Health, 1997. 13(2): p. 99-108.
31. Campbell, N. C., et al. Rural and urban differences in stage at diagnosis of colorectal and lung cancer. Cancer, 2001. 84(7): p. 910-914.
32. Smigal C., et al. Trends in breast cancer by race and ethnicity: update 2006. CA: A Cancer Journal for Clinicians, 2006. 56 (3): p. 168-183.
33. Howlander, N. et al. SEER Cancer Statistics Review, 1975-2008. Bethesda, MD: National Cancer Institute, 2010.
34. Reynolds, P. et al. Regional variations in breast cancer incidence among California women, 1988-1997. Cancer Causes and Control, 2005. 16: 139-150.
35. Hausauer et al. Recent trends in breast cancer incidence in US white women by county level, urban/rural and poverty status. BMC Medicine, 2009. 7(31).
36. Kuzmiak, C.M., et al., Insurance status and the severity of breast cancer at the time of diagnosis. Acad Radiol, 2008. 15(10): p. 1255-8.

37. Plescia, M. The National Prevention Strategy and Breast Cancer Screening: Scientific Evidence for Public Health Action. American Journal of Public Health, 2013. 103(1): p. 1545-1548.
38. US Preventive Task Force. Screening for breast cancer: US Preventive Services Task Force recommendation statement. Ann Intern Med, 2009. 151(10): p. 716 – 726.
39. McDavid, K., et al., Cancer survival in Kentucky and health insurance coverage. Arch Intern Med, 2003. 163(18): p. 2135-44.
40. Coburn, N., et al., Treatment Variation by Insurance Status for Breast Cancer Patients. The Breast Journal, 2008. 14(2): p. 128-134.
41. Costarellia, V., et al. Breast Cancer risk in women: the protective role of pregnancy. Nursing Standard, 2010. 24(18): p. 35-40.
42. Greenlee, R. T. and Howe, H. L. County-level poverty and distant stage cancer in the United States. Cancer Causes & Control, 2009. 20(6): p. 989-1000.
43. Henry, K. A. The joint effects of census tract poverty and geographic access on late-stage breast cancer diagnosis in 10 US States. Health & Place, 2013. 21: p. 110-121
44. Saquib N. et al. Mortality risk in former smokers with breast cancer: Pack-years vs. smoking status. International Journal of Cancer, 2013. 133: 2493-2497.
45. Lengerich, E. J. et al. Cancer Incidence in Kentucky, Pennsylvania, and West Virginia: disparities in Appalachia. Journal of Rural Health, 2005. 21(1): p. 39-47.
46. Lengerich, E. J. et al. The Appalachian Cancer Network: Cancer Control Research Among a Rural, Medically Underserved Population. Journal of Rural Health, 2004. 20(2): p. 181-187.
47. Wingo, P. A. et al. Cancer in Appalachia, 2001-2003. Cancer, 2008. 112(1): p. 181-192.

48. Edwards, K. K. et al. Impact of geographic barriers on the utilization of mammograms by older rural women. Journal of the American Geriatrics Society, 2002. 50(1): p. 62-68.

49. Calle, E. E. et al. Demographic Predictors of mammography and Pap smear screening in US women. American Journal of Public Health, 1993. 83(1): p. 53 – 60.

50. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2009;4: CD001877.

51. Osborne, C. et al. The influence of marital status on the stage at diagnosis, treatment, and survival of older women with breast cancer. Breast Cancer Research and Treatment, 2005. 93: p. 41-47.

52. MacKinnon, J. A. et al. Detecting an association between socioeconomic status and late stage breast cancer using spatial analysis and area based measures. Cancer Epidemiology Biomarkers and Prevention, 2007. 16(4): p. 756-762.

53. Centers for Disease Control and Prevention. National Breast and Cervical Cancer Early Detection Program. 2008 [cited January 22, 2014. Available from http://www.cdc.gov/cancer/nbccedp/]

54. Ghazali, S. M. et al. Non-Practice of Breast Self Examination and Marital Status are Associated with Delayed Presentation with Breast Cancer. Asian Pacific J Cancer Prev, 2013. 14: p. 1141-1145.

55. Woolf et al. The 2009 Breast Cancer Screening Recommendations of the US Preventive Services Task Force, 2010. JAMA. 303(2): p.34-36.

56. Robb-Nicholson, C. A Doctor talks about screening mammography. Harvard Women's Health Watch, 2010. Pages: 4-5.

57. Wang, F. et al. Late Stage Breast Cancer Diagnosis and Health Care Access. The Professional Geographer, 2008. 60(1): p. 54-69.
58. Anderson, R. T. Breast Cancer Screening, Area Deprivation, and Later Stage Breast Cancer in Applachian: Does Geography Matter? Health Services Research, 2011. 10(1): p.345-358.
59. Merkin et al. Geographic Socioeconomic Status, Race and Advanced Stage Breast Cancer in New York City. American Journal of Public Health. 92(2): p. 64-70
60. Dalton, et al. The Relation between socioeconomic and demographic factors on tumour stage in women diagnosed with breast cancer in Denmark, 1983-1999, British Journal of Cancer. 95(1): p.653-659.
61. Lin, G. and Zhang, T. Loglinear Residual Tests of Moran's I Autocorrelation and their Applications to Kentucky Breast Cancer Data. Geographical Analysis, 2007. 39: p. 293-310

XII. Biographical Sketch

Michael Jacob Sither is originally from Lexington, Kentucky. He received his Bachelor of Science degree in Biology from Centre College in Danville, KY in May of 2004. Currently, he is finishing his Doctor of Osteopathic Medicine (DO) degree at the University of Pikeville Kentucky College of Osteopathic Medicine and will complete his residency training in Internal Medicine at Lehigh Valley Health Network in Allentown, PA. He is a Master of Public Health candidate in Epidemiology at the University of Kentucky College of Public Health.

Name: Michael Jacob Sither Daytime phone: (859) 227- 4080
Email address: michaelsither@upike.edu or jacobsither@gmail.com