EPIDEMIOLOGY OF BREAST CANCER AMONG YOUNG BLACK WOMEN AND THE RISE IN YOUNG-ONSET DISTANT DISEASE IN THE U.S.

Lynn S. Chollet Hinton

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2017

Approved by:

Melissa A. Troester

Carey K. Anders

Jennifer L. Lund

Hazel B. Nichols

Andrew F. Olshan

ABSTRACT

Lynn S. Chollet Hinton: Epidemiology of Breast Cancer among Young Black Women and the Rise in Young-Onset Distant Disease in the U.S.
(Under the direction of Melissa Troester)

Purpose: Rates of distant (stage IV) breast cancer have significantly increased since 1976 among young women <40 years. Young-onset breast cancers tend to be more aggressive with poorer prognosis than older-onset disease, particularly among black women. This dissertation sought to clarify the impact of shifting incidence by (1) characterizing the epidemiology of young black women's breast cancer, and (2) investigating temporal shifts in breast cancer biology and diagnostic imaging use as contributors to rising young-onset distant disease.

Methods: We examined tumor characteristics and breast cancer risk factors associated with premenopausal young (<40) vs. older (≥40) black women's breast cancer in the African American Breast Cancer Epidemiology and Risk Consortium (2,008 cases; 5,144 controls) using unconditional logistic regression. Additionally, we examined longitudinal breast cancer incidence using joinpoint regression among young women (20-39 years) from 1992-2011 according to breast tumor characteristics in the Surveillance, Epidemiology, and End Results (SEER) program. Temporal patterns in imaging use (positron emission tomography, computed tomography, magnetic resonance imaging, bone scans) were examined separately among Medicare-eligible breast cancer cases using SEER-Medicare-linked data.

Results: Premenopausal black women <40 years had higher frequency of poorer-prognostic tumor characteristics compared to older (≥40) women, including negative estrogen and progesterone receptor (ER/PR) status, triple-negative subtype, high grade, higher stage, and larger tumor size. Adiposity, family history of breast cancer, and oral contraceptive use were associated with increased risk for young women while breastfeeding was more strongly protective. In SEER, the frequency of

favorable tumor characteristics significantly increased while less favorable characteristics declined among young women. Imaging use dramatically increased from 1992-2011 among SEER-Medicare cases and was significantly associated with less favorable characteristics, including ER/PR negativity, high grade, and tumor size >2cm.

Conclusions: Among premenopausal black women, young age (<40 years) was associated with more aggressive breast tumor biology. Modifiable risk factors including breastfeeding, adiposity, and oral contraceptive use may be important targets for mitigating harms of young-onset breast cancer. In SEER, the frequency of aggressive disease decreased while imaging use dramatically increased from 1992-2011, suggesting that stage migration rather than shifting tumor biology has contributed to rising incidence of young-onset distant breast cancer.

To Chris, with so much love and gratitude.

ACKNOWLEDGEMENTS

I am immensely grateful to so many people for their guidance and support throughout this work. First and foremost, I want to thank my advisor, Melissa Troester. I cannot imagine a more incredible mentor and I will always be grateful for your advice and friendship. You instilled in me a deep appreciation for breast cancer research and love of new challenges and collaborations. I cannot thank you enough for all the mentorship, encouragement, and laughs over the years.

I would also like to thank the members of my committee for their guidance. To Andrew Olshan, thank you for sharing your wealth of knowledge and expertise throughout this project and for your gracious support of many other research and professional opportunities. To Jennifer Lund, thank you for guiding me through a new world of methods with so much patience and positivity. To Hazel Nichols, thank you for providing such thoughtful and insightful feedback on complicated topics and for posing thought-provoking questions of your own. To Carey Anders, thank you for jumping into this project with such enthusiasm and for generously sharing your expertise. I have so enjoyed this opportunity to work with all of you and greatly appreciate your advice and conversation. Thank you for helping to turn a complex idea into a successful dissertation.

I am very thankful for many others for their help throughout this journey. To the Troester Lab members, past and present, I am deeply appreciative of your friendship and support. Thank you for all the constructive feedback and comradery, and for creating a collaborative and fun professional home all these years. To the friends I have made through this program, especially Greta Bushnell, Kathy Lancaster, Saira Khan, and my MSPH cohort, your friendship has been invaluable, and I look forward to new and exciting ventures to come. To the Epidemiology Department support staff, thank you for the encouragement and behind-the-scenes support to all students. To the Cancer Control Education Program, thank you for providing funding support, mentorship, and much-appreciated professional

advice. Thank you to all the participants and support staff involved in the studies included in this work, without whom this dissertation would not have been possible.

And last but certainly not least, I am forever grateful to my family. To my husband Chris, thank you for being the most wonderful partner in the truest sense of the word, helping to hold our family together and providing endless encouragement through all the ups and downs. To my daughter Maddie, thank you for your sweet smiles and strong hugs, and for keeping my life grounded and full of happy, hilarious surprises. Finally, thank you to all my family, especially Mom, Lucy, Ryan, and Clara, for your love and cheerleading across the miles.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMBER	African American Breast Cancer Epidemiology and Risk
AAPC	Average annual percent change
APC	Annual percent change
BMI	Body mass index
BWHS	Black Women's Health Study
CBCS	Carolina Breast Cancer Study
CI	Confidence interval
CMS	Centers for Medicare and Medicaid Services
CPT	Current Procedural Terminology
CT	Computed tomography
DAG	Directed acyclic graph
EMM	Effect measure modifier
ER	Estrogen receptor
HCPCS	Healthcare Common Procedure Coding System
HER2	Human epidermal growth factor receptor 2
ICD	International Classification of Diseases
IHC	Immunohistochemical
kg	Kilogram
m	Meter
MEC	Multi-Ethnic Cohort
MHT	Menopausal hormone therapy
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
NHANES	National Health and Nutrition Examination Survey
OR	Odds ratio

PET	Positron emission tomography
PR	Progesterone receptor
ref	Referent
SEER	Surveillance, Epidemiology, and End Results
TMA	Tissue microarray
TPL	Translational Pathology Lab
UNC	University of North Carolina at Chapel Hill
U.S.	United States
WCHS	Women's Circle of Health Study
WHR	Waist-to-hip ratio

CHAPTER 1: BACKGROUND

1.1 Overview: Breast Cancer in the U.S.

Breast cancer is the most common cancer and second leading cause of death among women in the United States (U.S.). Approximately 1 in 8 women are diagnosed with breast cancer during their lifetime, totaling nearly 300,000 new cases and 40,000 deaths each year(1, 2). In the U.S., the American Cancer Society recommends annual mammographic screening to all women over 40 years of age as a method for identifying breast cancers early, and as a result the majority of breast cancers are early-stage (*in situ* or stage I/II) rather than advanced-stage (stage III/IV) at time of diagnosis(3, 4). As with many other cancer sites, breast cancers that are localized to breast tissue (early-stage) have much improved treatment options and prognoses than do regional (stage III) or distant metastatic (stage IV) cases that have spread to other organ systems. However, breast cancer is a complex and heterogeneous disease, and trends in breast cancer incidence vary by age, race, and stage at diagnosis. Continued advances in breast cancer epidemiology are needed to further breast cancer prevention efforts and reduce the burden of advanced-stage breast cancers.

1.2 Breast cancer biologic heterogeneity

Although breast cancer as a whole has a high burden of morbidity and mortality in the U.S., breast cancer is not considered to be a single disease. Rather, breast cancer is comprised of a group of disease subtypes with distinct molecular, morphological, and clinical features that are associated with diverse incidence, prognosis, and survival patterns. Based on analyses of tumor gene expression, breast cancers have been divided into five subtypes (or six with the inclusion of normal breast-like tumors), including luminal A, luminal B, basal-like, HER2 over-expressing, and the recently identified Claudin-low tumors(5, 6). However, in the absence of tumor genomic data, these subtypes are approximated via immunohistochemical tests of hormone receptors (estrogen and progesterone

(ER/PR) receptors and human epidermal growth factor receptor 2 (HER2)), resulting in four conventional subtypes: luminal A (ER positive and/or PR positive and HER2 negative), luminal B (ER positive and/or PR positive and HER2 positive), basal-like (negative for all three receptors), and HER2-enriched (ER/PR negative and HER2 positive). Subtyping newly diagnosed breast cancers by immunohistochemistry has become routine clinical practice, revolutionizing the treatment of breast cancer by tailoring therapies to target to specific tumor subtypes.

However, tumor subtypes differ with respect to the availability of targeted therapies, aggressiveness, and clinical outcomes, resulting in disparities in incidence and prognosis across subtypes. Tumors with ER or PR positivity (luminal tumors) can be targeted with anti-estrogenic therapies in addition to conventional breast cancer treatments and tend to be lower grade and slower proliferating, resulting in favorable prognosis and improved survival compared to other subtypes(6-10). In contrast, HER2-enriched breast cancers are associated with aggressive, higher grade, and faster proliferating disease, making these breast cancers more likely to recur and metastasize following initial treatment and leading to poor patient outcomes(11-13). However, the introduction of HER2-targeting adjuvant treatments has substantially improved disease-free survival rates among women with HER2-positive breast cancers(14-16). Only basal-like breast cancers have no targeted therapies; these tumors grow independently of ER, PR, and HER2 expression and can have varied response to chemotherapy, greatly limiting treatment options(10). Furthermore, like HER2-enriched cancers, basal-like breast cancers tend to be more aggressive with higher grade and highly proliferative disease, thereby resulting in higher rates of disease recurrence and significantly reduced survival(8, 10). Figure 1.1 illustrates the survival curves for each subtype prior to the availability of targeted treatments, demonstrating the marked variation in survival for each subtype(6).

These biological differences across breast cancer subtypes suggest distinct etiologies, and epidemiologic studies have identified heterogeneity in the associations between breast cancer risk factors and risk of each subtype, particularly when comparing ER-positive and ER-negative disease. While some risk factors (e.g., family history of breast cancer, alcohol consumption, smoking history,

oral contraceptive use, and hormone replacement therapy use) appear to impact breast cancer risk similarly across all subtypes, reproductive and body size exposures are differentially associated with luminal (ER-positive) and basal-like (ER-negative) tumors(17, 18). Specifically, luminal breast cancers are associated with older age at menarche, later age at first birth, lower parity, and postmenopausal adiposity, while basal-like breast cancers are associated with younger age at menarche, younger age at first birth, higher parity, lack of breastfeeding, and pre- and postmenopausal adiposity(17-19). Given the disparate associations with survival between subtypes, differences in risk factor distributions in luminal vs. basal-like breast cancers can have substantial implications for breast cancer aggressiveness, prognosis, and patient outcomes. Furthermore, the associations between risk factors and tumor subtypes are modified by both age and race.

1.3 Breast cancer heterogeneity by age

Increasing age is the strongest risk factor for breast cancer, with approximately 78% of all new breast cancers occurring among women over 50 years of age(20). However, age is a complex risk factor, serving as a proxy for known and unknown age-associated exposures (e.g., genetic, behavioral, or environmental exposures that shift throughout the life course) that may promote cancer development(21). As such, many breast cancer risk factors vary with respect to age, including reproductive, behavioral, body size, and environmental exposures. Thus, examining age trends in breast cancer incidence can reveal heterogeneity in how risk factor exposure contributes to breast cancer risk over the life course.

As shown in Figure 1.2, age at diagnosis is strongly associated with overall breast cancer incidence, and the incidence of breast cancer among young women under 40 years of age is low. Indeed, less than 7% of all breast cancers in the U.S. are diagnosed among women <40 years(1, 22). However, although breast cancer among young women <40 years is rare, it is well-established that young women's breast cancer is more aggressive and proliferative than disease among older women. Young women tend to have a higher prevalence of basal-like breast cancers, larger and higher-grade primary tumors, ER and PR negativity, HER2 overexpression, p53 mutations, lymphovascular

invasion, and increased risk of regional and distant metastases than women diagnosed at a later age(22-27). Additionally, breast cancer tumors arising in young women have distinct patterns of gene expression compared to those occurring in older women, suggesting that breast cancers among young vs. older women are biologically distinct(27). The poor prognosis associated with younger-onset breast cancer has been shown to significantly reduce overall and relapse-free survival(22, 26-30). Additionally, women <40 years typically do not receive mammographic screening, and the incidence and survival patterns associated with older, screened populations (e.g., early-stage cancers, increased ER/PR-positive disease, and smaller, lower grade tumors) are not evident within this demographic. Indeed, the majority of breast cancers among young women are self-detected and more advanced at time of diagnosis(29, 31), contributing to a higher incidence of advanced-stage disease.

Epidemiologic studies have identified differences in the distribution of breast cancer risk factors with respect to age at diagnosis that may contribute to etiologic differences in young- vs. older-onset breast cancer(32-36). However, previous studies have used inconsistent definitions of young age, with some studies using varying young age cutoffs (40, 45, or 50 years) and others using pre- vs. postmenopausal status to represent young vs. older age(17, 34, 35, 37-40). Furthermore, the reported associations between breast cancer risk factors and age at diagnosis have been mixed, yielding inconsistent conclusions regarding whether breast cancers arising in young and older women have distinct etiologies.

Reproductive and body size exposures are the most established risk factors with heterogeneity in the effect on breast cancer risk across age. For example, risk of breast cancer has been shown to be temporarily elevated in the first several years following a pregnancy, after which parity is associated with reduced risk of breast cancer later in life(41-43). Breast cancers arising during or following pregnancy are known as pregnancy-associated breast cancers and are associated with more aggressive disease, higher likelihood of metastasis, and poor clinical outcomes(41, 42). The period of increased risk following pregnancy occurs among women of all ages, although the risk for women with later age at first birth (over age 30-35 years) is significantly higher than that for

women with younger age at first birth(41, 42). Because young women <40 years of age are more likely to have been recently parous than older women, young parous women are at increased risk of breast cancer while older parous women experience a protective effect(35, 42, 43). Additionally, oral contraceptive use has been reported to increase breast cancer risk among young women but not older women, a finding thought to be due to recency of oral contraceptive use among young women during their reproductive years(34, 35). Finally, obesity is well-known to be associated with reduced risk of breast cancer among young women and elevated risk among older women(44-47); however, this inverse relationship appears to be limited to ER/PR positive breast cancers, as obesity does appear to increase the risk of basal-like breast cancer among young women(17, 19).

Breast cancer risk modification by age for other established risk factors has been less clear (e.g., age at menarche, age at last birth, breastfeeding history, alcohol consumption, and smoking history)(32, 35, 36). Because breast cancer among young women is rare, previous work has generally been conducted in study populations with limited representation of young women(34-36, 38), resulting in uncertainty about the epidemiology of young women's breast cancer. Age differences in breast cancer tumor biology and aggressiveness are well-established, and further work examining the epidemiology of young-onset breast cancer in larger populations of young women is needed to identify modifiable targets for breast cancer prevention within this demographic.

1.4 Racial differences in breast cancer incidence

Breast cancer incidence patterns are known to differ significantly between black and white U.S. women. The "black-white crossover" is a well-described phenomenon that has been observed when comparing breast cancer incidence by age and race. This crossover, illustrated in Figure 1.3, refers to the incidence shift that occurs around age 40: black women have higher incidence rates than white women until age 40, after which black women have lower incidence rates than white women(2, 48, 49). However, black women tend to have more aggressive disease and higher mortality rates than white women at all ages(2), and these racial differences in breast cancer incidence and patient outcomes suggest etiologic differences by race.

Indeed, numerous studies have identified differences in the presentation of breast cancers among black and white women that are not entirely explained by differences in cultural or socioeconomic factors. Black women are more likely to have higher stage disease at diagnosis as well as larger, higher grade tumors and an increased prevalence of ER-negative, p53 mutated, and basal-like tumors than are white women (8, 21-23, 48-57). Furthermore, black women are less likely to report having a mammogram in the previous three years(53), more likely to experience treatment delays (particularly among young black women)(58), and less likely to receive a first treatment course that meets national cancer treatment standards(55), all contributing to poorer clinical outcomes and increased mortality among black women. Taken together, these established racial differences in breast cancer presentation and clinical outcomes suggest that breast cancers among black and white women may be biologically distinct.

Racial differences in breast cancer epidemiology are also important in breast cancer disparities. Aggressive breast cancers are more prevalent among young women <40 years, and a higher proportion of young women with breast cancer are black(23, 48). Indeed, black and white women appear to have distinct age at incidence curves (Figure 1.4), with a higher proportion of young cases being black and a higher proportion of older cases being white(48). Both races appear to have bimodal age at incidence trends with clusters of breast cancer diagnoses around 40 and 70 years of age, although black women tend to be diagnosed at earlier ages than white women (Figure 1.4). In studies comparing risk factor exposure distributions by race, black women tend to have higher obesity rates, reduced breastfeeding, and higher parity than white women(17, 56, 59), factors that are all associated with increased risk of basal-like breast cancer. Indeed, approximately 20% of the difference in late-stage disease between black and white women has been attributed to obesity alone(56), while 53% of basal-like breast cancers occurring in black women are thought to be preventable by reducing obesity and increasing breastfeeding among black women(17).

However, young black women under age 40 years are at greatest risk for aggressive breast cancers, and very little research has investigated whether breast cancer risk factor exposure differs

with respect to age among black women. Hall et al.(59) reported increased risk of breast cancer among younger (<50 years) black multiparous women (3 or more pregnancies) compared to older (≥50 years) multiparous black or white women. Mayberry et al.(60) observed that young (<40 years) black women with long oral contraceptive use (at least 10 years) and obesity had elevated risk compared to older black women with similar oral contraceptive use and obesity, while age at menarche under 13 years was protective against breast cancer for young but not older black women. Other work examining the epidemiology of young- vs. older-onset breast cancer among black women has been limited and reported null or imprecise effect estimates due to small sample sizes(17). However, these studies suggest heterogeneity in the epidemiology of breast cancer by age among black women, and additional work in larger study populations with sufficient representation of young black women is needed.

1.5 Longitudinal incidence trends in breast cancer

Recent studies of longitudinal breast cancer incidence trends have revealed shifts in breast cancer incidence over time related to age, race, and specific tumor characteristics. Overall, breast cancer incidence trends have varied substantially over the past several decades, largely due to advancements in screening and changes in exposure patterns to breast cancer risk factors (e.g., rising obesity rates, delayed childbearing, and decreased hormone replacement therapy use)(20). With the introduction of widespread mammographic screening in the late 1970s, overall breast cancer incidence rose dramatically in the 1980s and 1990s due to marked improvements in the detection of early-stage (*in situ* and stage I and II) disease (Figure 1.5)(3, 4). Following this surge in incidence, breast cancer rates began to decline in the late 1990s, a trend that has since been attributed primarily to decreased use of menopausal hormone therapy (MHT) following the highly publicized risks of MHT use reported by the Women's Health Initiative(3, 61). However, evidence suggests that the rates of early-stage breast cancers have again been rising in recent years(3, 4), and the underlying reasons for this recent increase are unknown. In comparison, the incidence of late-stage breast cancers (stage III and IV) has been markedly lower since the advent of widespread mammography, and the majority

of breast cancers diagnosed in the U.S. are early-stage (Figure 1.5). Over the past several decades, late-stage breast cancer incidence has been relatively stable with a slight but significant decrease in the incidence of regional and distant metastatic breast cancers(3, 4).

Because breast cancer among young women is rare, breast cancer incidence trends among U.S. women overall more strongly represent older women. Indeed, recent evidence suggests that longitudinal breast cancer incidence rates differ among young vs. older women, particularly with respect to stage at diagnosis. Among women <40 years of age, the incidence of localized and regional disease at diagnosis has been stable since 1976; however, rates of distant metastatic breast cancer have significantly and consistently increased since 1976 among young women(21, 62), with one study estimating that distant disease rates have nearly doubled within this demographic(4). As shown in Figure 1.6, distant disease has been increasing among young women more sharply in recent years, with a 3.6 percent increase in incidence each year from 2000-2009(4). Although breast cancer incidence is low among young women, the prognosis of metastatic breast cancer is poor (5-year survival is 25%)(63), and the increasing trend has inspired substantial concern. Johnson et al.(4) also found that distant disease significantly increased among women aged 40-54 years, albeit at a slower rate of 0.6% per year. In contrast, the incidence of distant disease remained stable for women 55-69 years and decreased slightly but significantly for women aged 70-84 years at diagnosis(4). In older women, Anderson et al.(21) reported stable or decreasing rates of distant disease and observed a stage-shift from distant/regional to local disease. Thus, women of older ages appear to mirror incidence trends of U.S. women overall, with increasing early-stage and decreasing late-stage breast cancers over time, while incidence rates of distant disease are significantly increasing among young women at an increasing rate over time.

Given that young black women have a higher prevalence of aggressive and late-stage breast cancers compared to young white and older women, increased distant disease among women <40 years may have been influenced by incidence among young black women. While several studies have examined whether longitudinal breast cancer incidence trends vary by race in the U.S.(1, 2, 21, 48,

64, 65), research investigating incidence patterns by age among black women is limited(21, 48, 64). Anderson et al.(21) identified increasing rates of distant disease among both young black and white women from 1974-2003, though black women aged 20-29 years had the greatest increase of any demographic group. Later, Anderson et al.(48) reported that more white than black women were diagnosed at an early age from 1975 to 1995, after which black women represented a larger proportion of early-onset disease. Similarly, Hou et al.(64) identified that young black women had higher incidence rates than young white women from 2000-2009, although they did not examine how trends in young-onset distant disease varied by race. While these studies demonstrate racial differences in breast cancer incidence by age, no known studies have investigated stage-specific incidence trends in young black vs. white women past 2003, when rates of distant disease were increasing most rapidly(4).

Several studies have examined how breast cancer incidence has varied by tumor characteristics over time. Evidence suggests that among U.S. women overall and among black and white women under age 50 years, ER positive disease has significantly increased from 1974-2010 (combined study years)(1, 21, 64). These studies have also reported reduced incidence of ER negative disease overall and either reduced(1, 64) or stable(21) ER negative disease among younger black and white women. These trends are consistent with expectation given large representation of older women. However, the observed decrease in ER negative disease among younger black and white women suggests that more aggressive ER negative tumors may be declining; thus, the increase in distant disease within this demographic may be attributable to factors other than temporal shifts in ER status. However, trends in ER status have not been examined in conjunction with stage at diagnosis, and ER status may be independently associated with trends in distant disease. With regard to other tumor characteristics, Anderson et al.(21) investigated how tumor grade varied over time among young women but found that both high and low grade disease increased in parallel among young black and white women through 2003 with no evidence that tumors of either grade were increasing more rapidly. Only one known study has investigated trends in other tumor characteristics (i.e., PR

status and tumor size), and this study did not include young women <40 years nor consider differences in trends by race. Shifts in other aggressiveness markers (e.g., PR and HER2 status, tumor size, grade, tumor subtype, and lymph node positivity) may have contributed to the rise in distant disease among young women and merit further study.

Finally, while temporal shifts in tumor aggressiveness may have impacted longitudinal breast cancer incidence trends, stage migration due to advances in diagnostic imaging technology may have also contributed to the recent rise in distant breast cancer among young black and white women. This phenomenon refers to a shift in the distribution of cancer stages within a population that is unrelated to cancer biology. Rather, changes to the cancer staging system or the use of improved diagnostic techniques results in the classification of cancer patients into different stages than would previously have been assigned(66). Evidence from studies of lung cancer suggests that the use of new imaging technology (i.e., computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI)) identifies previously undetectable distant metastases, resulting in misleading stage-specific incidence and survival trends as patients are classified as stage IV rather than stage III without a change in the actual state of their disease(66-69). It is unknown whether stage migration has contributed to the rise in distant metastatic breast cancer as well. The influence of imaging use on breast cancer staging has not been investigated, and shifts in the clinical recommendations for advanced imaging over time based on breast cancer tumor characteristics may have impacted national trends in breast cancer incidence. Investigating whether new imaging technologies have contributed to stage-specific incidence trends over time will reveal whether shifting trends in distant disease could simply be spurious artifacts of stage migration and shed light on the clinical use of these technologies.

1.6 Conclusions and study rationale

In summary, breast cancer is a complex disease composed of multiple distinct subtypes with incidence patterns that vary by age and race over time. Young black women less than 40 years of age are particularly at risk for aggressive, advanced-stage breast cancers with fewer treatment options,

worse prognosis, and decreased survival. The rate of distant disease at diagnosis has significantly increased among young black and white women since 1973, with the sharpest increases occurring since 2000. This rise in advanced-stage disease among women most at risk for poor clinical outcomes is concerning, and the underlying reasons for the increasing trend are unknown. These young-onset breast cancers appear to have a distinct etiology from older-onset breast cancers; however, studies investigating differences in breast cancer aggressiveness and epidemiology among young vs. older women have been rare and generally underpowered. Furthermore, the epidemiology of breast cancers arising in young black women is poorly understood. Further study investigating the role of breast cancer biology and stage migration on incidence trends in young black women's breast cancer is needed to understand the changing burden of breast cancer incidence.

Figure 1.1. Overall survival for the five original gene expression-based tumor subtypes. Sorlie et al.(6).

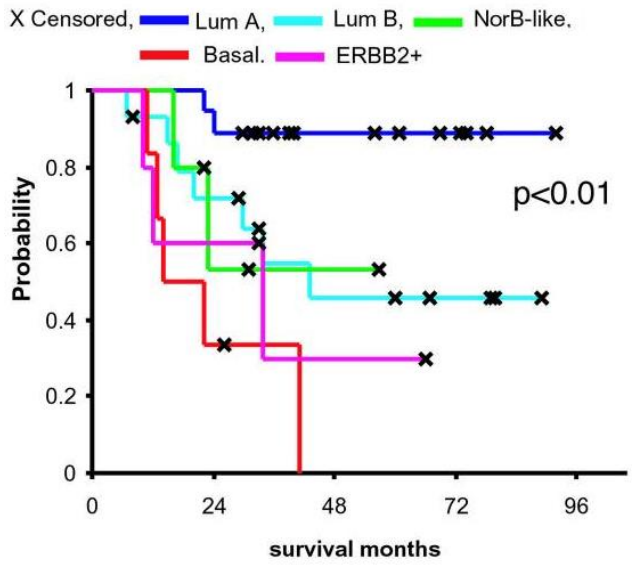


Figure 1.1 legend. ERBB2+ represents HER2+ subtype.

Figure 1.2. Age-specific breast cancer incidence rates in SEER, 1973-2010.

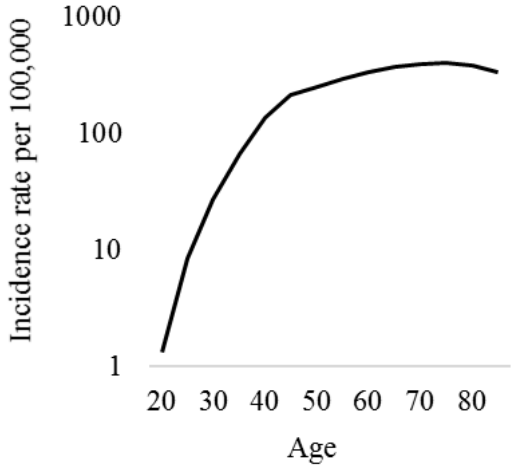


Figure 1.3. Age-specific incidence rates for breast cancer among white and black women in SEER, 1975-2004. Anderson et al.(48).

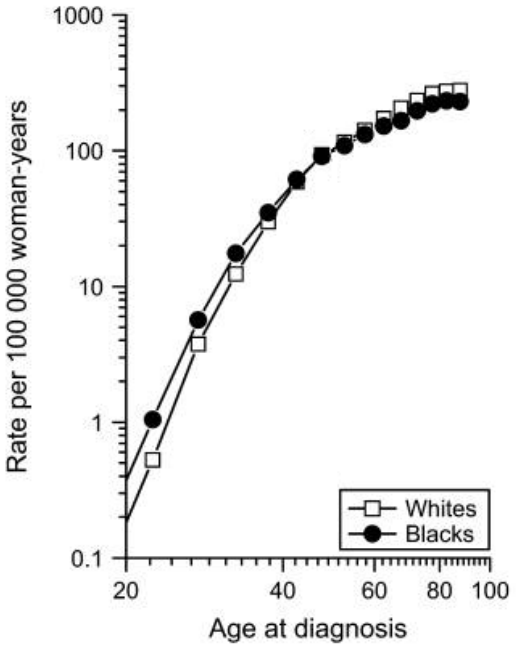


Figure 1.4. Age distributions at diagnosis by race in SEER, 1975-2004. Anderson et al.(48).

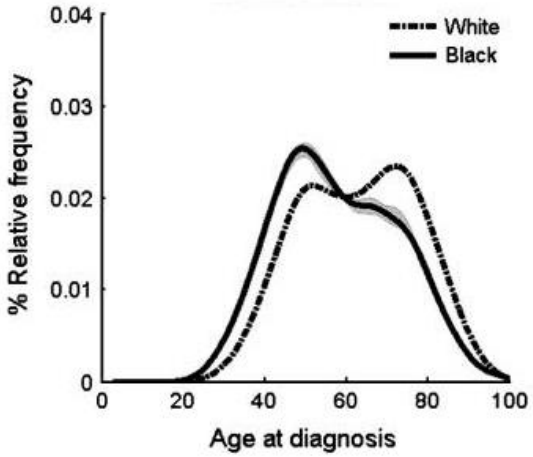


Figure 1.5. Incidence of stage-specific breast cancer in the U.S. in SEER, 1976-2008. Bleyer et al.(3).

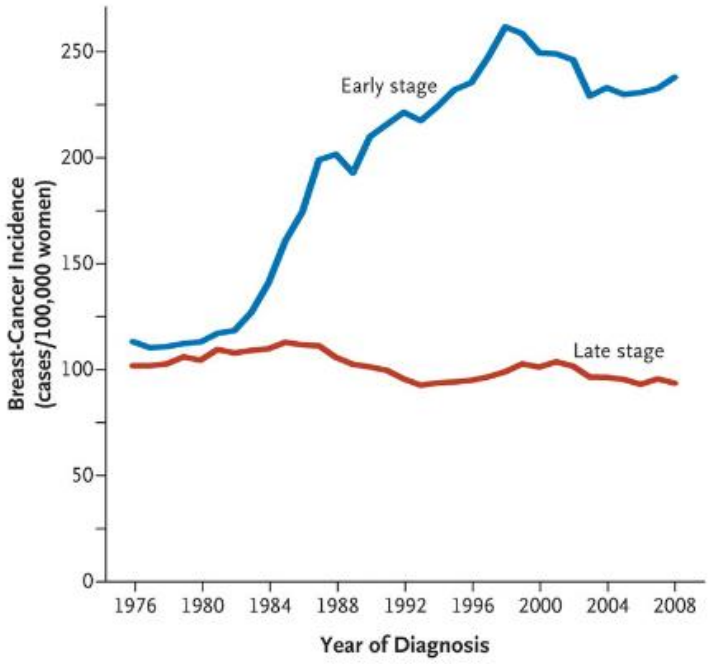
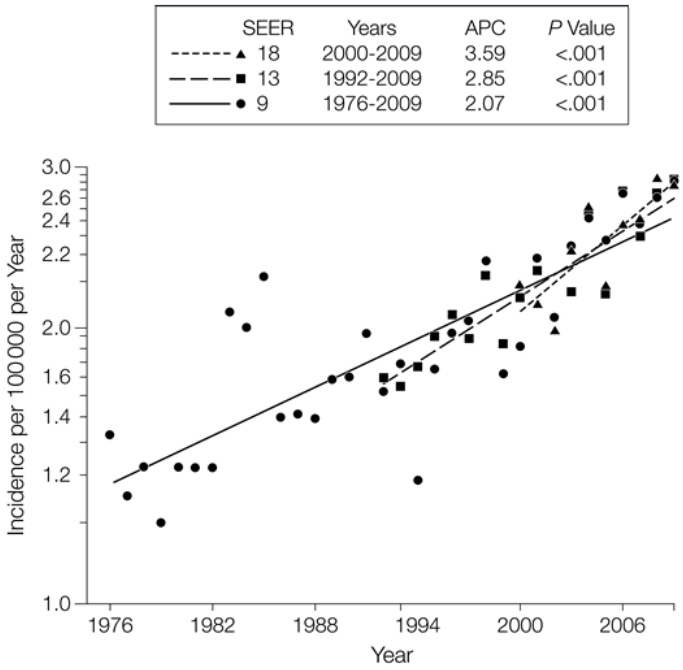


Figure 1.6. Annual incidence of distant breast cancer among young women (25-39 years) from 1976-2009 by SEER registry and era. Johnson et al.(4).



CHAPTER 2: SPECIFIC AI MS

Breast cancer is the most commonly diagnosed cancer and the second leading cause of death among women in the U.S.(2). Significant advancements in breast cancer epidemiology, screening, diagnosis, and treatment have resulted in improved breast cancer detection and survival over time(2, 20, 61, 70, 71). However, recent epidemiologic evidence suggests that rates of distant metastatic (stage IV) breast cancer have increased since 1976, especially among young women <40 years of age(4, 21, 62). These observed trends are particularly pronounced among young (<40 years) black women, among whom distant disease rates have increased more sharply compared to those for white women(4, 21). Although breast cancer incidence is low within this age group relative to older women, these findings are of concern given that breast cancer among young black women is associated with more aggressive and proliferative breast cancer subtypes, poorer prognosis, fewer therapy options, and reduced survival compared to diagnoses at a later age(17, 22, 28). The factors contributing to this rise in distant metastatic breast cancer are unknown, and an investigation into the epidemiology of young-onset breast cancer among black women is needed to identify targets for mitigating or preventing advanced-stage disease. Aggressive breast cancers are more prevalent among young black women(8, 17, 50), and we hypothesize that temporal shifts in breast cancer aggressiveness may have contributed to the recent rise in distant disease within this demographic.

Aim 1. To identify tumor characteristics and risk factors for young-onset (<40 years) breast cancer among premenopausal black women in the AMBER Consortium. To describe the characteristics of breast cancer among young black women, we will identify (a) tumor characteristics (ER, PR, and HER2 positivity; subtype; grade; stage; tumor size; and lymph node status) and (b) breast cancer risk factors (BMI, parity, breastfeeding history, oral contraceptive use, etc.) that are differentially associated with breast cancer according to age at diagnosis (<40 vs. ≥40 years).

Hypothesis: Young age at diagnosis (<40 years) will be associated with more aggressive tumors (ER/PR negativity, HER2 positivity, high grade, advanced stage, large primary tumor size, and lymph node positivity) and with distinct breast cancer risk factor distributions (especially in relation to reproductive exposures).

Aim 2. To describe the biological and imaging patterns associated with distant disease among black and white women.

Aim 2a. To examine how tumor biological characteristics contribute to temporal trends in distant disease, we will use joinpoint regression(72) and SEER 1992-2011 data to estimate the association between tumor characteristics (ER/PR status, tumor grade, and primary tumor size) and stage-specific incidence rates among young women <40 years, stratified by race (black vs. white).

Hypothesis: Young cases with distant stage disease will have increasing markers for aggressive disease over time, supporting the "biological shift" hypothesis(17, 37, 73-75).

Aim 2b. Shifts in imaging technology use may have inflated trends in the rates of distant disease due to improved ability to detect asymptomatic metastases at diagnosis(66-69). To evaluate whether stage migration due to increased imaging technology use (CT, PET, MRI, and/or bone scans) over time has contributed to rising rates of distant breast cancer among young women, joinpoint regression and SEER-Medicare linked data (1992-2011) will be used to describe how imaging use has changed over time and whether imaging use patterns differ by breast cancer tumor characteristics (ER/PR status, tumor grade, and primary tumor size) among older (≥65 years) U.S. women.

Hypothesis: Imaging use will have increased over time and will vary by tumor characteristics, with the greatest increases in use occurring among women with aggressive tumor characteristics, supporting the "stage migration" hypothesis.

The proposed study will investigate whether shifts in breast cancer aggressiveness contributed to the rise in distant metastatic breast cancer among young women. Characterizing the biology, epidemiology, and diagnostic imaging characteristics of aggressive, advanced-stage breast cancers over time will yield targets for intervention and clarify the public health impact of these trends.

CHAPTER 3: METHODS

3.1 Study Design Overview

This dissertation involved two overarching aims and analysis strategies to characterize young-onset breast cancer among black women and examine temporal trends in the incidence of distant stage disease. Figure 3.1 illustrates a conceptual model of these aims. In Aim 1, case-case and case-control analyses were used to identify tumor characteristics and breast cancer risk factors that were differentially associated with young- vs. older-onset breast cancer among black women in the African American Breast Cancer Epidemiology and Risk (AMBER) Consortium. Second, Aim 2 used joinpoint regression to examine how temporal trends in breast tumor characteristics (Aim 2a) and imaging use (Aim 2b) have contributed to national trends in stage-specific breast cancer incidence rates among young (<40 years) U.S. women in the Surveillance, Epidemiology, and End Results (SEER) 13 program from 1992-2011. Using SEER-Medicare linked data from 1992-2011, Aim 2b also characterized whether imaging use varied by stage at diagnosis and breast tumor characteristics.

3.2 Study Populations

To address the aims of this dissertation, data from three study populations were used: the African American Breast Cancer Epidemiology and Risk (AMBER) Consortium, the Surveillance, Epidemiology, and End Results (SEER) 13 program from 1992-2011, and the SEER-Medicare linked program from 1992-2011.

3.2.1 AMBER Consortium

The African American Breast Cancer Epidemiology and Risk (AMBER) Consortium is a collaboration between four epidemiologic studies of breast cancer among black women: the Carolina Breast Cancer Study (CBCS), the Black Women's Health Study (BWHS), the Women's Circle of Health Study (WCHS), and the Multi-Ethnic Cohort (MEC) Study(76). These four studies combined

breast cancer risk factor exposure data (genetic, biologic, reproductive, and lifestyle) for black women with and without breast cancer, as well as molecular and genetic tumor characteristics for breast cancer cases. Taken together, these parent studies represent a pooled study population of over 10,000 U.S. black women with data extending from 1993-2014, making the AMBER Consortium one of the largest and most extensive breast cancer studies of black women in the U.S.

Carolina Breast Cancer Study (CBCS). The CBCS is a population-based, case-control study involving 24-44 counties of central and eastern North Carolina. Women of all races were eligible for inclusion if they were 20-74 years of age at the time of diagnosis (cases) or study recruitment (controls). As described previously(17, 58, 77, 78), the CBCS collected extensive clinical, molecular, and epidemiologic data for *in situ* and invasive breast cancer cases diagnosed from 1993-1996 (Phase I), 1996-2001 (Phase II), and 2008-2013 (Phase III). Cases were identified through rapid case ascertainment by the North Carolina Central Cancer Registry; controls were recruited (for Phases I and II only) through North Carolina Department of Motor Vehicles records and Health Care Financing Administration records for Medicare enrollment. Phase III recruited only invasive breast cancer patients and expanded the case catchment area to 44 North Carolina counties. Randomized recruitment was used to oversample younger (<50 years) and African American cases as well as to frequency-match controls to cases by age (<50 vs. ≥50 years) and self-reported race (African American vs. non-African American)(77). The CBCS contributed 701 premenopausal African American breast cancer cases from Phases I-III and 298 controls from Phases I-II to this analysis.

Black Women's Health Study (BWHS). The BWHS is a prospective cohort study of U.S. black women's health from 17 mainland states; the regional distribution of participants is 28% Northeast, 30% South, 23% Midwest, and 19% West(79). Participants were recruited in 1995 through mail questionnaires sent to *Essence* Magazine subscribers, members of the Black National Education Association and Black Nurses' Association, and friends and relatives of respondents(79). A total of 59,000 black women aged 21 to 69 were enrolled following completion of a baseline questionnaire assessing demographic information, medical and family history, and biologic, reproductive, and

lifestyle exposures(79). Since 1997, follow-up questionnaires have been mailed to all participants biennially to update exposure history and record incident disease and mortality data, with a follow-up success rate of 80%(80). For participants reporting breast cancer diagnoses, clinical information relating to the diagnosis and tumor pathology was abstracted from medical records or obtained from local cancer registries(80). The BWHS contributed 738 premenopausal cases and 4,281 controls to this analysis.

Women's Circle of Health Study (WCHS). The WCHS is a case-control study of breast cancer involving black and white women living in New York and New Jersey. Initially, the WCHS was a hospital-based case-control study recruiting English-speaking white and black women aged 20-75 years from New York boroughs (Manhattan, Bronx, Brooklyn, and Queens) beginning in January 2002(81). Women with histologically-confirmed incident breast cancers were defined as cases and women with no previous diagnosis of cancer (other than non-melanoma skin cancer) were defined as controls. However, in March 2006 the study later expanded to New Jersey and became population-based, with cases identified from the New Jersey Cancer Registry and controls recruited by random digit dialing or through church and health fair events occurring in communities from which cases were recruited(81). Controls were frequency-matched to cases by race and 5-year age groups. The WCHS is ongoing, with current enrollment limited to black women living in 10 New Jersey counties; New York recruitment was discontinued in December 2008(81). All cases and controls are interviewed in their homes by study personnel, and all participants provide biologic, lifestyle, reproductive, and medical history information. The WCHS contributed 569 premenopausal black cases and 565 controls to this analysis.

Multi-Ethnic Cohort (MEC). The MEC is a prospective cohort study conducted in Hawaii and Southern California examining risk factors associated with cancer in an ethnically diverse study population of white, black, Asian American, Japanese American, Latino, and Native Hawaiian men and women. Over 215,000 English- and Spanish-speaking participants (16,594 black women) aged 45 to 69 years were recruited from 1993-1996 through a mailed baseline questionnaire assessing

demographic, medical, reproductive, lifestyle, and family histories. Participants completed a five-year follow-up questionnaire in 1999-2001 and repeated the baseline questionnaire in 2003-2008 to update exposure data. Incident breast cancer cases were identified by linkage with the Hawaii Tumor Registry and Los Angeles County Cancer Surveillance Program, and mortality information is obtained through linkage with the California and Hawaii state death certificate files as well as the National Death Index. While the MEC contributes approximately 1,053 incident breast cancer cases to AMBER, the MEC was excluded from the present analysis, as their exclusion criteria limited participants to ≥45 years of age.

Pooled AMBER data. The final AMBER study population includes black premenopausal breast cancer cases and controls from the CBCS, BWHS, and WCHS parent studies. The two case-control studies (CBCS and WCHS) contribute all enrolled premenopausal cases and controls to AMBER, while the BWHS provided nested case-control data comprised of all premenopausal incident breast cancer cases and up to four controls for each case. The BWHS controls were randomly selected and matched to cases by year of birth and by completion of the same follow-up questionnaire prior to the case's diagnosis(76). Data collection in AMBER is ongoing, with additional cases and controls added to the pooled study population as follow-up continues in the ongoing parent studies and as breast tumor tissue samples are processed for genetic and molecular data.

The AMBER study population for Aim 1 of this dissertation included all premenopausal black cases diagnosed with invasive breast cancer and matched controls from the CBCS (701 cases, 298 controls), WCHS (569 cases, 565 controls), and BWHS (738 cases, 4,281 controls), for a total study population of 7,152 women (2,008 cases, 5,144 controls). All postmenopausal women (defined based on self-reported cessation of menstruation, bilateral oophorectomy, or ovary irradiation), and women with unknown menopausal status were excluded to estimate age effects independent of menopausal status. Each study and the AMBER Consortium collaboration were approved by Institutional Review Boards at participating institutions, and all participants gave written informed consent.

3.2.2 SEER 1992-2011

The SEER research program, led by the National Cancer Institute (NCI), has collected cancer incidence and survival data from a collection of U.S. cancer registries since 1973. The original SEER program involved 9 cancer registries and has since expanded to 18 registries throughout the U.S. (Table 3.1). The SEER expansions were completed to diversify the study population and increase heterogeneity by race/ethnicity, thus improving the external validity of the SEER program over time(82). SEER currently covers approximately 28% of the total U.S. population, and the registries participating in SEER are carefully selected to ensure that included cancer cases are representative of the general U.S. population in terms of education status and poverty level and that minority races and ethnicities are adequately represented(82, 83). In general, the SEER program is highly generalizable in terms of education, poverty level, and race/ethnicity, although SEER tends to have slightly higher proportions of urban and foreign-born persons than the total U.S. population(82).

Because the SEER program is a collection of cancer registries, all incident cancer cases diagnosed in participating areas are reported to SEER by local hospitals, clinicians, and pathology laboratories. SEER requires that each registry report all cases within two years of diagnosis, after which cases are followed for demographic, clinical, and mortality data(82). This selection process results in a complete population of all cancer cases within participating geographical areas that together are representative of the general U.S. population. In total, the SEER program captures incidence and survival data for over 7.7 million cancer cases throughout an almost 40-year study period, making the SEER program the largest and most comprehensive population-based cancer epidemiologic data source in the country.

This dissertation involved SEER data collected from 1992-2011, and thus this subpopulation included breast cancer cases in the SEER 13 program, as these registries have continuously collected data throughout the study period (Table 3.1). Because SEER registries did not routinely report tumor characteristics prior to 1990 and 4 racially-diverse SEER registries were added in 1992(83), the subpopulation was limited to all breast cancer cases with diagnosis dates beginning in 1992 in order

to ensure consistent reporting of tumor characteristics variables and to maximize racial heterogeneity. The inclusion criteria for this study were 1) U.S.-residing women 20-39 years of age, and 2) a first primary breast cancer diagnosis (ICD-O-3 cancer sites C50.0-50.9) between January 1, 1992 and December 31, 2011 that was reported to SEER. Women with a personal history of cancer prior to the first primary breast cancer diagnosis or a primary breast cancer diagnosed at time of death or autopsy were excluded. The final SEER 13 study population (1992-2011) included 30,407 incident breast cancer cases among young women (<40 years).

3.2.3 SEER-Medicare 1992-2011

The SEER-Medicare program links the cancer incidence data within SEER to healthcare utilization claims data available in Medicare. This data linkage has existed since 1991 and includes all SEER cancer cases diagnosed after 1991 who were eligible for Medicare (≥65 years of age or diagnosed with end-stage renal disease, amyotrophic lateral sclerosis, or medical disability)(84). The SEER-Medicare linkage is updated every 2-3 years by the NCI, SEER, and CMS through a collaborative process in which eligible cancer cases are matched by personal identifiers (name, social security number, sex, and date of birth) to Medicare claims(85, 86). An estimated 93% of all eligible SEER cancer cases diagnosed since 1991 have been successfully linked to Medicare claims(86). Although the SEER-Medicare population has been found to have lower poverty rates, higher proportion of minority races, greater urban vs. rural living, and decreased cancer mortality than the general U.S. elderly population(85), the magnitude of these differences is small and the external validity of these data to the elderly U.S. population remains very high considering that 97% of elderly Americans are enrolled in Medicare(85).

The current study utilized SEER-Medicare data for breast cancer cases diagnosed from 1992-2011. Inclusion criteria were: 1) U.S.-residing women ≥65 years of age, 2) a first primary breast cancer diagnosis (ICD-O-3 cancer sites C50.0-50.9) between January 1, 1992 and December 31, 2011 that was reported to SEER, and 3) Medicare enrollment at time of diagnosis (parts A/B coverage and no health maintenance organization (HMO) enrollment). Women with a personal history of cancer

prior to the first primary breast cancer diagnosis or a primary breast cancer diagnosed at time of death or autopsy were excluded. The final subpopulation included 142,051 breast cancer cases ≥65 years from the SEER 13 registries that are linked to Medicare claims data.

The purpose of including SEER-Medicare data in addition to the full SEER program was to describe the use of diagnostic imaging technology over time in a national U.S. cancer population during a period of dramatic change in the use of imaging for breast cancer staging (1992-2011). Medicare data include the type of imaging technology used, diagnostic codes, and date of service, and the linkage to SEER enables an investigation into imaging trends over time within a national U.S. cancer population. While we were unable to directly assess temporal trends in imaging use among young women within this population, this aim sought to identify imaging patterns in a national population of breast cancer cases and evaluate how clinical recommendations for advanced imaging vary by disease characteristics in the U.S.

3.3 Data Collection

3.3.1 Aim 1: AMBER Consortium

Breast cancer risk factors. The AMBER Consortium's breast cancer risk factor data were collected through either self-reported mailed questionnaires (BWHS) or home interviews conducted by study nurses (CBCS and WCHS). Although the AMBER Consortium involved harmonizing risk factor data obtained from different study populations, the demographic, reproductive, and medical history data relevant to this proposed study were collected using very similar questions and coding schemes across all contributing studies. Participants were asked questions regarding their medical and family histories as well as biologic, anthropomorphic, reproductive, and lifestyle exposures. For CBCS and WCHS, interviewers also measured body weight, height, and waist and hip circumferences during home interviews; for the BWHS, these measures were self-reported on questionnaires by study participants. Questionnaire and interview data from each study were then harmonized by the AMBER Biostatistics and Data Management core within AMBER to create a central database with consistent exposure definitions across studies. Breast cancer risk factors were categorized as: age at menarche

(<13, ≥13 years), parity (nulliparous, 1-2, ≥3 births), age at first live birth (<25, ≥25 years), age at last live birth (<30, ≥30 years), time since last birth (<10, ≥10 years), lifetime duration of breastfeeding (never, <3 months, ≥3 months), oral contraceptive use (never, ever), duration (never/<1 year, 1-4 years, 5-9 years, ≥10 years) and recency (never, <10 years, ≥10 years), and first-degree family history of breast cancer (no, yes). Body mass index (BMI) was defined as body weight/height (kg/m2) using categories from the National Heart, Lung, and Blood Institute (<25 normal/underweight, 25.0-29.9 overweight, and ≥30 obese)(87). Waist-to-hip ratio (WHR) was calculated as the ratio of waist/hip circumference (cm) and categorized in tertiles as <0.77, 0.77-0.83, and ≥0.84, consistent with previous work(88).

Tumor characteristics. Tumor characteristic data were obtained from patient medical records or from tumor tissue blocks collected at time of surgery for all breast cancer cases. The three studies included in this analysis (CBCS, BWHS, and WCHS) contributed breast tumor tissue when available to two core research facilities (the Translational Pathology Lab (TPL) at the University of North Carolina at Chapel Hill (UNC) for the CBCS and the Roswell Park Cancer Institute for the WCHS and BWHS) where tissue microarrays (TMAs) were constructed for available tumor specimens. Immunohistochemistry (IHC) assays were conducted on all TMAs at UNC's TPL to define expression of estrogen and progesterone receptors (ER/PR) and human epidermal growth factor receptor 2 (HER2)(89). Positive expression was defined as ≥1% staining for ER and PR, and ≥10% staining at the 3+ level for HER2 consistent with previous work(89). Breast cancer subtype was defined as four groups based on positivity of three IHC markers: luminal A (ER+ or PR+, HER2-), luminal B (ER+ or PR+, HER2+), HER2+/ER- (ER-, PR-, HER2+), and basal-like (ER-, PR-, HER2-). For cases with missing IHC-based tumor characteristics, ER, PR, HER2, and subtype data were defined from medical records; cases with both IHC-based and clinical hormone receptor data showed high agreement for the two measures (κ statistic range=0.68-0.76, concordance range=88-91%). Tumor grade was centrally reviewed by a study pathologist for 44% of cases, with grade data obtained from medical records for remaining cases (κ statistic=0.95, concordance=96% for both grade

measures). Other tumor characteristics (including stage (1-4), lymph node status (positive vs. negative), and estimated tumor size (<2, 2-4.9, ≥5cm)) were acquired from medical records.

3.3.2 Aim 2: SEER and SEER-Medicare

Tumor characteristics. The SEER program records several markers of breast tumor aggressiveness (ER/PR status, tumor grade, primary tumor size, and stage) for all breast cancer cases. These tumor characteristics are obtained from pathology reports or medical records, and the accuracy of these markers is verified to be at least 98% through random medical record reviews(82, 85). ER/PR status was recorded as positive, borderline, negative, and unknown. Given that clinical standards defining hormone receptor positivity changed over the study period, varying from 1-10%, we combined positive and borderline cases to approximate more recent guidelines of 1% positivity recommended by the American Society of Clinical Oncology/College of American Pathologists(90). Tumor grade was defined at time of diagnosis by a pathologist as low (well-differentiated), moderate (moderately-differentiated), high (poorly-differentiated) and undifferentiated (anaplastic). Undifferentiated tumors (defined as grade 4) were uncommon (approximately 3% of cases) and were excluded from all grade analyses. Primary tumor size was defined as the single largest tumor dimension prior to any neoadjuvant cancer therapy and dichotomized as ≤2cm vs. >2cm.

Stage at diagnosis was defined using SEER's historic stage A coding of four summary stage categories, as this coding has remained consistent in SEER throughout the study period. While breast cancer staging systems have been refined and improved over time, the SEER and SEER-Medicare programs utilize an extent-of-disease reporting scheme that has been in place since 1988 to create SEER summary stages that have remained consistent over time(82). Specifically, stage at diagnosis was defined as *in situ* (noninvasive cancer), local (invasive cancer confined to the breast), regional (tumor extension to breast skin, chest wall, and/or regional lymph nodes), and distant (metastasis to non-breast tissues).

Diagnostic imaging technology: SEER-Medicare includes data regarding the use of diagnostic imaging technologies while enrolled in Medicare for all eligible cancer cases. In this

dissertation, diagnostic imaging technologies referred to all imaging technologies that have been used from 1992-2011 to identify whether a breast cancer has metastasized at the time of diagnosis (i.e., CT, PET, MRI, and/or bone scans). In this study, imaging use was defined as having at least one imaging claim in a given calendar year. Use of these technologies was defined using relevant Healthcare Common Procedure Coding System (HCPCS) and Current Procedural Terminology (CPT) imaging codes corresponding to these technologies from 1992-2011 (Table 3.2). Breast cancer cases were considered "imaged" if at least one of these four imaging scans was received between 2 months prior and 4 months following the primary diagnosis date within SEER. This exposure window was selected to identify imaging used incidentally prior to cancer diagnosis and to allow sufficient time for complete disease staging post-diagnosis. Breast cancer is not typically diagnosed and staged using a single diagnostic test: the disease is generally diagnosed via biopsy of a detected lesion in the breast(91-94), and final staging occurs following surgical removal of the lesion, lymph node biopsy/dissection, and/or diagnostic imaging. Time since initial diagnosis to complete staging has been reported to range from less than 1 month to over 3 months(57, 95-97), and similar studies examining the use of imaging technology at time of cancer diagnosis have previously utilized this 6-month timeframe to assess technology use throughout the diagnostic time period(68, 98).

3.4 Analysis methods

3.4.1. Aim 1. To identify tumor characteristics and risk factors for young-onset (<40 years) breast cancer among premenopausal black women in the AMBER Consortium.

Case-case and case-control analyses were conducted to identify differences in the associations between tumor characteristics and epidemiologic risk factors and breast cancer by age at diagnosis (<40 vs. ≥40 years) among premenopausal black women (age range 22-59 years). Case-case analyses of tumor characteristics associated with young- vs. older-onset disease included all cases (N=2,008), while case-control analyses of risk factors included all cases and controls except cases from Phase III of the CBCS (total N=6,736), as no matched controls were available for this study phase. Case-control analyses examined risk factor associations for breast cancer among young and

older women overall and further stratified by ER status among young women, in which ER-positive and ER-negative cases were compared separately to all controls. Unconditional logistic regression models were used to estimate adjusted odds ratios (ORs) and 95% confidence intervals (CIs) to assess differences in tumor characteristics and breast cancer risk factors associated with breast cancer by age at diagnosis for all analyses. In case-control analyses, effect measure modification by age was evaluated using likelihood ratio tests in which the estimated log-likelihood of the adjusted model was compared to that of the same model including a multiplicative interaction term for age and the corresponding risk factor. Statistically significant modification was assessed using an α-level of 0.1. Heterogeneity in risk factor associations by ER status among young women was assessed by comparing case-case odds ratios (ORs), with ER status defined as the outcome and each risk factor as the explanatory variable. These case-case ORs represent the ratio of case-control ORs for risk factors associated with ER-positive vs. ER-negative disease, and statistical significance was defined using an α-level of 0.05. Additionally, we conducted sensitivity analyses to examine whether patterns of tumor characteristics and risk factors associated with young- vs. older-onset disease were impacted by the cutpoint used to define young age (40, 45, and 50 years). All models controlled for study, diagnosis year, geographic region, and education status to account for differences between studies. Case-control models additionally adjusted for other risk factors that were identified *a priori* via directed acyclic graphs as potential confounders of each risk factor association. Models for age at first live birth, age at last live birth, time since last birth, and lifetime breastfeeding duration were restricted to parous women. Statistical significance was defined at an α-level of 0.05. All analyses were performed using SAS software, version 9.4 (SAS Institute, Inc., Cary, North Carolina).

3.4.2 Aim 2. To describe the biological and imaging patterns associated with distant disease among black women.

3.4.2.1 Aim 2a. To examine how tumor biological characteristics contribute to temporal trends in distant disease among young women <40 years in the SEER 13 program from 1992-2011.

To assess trends in distant breast cancer among young women <40 years of age during our 20-year study period, we used SEER*Stat software (version 8.3.2(99)) to calculate annual stage-specific breast cancer incidence rates from 1992-2011 in the SEER 13 program, adjusted for age using the 2000 U.S. standard population. Incidence rates were calculated for young women overall and stratified by black and white race to identify whether temporal trends in stage at diagnosis vary according to race.

Both Aims 2a and 2b used joinpoint regression(72) (version 4.4.0.0(100)) to model longitudinal SEER and SEER-Medicare data. Joinpoint regression, also known as piecewise regression, is an established statistical method for modeling cancer incidence rates over time, and the model estimation process has been described in detail previously(72). To summarize, logarithm-transformed incidence rates or proportions are modeled over time by fitting a series of linear trends that pivot at "joinpoints" corresponding to points in time at which the slope of the cancer incidence trend significantly changes. The total number of joinpoints is determined through a permutation process: the regression model first fits the simplest model with 0 joinpoints (meaning a single linear trend over the entire time period), after which the number of joinpoint terms is serially increased to test whether including additional joinpoints significantly improves model fit, up to a maximum of 3 joinpoints for the current study, based on time period length. Following this forward estimation process, the final model estimates the best-fit series of linear trends with slope changes at statistically significant joinpoints (see Figure 3.2 for example). The slope parameters estimated for each linear segment are known as annual percentage change (APC) estimates, and p-values for each APC are calculated via Monte Carlo permutation methods. Weighted average APCs (AAPCs) are then calculated to generate a single percent change estimate for the time period and to quantitatively compare trends across analyses(101).

To address Aim 2a and examine the "biological shift" hypothesis, we calculated annual stage-specific incidence rates and annual incidence rates according to ER/PR status, tumor grade, and primary tumor size, overall and among cases with distant stage disease at diagnosis. All incidence

rates were adjusted for age using the 2000 U.S. standard population, and analyses of tumor characteristics were restricted to invasive breast cancer cases (N=26,870). We then used joinpoint regression to calculate AAPCs estimating temporal trends in stage and breast tumor characteristics from 1992-2011 among young women overall and according to race (black vs. white). Substantial missing data has been documented for tumor characteristics reported to SEER, particularly for ER and PR status (approximately 25% in 1992), which may bias longitudinal studies of incidence rates(102, 103). To clarify whether data are missing at random across characteristics, we characterized patients with missing data by estimating descriptive statistics with chi-square significance testing for cases with and without missing data, and used joinpoint regression to examine whether the distribution of breast cancer cases (as measured by percent of cases) varied over time with respect to each tumor characteristic.

3.4.2.2 Aim 2b. To evaluate whether stage migration due to increased imaging technology use (CT, PET, MRI, and/or bone scans) over time has contributed to rising rates of distant breast cancer among young women <40 years.

To investigate the "stage migration" hypothesis, annual rates of imaging use were calculated overall and by imaging type (CT, PET, MRI, and bone scans) among older breast cancer cases in SEER-Medicare from 1992-2011. Joinpoint regression was used as described above to estimate AAPCs for all longitudinal trends. Sensitivity analyses were conducted excluding elderly women (over age 75) as well as varying the 6-month imaging window to 3 months post-diagnosis; however, these exclusions did not substantially impact results beyond reductions in study power and therefore were not employed in final analyses. Unconditional logistic regression was used to estimate age-adjusted odds ratios (ORs) and 95% confidence intervals (CIs) to assess the associations between imaging use and tumor biological characteristics. Logistic regression analyses were performed using SAS software, version 9.4 (SAS Institute, Inc., Cary, North Carolina). Statistical significance for all analyses was defined at an α-level of 0.05.

The concerns described above (3.4.2.1 Aim 2a) for missing breast tumor characteristics data in SEER also apply to the SEER-Medicare data. Unlike Aim 2a, Aim 2b did not involve the estimation of temporal trends in tumor characteristics with substantial time-varying missing data, and thus complete-case analyses were conducted to evaluate the association between imaging use and breast tumor characteristics. Additionally, the use of administrative claims data carried concerns about the accuracy of procedure codes (such as the CPT and HCPCS codes used in this study) that are reported in claims data. It is possible that procedure code errors resulted in misclassification of the imaging use exposure and thus created bias in the proposed statistical analyses. While no known validation studies have been conducted to evaluate the sensitivity and specificity of the relevant imaging technology codes, similar studies examining imaging technology use in cancer populations have previously utilized these same procedure codes with success(68, 95, 104). Additionally, these imaging technologies are expensive healthcare procedures that are more likely to be reported accurately to ensure appropriate provider reimbursement compared to other lower cost procedures, suggesting that imaging use data are likely to be well-represented in healthcare utilization data sources such as the SEER-Medicare program.

3.5 Strengths and Limitations

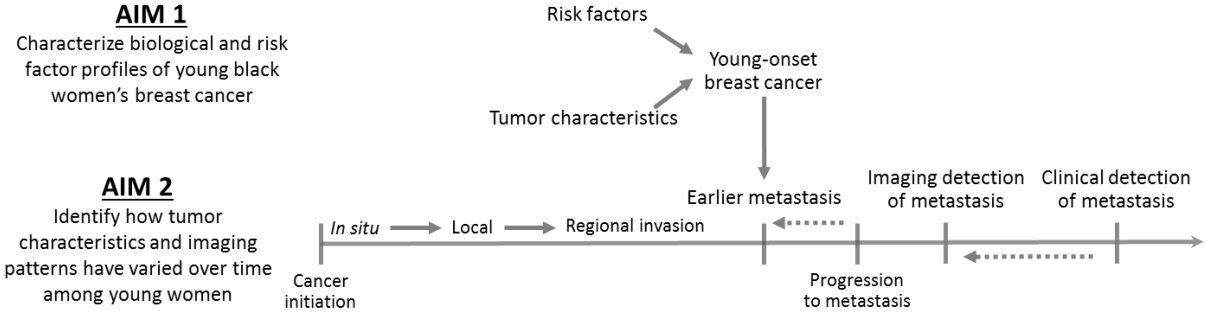
To our knowledge, this study is the first to investigate whether temporal shifts in breast cancer biology have contributed to the recent increase in distant metastatic breast cancer among young women. This work characterized the epidemiology of breast cancer among young black women by examining how a comprehensive list of tumor characteristics and breast cancer risk factors vary with respect to age at diagnosis. Previous studies of young-onset breast cancer among black women have been limited by small sample sizes, and by using AMBER data this study extended previous work using one of the largest available breast cancer epidemiology data sources for black women in the U.S.

Additionally, the SEER program is the most extensive and comprehensive population-based cancer epidemiologic data source in the U.S., and the use of this very large study population enabled

an examination of longitudinal incidence trends among young women <40 years, stratified by race and tumor characteristics. The use of SEER data also allowed for an investigation into shifts in tumor biology as an underlying reason for the increased rates of distant disease among young women within the data source originally used to identify those incidence trends. Furthermore, the SEER and SEER-Medicare data spanned from 1992 to 2011, a time period that encompasses the period of greatest change in distant disease incidence among young women and a period of rapid advancements in imaging technologies. The use of SEER-Medicare data allowed an exploration into the roles of shifting imaging use and stage migration in stage-specific breast cancer incidence trends among elderly U.S. women and evaluated how imaging use recommendations may vary clinically by tumor characteristics. Additionally, while missing tumor characteristic data was of concern for the early years of the SEER program, verified multiple imputation methods and careful evaluation of missingness within the SEER data were employed to address the missing data for ER/PR status and any other predictors.

Ultimately, this dissertation thoroughly characterized the epidemiology of young-onset breast cancer among black women and examined breast cancer biology and imaging use as potential underlying reasons for temporal shifts in stage-specific breast cancer incidence rates among young women <40 years in the U.S. This work sought to identify potentially modifiable targets for public health intervention against young-onset breast cancers among black women and clarify the public health impact of the rising incidence of distant disease among young women.

Figure 3.1. Summary of aims and hypotheses.



Hypotheses: (1) Young age (<40 years) is associated with distinct patterns of breast tumor biology and risk factor exposure.
(2) Etiologic shifts to more aggressive breast cancers and/or stage migration due to imaging use have increased the incidence of distant stage breast cancer among young women from 1992-2011.

Figure 3.2. Sample joinpoint regression model.

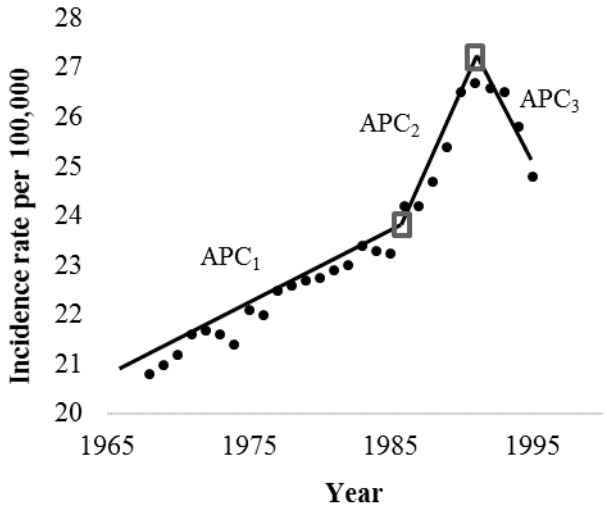


Figure 3.2 legend. Boxes mark two joinpoints resulting in three annual percent change (APC) estimates.

Table 3.1. Timeline of SEER program expansions, 1973-2011.(82)

SEER program	Data years	Included cancer registries
SEER 9	1973–2011	Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, Utah
SEER 11	1992–2011	SEER 9, Los Angeles, San Jose-Monterey
SEER 13	1992–2011	SEER 11, Rural Georgia, Alaska Native Tumor Registry
SEER 17	2000–2011	SEER 13, Greater California, Kentucky, Louisiana, New Jersey
SEER 18	2000–2011	SEER 17, Greater Georgia

Table 3.2. Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) codes defining diagnostic imaging technology use.

Imaging Technology	Category	CPT/HCPCS Codes
CT	Brain/Head	70450-70492
	Breast/Chest	72125-72133
	Abdomen	71250-71270, 72192-72194, 74150-74170, 74176-74178
	Bone	73200-73202, 73700-73702
	Body	76380, 76497
PET	Body	78608, 78609, 78810-78816, G0125, G0126, G0165, G0210-G0228, G0231-G0235, G0252-G0254, G0296, G0330, G0331
MRI	Brain/Head	70336, 70540, 70542, 70543, 70551-70553, 70557-70559
	Breast/Chest	71550-71552, 75552, 75553, 75557, 75561, 76093, 76094, 77058, 77059, C8903-C8908
	Abdomen	72195-72197, 74181-74183
	Bone	72141, 72142, 72146-72149, 72156-72158, 73218-73223, 73718-73723
	Body	76498
Bone scan (nuclear medicine)	Bone/Body	76400, 78800-78804, 78102-78104, 78300-78320, 78399, 78999

CHAPTER 4: BIOLOGY AND ETIOLOGY OF YOUNG-ONSET BREAST CANCERS AMONG PREMENOPAUSAL BLACK WOMEN: RESULTS FROM THE AMBER CONSORTIUM

4.1 Introduction

Black women have increased rates of breast cancer with greater prevalence of advanced stage, larger size, higher grade, hormone receptor negative, and basal-like disease compared to white women(8, 21, 22, 48-53, 105). Similar tumor biology is also evident in young-onset (<40 years) breast cancers, which are more common among black women(22, 23, 26, 27). Differences in risk factor profiles for young and older women may reflect distinct etiologies for breast cancers arising in young and black women. Risk factors such as parity, age at first birth, oral contraceptive use, and obesity have been shown to differentially affect the risk of breast cancer according to age at diagnosis(17, 19, 34, 35, 43, 45-47, 88). These same risk factors are also differentially associated with hormone receptor positive and negative disease(17-19, 106, 107), which may confound observed age-related patterns. However, population-based studies examining whether risk factor associations vary by age at diagnosis among black women are rare and have been hampered by small sample sizes, overall and by age(17, 59, 60). Furthermore, previous studies of young women's breast cancer have used inconsistent definitions of young age, defining young with varying age cutpoints or confounding age and menopausal status(17, 34-36, 38, 59, 60, 108, 109), complicating comparisons across studies.

The present study investigated risk factors for young black women's breast cancer in the African American Breast Cancer Epidemiology and Risk (AMBER) Consortium, a large collaboration of breast cancer studies among black women with extensive clinical, molecular, and epidemiologic data. We restricted our analysis to premenopausal women, as previous work has suggested that age and menopausal status may have independent roles in young women's breast cancer(36, 88). Our objectives were two-fold: first, to characterize the biology of breast cancers

diagnosed among young and older premenopausal black women in the AMBER Consortium, and second, to identify epidemiologic risk factors associated with premenopausal young- vs. older-onset breast cancers overall and by estrogen receptor (ER) status. We hypothesized that more aggressive breast tumor characteristics and distinct patterns of breast cancer risk factors would be associated with young-onset breast cancers (<40 years), and ER status would modify observed risk factor associations by age at diagnosis.

4.2 Methods

4.2.1 Study population

The African American Breast Cancer Epidemiology and Risk (AMBER) Consortium is a collaboration of four of the largest epidemiologic studies of breast cancer among black women(76). Included are two case-control studies, the Carolina Breast Cancer Study (CBCS) (17, 58, 77) and Women's Circle of Health Study (WCHS)(81), as well as two prospective cohort studies, the Black Women's Health Study (BWHS)(79) and the Multiethnic Cohort Study (MEC)(110). The AMBER Consortium and participating studies have been described in detail previously(76). Briefly, the CBCS recruited breast cancer cases and controls aged 20-74 years across 24-44 North Carolina counties in three phases (Phase 1: 1993-1996, Phase II: 1996-2001, and Phase III (cases only): 2008-2013). The WCHS recruited cases and controls aged 20-75 years in New York (2002-2008) and New Jersey (2006-present). The BWHS enrolled participants aged 21-69 years from 17 continental states in 1995 with biennial follow-up to record changes to exposure history, incident disease, and mortality. The MEC recruited participants aged 45-69 years in Hawaii and southern California from 1993-1996, with 5-year follow-up questionnaires. Cohort studies (BWHS and MEC) provided nested case-control data to the Consortium comprised of all incident breast cancer cases and up to four matched controls for each case(76).

The present study included premenopausal black cases diagnosed with invasive breast cancer and matched premenopausal controls from the CBCS (701 cases, 298 controls), WCHS (569 cases, 565 controls), and BWHS (738 cases, 4,281 controls), for a total study population of 7,152 women

(2,008 cases, 5,144 controls). All postmenopausal women (defined based on self-reported cessation of menstruation, bilateral oophorectomy, or ovary irradiation), and women with unknown menopausal status were excluded to estimate age effects independent of menopausal status. Additionally, the MEC was excluded from this analysis due to their exclusion criteria limiting to women ≥45 years of age. Each study and the AMBER Consortium collaboration were approved by Institutional Review Boards at participating institutions, and all participants gave written informed consent.

4.2.2 Data collection

The collection of tumor characteristic and risk factor exposure data in the AMBER Consortium has been described previously(76, 89). Briefly, each study contributed paraffin-embedded breast tumor tissue for all cases to two core research facilities (the Translational Pathology Lab (TPL) at the University of North Carolina at Chapel Hill (UNC) for the CBCS and the Roswell Park Cancer Institute for the WCHS and BWHS) where tissue microarrays (TMAs) were constructed for available tumor specimens. Immunohistochemistry (IHC) assays were conducted on all TMAs at UNC's TPL to define expression of estrogen and progesterone receptors (ER/PR) and human epidermal growth factor receptor 2 (HER2)(89). Positive expression was defined as ≥1% staining for ER and PR, and ≥10% staining at the 3+ level for HER2 consistent with previous work(89). Breast cancer subtype was defined as four groups based on positivity of three IHC markers: luminal A (ER+ or PR+, HER2-), luminal B (ER+ or PR+, HER2+), HER2+/ER- (ER-, PR-, HER2+), and triple-negative (ER-, PR-, HER2-). For cases with missing IHC-based tumor characteristics, ER, PR, HER2, and subtype data were defined from medical records; cases with both IHC-based and clinical hormone receptor data showed high agreement for the two measures (κ statistic range=0.68-0.76, concordance range=88-91%). Tumor grade was centrally reviewed by a study pathologist for 44% of cases, with grade data obtained from medical records for remaining cases (κ statistic=0.95, concordance=96% for both grade measures). Other tumor characteristics (including stage (1-4), lymph node status (positive vs. negative), and estimated tumor size (<2, 2-4.9, ≥5cm)) were acquired from medical records.

Risk factor exposure data for cases and controls were obtained via in-home interviews by study staff (CBCS and WCHS) or mailed questionnaire (BWHS), as described previously(76). Participants were asked questions regarding their medical and family histories as well as biologic, anthropomorphic, reproductive, and lifestyle exposures. For CBCS and WCHS, interviewers also measured body weight, height, and waist and hip circumferences during home interviews; for the BWHS, these measures were self-reported on questionnaires by study participants. Questionnaire and interview data from each study were then harmonized by the AMBER Biostatistics and Data Management core within AMBER to create a central database with consistent exposure definitions across studies. Breast cancer risk factors were categorized as: age at menarche (<13, ≥13 years), parity (nulliparous, 1-2, ≥3 live births), age at first live birth (<25, ≥25 years), age at last live birth (<30, ≥30 years), time since last live birth (<10, ≥10 years), lifetime duration of breastfeeding (never, <3 months, ≥3 months), oral contraceptive use (never, ever), duration (never/<1 year, 1-4 years, 5-9 years, ≥10 years) and recency (never, <10 years, ≥10 years), and first-degree family history of breast cancer (no, yes). Body mass index (BMI) was defined as body weight/height (kg/m^2) using categories from the National Heart, Lung, and Blood Institute (<25 normal/underweight, 25.0-29.9 overweight, and ≥30 obese)(87). Waist-to-hip ratio (WHR) was calculated as the ratio of waist/hip circumference (cm) and categorized in tertiles as <0.77, 0.77-0.83, and ≥0.84, consistent with previous work(88).

4.2.3 Statistical analysis

Case-case and case-control analyses were conducted to identify differences in the associations between tumor characteristics and epidemiologic risk factors and breast cancer by age at diagnosis (<40 vs. ≥40 years) among premenopausal black women (age range 22-59 years). Case-case analyses of tumor characteristics associated with young- vs. older-onset disease included all cases (N=2,008), while case-control analyses of risk factors included all cases and controls except cases from Phase III of the CBCS (total N=6,736), as no matched controls were available for this study phase. Case-control analyses examined risk factor associations for breast cancer among young and older women overall and further stratified by ER status among young women, in which ER-positive

and ER-negative cases were compared separately to all controls. Unconditional logistic regression models were used to estimate adjusted odds ratios (ORs) and 95% confidence intervals (CIs) to assess differences in tumor characteristics and breast cancer risk factors associated with breast cancer by age at diagnosis for all analyses. In case-control analyses, effect measure modification by age was evaluated using likelihood ratio tests in which the estimated log-likelihood of the adjusted model was compared to that of the same model including a multiplicative interaction term for age and the corresponding risk factor. Statistically significant modification was assessed using an α-level of 0.1. Heterogeneity in risk factor associations by ER status among young women was assessed by comparing case-case odds ratios (ORs), with ER status defined as the outcome and each risk factor as the explanatory variable. These case-case ORs represent the ratio of case-control ORs for risk factors associated with ER-positive vs. ER-negative disease, and statistical significance was defined using an α-level of 0.05. Additionally, we conducted sensitivity analyses to examine whether patterns of tumor characteristics and risk factors associated with young- vs. older-onset disease were impacted by the cutpoint used to define young age (40, 45, and 50 years). All models controlled for study, diagnosis year, geographic region, and education status to account for differences between studies. Case-control models additionally adjusted for other risk factors that were identified *a priori* via directed acyclic graphs as potential confounders of each risk factor association. Models for age at first live birth, age at last live birth, time since last birth, and lifetime breastfeeding duration were restricted to parous women. Statistical significance was defined at an α-level of 0.05. All analyses were performed using SAS software, version 9.4 (SAS Institute, Inc., Cary, North Carolina).

4.3 Results

4.3.1 Breast tumor biology varies by age at diagnosis among black women (case-case analyses)

Among premenopausal black women, young age (<40 years) at breast cancer diagnosis was associated with poorer-prognostic tumor characteristics compared to older age at diagnosis (≥40 years) (Table 4.1). Young women were significantly more likely to have higher stage and triple-negative tumors. While not significant, both luminal B and HER2-enriched tumors were associated

with younger age at diagnosis. Young-onset breast cancers were also significantly more likely to be ER and PR negative, with markedly higher grade and larger tumor size. No age associations were observed for lymph node positivity.

4.3.2 Age modifies breast cancer risk factor associations among black women (case-control analyses)

To examine whether breast cancers arising among young and older premenopausal black women are etiologically distinct, we estimated case-control ORs for risk factor associations among premenopausal women stratified by age (Table 4.2). Age at diagnosis most strongly modified associations with first-degree family history of breast cancer, with a three-fold increase in risk among young women that was attenuated among older women(interaction p=0.005). Likelihood ratio tests also showed significant age modification for associations with waist-to-hip ratio (p=0.06) and breastfeeding duration (p=0.1). Higher WHR was more strongly associated with young- compared to older-onset breast cancer, while breastfeeding, regardless of duration, had a reduced OR for young- but not older-onset disease. ORs for BMI were not significantly modified by age, though obese BMI ($\geq$30 kg/m^2) was more strongly associated with a reduced association among young women. Associations with parity were not modified by age, though the association with nulliparity was slightly reduced among young women and parity $\geq$3 births appeared to reduce the odds of disease among older women. Later age at first birth was associated with older-onset but not younger-onset breast cancer, while longer time since last birth appeared to reduce odds of breast cancer for older women. Oral contraceptive use showed similar patterns of association across age groups, with ever and more recent use as well as longer use duration associated with an increased OR among young and older women (p for interaction=0.3 (duration), 0.4 (recency); results not shown). Later age at menarche was not associated with young-onset breast cancer but showed a significantly reduced OR for older-onset breast cancer. In summary, we observed substantial differences in risk factor patterns by age among premenopausal black women, with the strongest differences for family history, body size, and reproductive exposures.

Given the difference in tumor characteristics observed between young and older-onset breast cancer in black women, we examined whether breast cancer biology modified the etiologic patterns we observed, specifically among young women. While some analyses generated ORs that were imprecise, ORs stratified by ER status showed little evidence for differences in etiology according to ER (Table 4.3). Increased odds of young-onset breast cancer associated with higher WHR was limited to ER-negative disease (OR=1.64, 95% CI=0.98, 2.75), conversely, higher BMI was more strongly associated with reduced ER-positive disease (OR=0.61, 95% CI=0.38, 0.98). Additionally, family history of breast cancer was positively associated with young-onset disease regardless of ER status, though the association was stronger for ER-negative disease. However, no statistically significant differences by ER status were observed for these or any other risk factor associations that we examined, suggesting that etiologic associations for young-onset breast cancer are not strongly modified by disease subtype.

4.3.3 Age-dependent risk factor associations are most pronounced with age 40 cutpoint

To examine whether our findings were sensitive to the cutpoint used to define young age, we repeated our analyses of tumor characteristics and risk factors associated with young-onset disease using older cutpoints of 45 and 50 years. We observed the strongest age-related heterogeneity when comparing the youngest women (<40) to women at least 40 years of age. Figure 4.1 shows ORs and 95% CI for risk of young-onset breast cancer defined as <40, <45, and <50 in our cohort for the three risk factors showing the strongest heterogeneity by age: breastfeeding history (ever/never), waist-to-hip ratio (highest/lowest tertile), and family history of breast cancer (yes/no). For all three factors, the associations for young-onset breast cancer were attenuated when defining young women as <45 or <50 at diagnosis.

4.4 Discussion

Using data from one of the largest and most comprehensive study of breast cancer biology and epidemiology among black women to date, the AMBER Consortium, we observed substantial differences in tumor characteristics and some evidence for etiologic heterogeneity of premenopausal

young- and older-onset breast cancers. The etiologic associations that vary by age appear not to be driven by differences in ER status, since few associations among young women were modified by ER status. Furthermore, age-dependent heterogeneity of risk factor associations with breast cancer were greatest when comparing the youngest women (<40) to older (≥40) premenopausal women.

The age-related patterns of tumor characteristics we observed are consistent with previous findings(8, 21-23, 27, 48-53, 57, 88, 105), and our work supports the growing hypothesis that breast cancers diagnosed among young women <40 years are biologically distinct from those diagnosed in older women. It is well-established that black women are more likely to be diagnosed with breast cancer under 40 years of age compared with white women(22, 23, 26, 27), highlighting the importance of identifying prevention strategies for young women's breast cancer, particularly for black women.

Some risk factors for young-onset breast tumors are potentially modifiable. In our study, breastfeeding had a reduced OR for breast cancer in young women while higher WHR was associated with an increased odds of young-onset disease. Both risk factors showed the strongest associations among ER-negative tumors. In contrast, higher BMI showed an inverse association with young-onset disease that was strongest among ER-positive cancers, consistent with previous work(17, 19, 44-47). The observed differences between BMI and WHR underscore these factors as distinct measures of body size and suggest that abdominal adiposity, as represented by WHR, is an important factor contributing to young-onset disease(111-113). Few studies have examined etiologic differences according to age and breast cancer subtype in populations of black women. Millikan et al.(17) and Bertrand et al.(108) also reported that a lack of breastfeeding and higher WHR were significantly positively associated with young-onset and triple-negative (or ER-negative) breast cancers among black women in the CBCS and BWHS, respectively. Other studies in predominately white populations have observed similar associations(36, 88, 109), suggesting that interventions to improve breastfeeding rates and reduce abdominal adiposity may benefit young women of all races. Given that black women tend to breastfeed at lower rates and for shorter durations than white women(114) and

are more likely to have ER-negative disease, breastfeeding-related interventions may be particularly relevant for reducing risk of young-onset disease among black women. Additionally, oral contraceptive use ≥5 years was associated with significantly increased ORs regardless of age, with a stronger association among young women that did not vary according to ER status. Others have shown similarly increased risk with longer and more recent oral contraceptive use for young and AA women(35, 60, 115-117), highlighting that reduced oral contraceptive use may mitigate breast cancer risk within this demographic.

Several exposures associated with young women's breast cancer are not targetable for prevention. Family history of breast cancer showed the greatest heterogeneity according to age in our study, with a markedly higher OR among young women and a moderately elevated OR for older women. Family history often serves as a surrogate for genetic susceptibility for breast cancer, and other work has shown that women diagnosed with breast cancer at an early age have a greater frequency of genetic mutations related to tumorigenesis(118, 119). However, an individual's family history is variable over time and changes with age; older women are more likely to have a positive family history than young women given that breast cancer risk increases with age. Thus, the attenuated risk associations that we observed among older women may be explained by a stronger contribution of environment (relative to germline genetics) in family history of older women. This also underscores that a positive family history in a young woman is a strong marker of familial/genetic risk.

Reproductive exposures have most consistently shown differential patterns with breast cancer risk by age, as young women are more proximal to reproductive years than older women. In contrast to other studies, we did not observe the expected dual risk associations for parity, in which higher parity is associated with increased risk among young women but reduced risk for older women(41-43). We observed suggestions of this relationship in that nulliparity was associated a slightly reduced risk of breast cancer among young women, though no associations with parity were statistically significant. However, other associations between reproductive factors and breast cancer were

consistent with previous work, showing younger age at first live birth and longer time since last birth as protective for older-onset but not young-onset breast cancer (17, 36, 88, 106, 107).

Prior epidemiologic studies of young women's breast cancer have used inconsistent cutpoints to classify young women, ranging from 35-50 years of age(17, 34-36, 38, 59, 60, 88, 108, 109). Many studies have also included limited representation of young women, as women <40 years of age represent less than 7% of all breast cancers diagnosed in the United States(1). As such, conclusions regarding whether young- and older-onset breast cancers have distinct etiologies have been mixed, and different studies have yielded varied directions and magnitudes of associations for many risk factors. However, reproductive (particularly parity, breastfeeding history, and age at first birth) and body size exposures have consistently shown the strongest differences in patterns of association for young and older women. We identified that varying the age cutpoint in our study population from 40 to 50 years resulted in attenuated effect estimates with increased age for many risk factor associations. Additionally, dichotomizing our cohort at age 40 enabled a comparison of younger and older premenopausal women, as we previously showed that age and menopausal status are best considered as separate factors in studies of young women's breast cancer(88). Taken together, our findings suggest that age-dependent heterogeneity in risk factor associations are most pronounced when classifying young women as <40 years.

Our results should be interpreted in light of some limitations. Differences in breast cancer screening rates and/or adherence between young and older black women may have influenced some tumor characteristics among younger women, although screening data were unavailable in the Consortium. Breast cancers detected via screening tend to have more favorable tumor characteristics than self- or clinically-detected tumors(3, 31). However, interval cancers, or those diagnosed between regular screening intervals, are more likely to be aggressive and may be present regardless of screening(120). While screening differences may contribute to differences in observed tumor characteristics, screening is unlikely to have influenced the etiologic associations we described by age and ER status.

4.5 Conclusions

In summary, we found strong evidence that breast cancers diagnosed among young black women have tumor characteristics suggestive of poorer prognosis, underscoring the need for greater understanding of the etiology of young-onset disease. In the largest epidemiologic study of young black women's breast cancer to date, our findings suggest that potentially modifiable risk factors, such as breastfeeding and adiposity, are associated with young-onset breast cancer, in addition to other non-modifiable factors such as family and reproductive history.

Table 4.1. Case-case ORs of tumor characteristics by age among premenopausal cases in the AMBER Consortium (N=2,008).

	≥40 years (ref)	<40 years	
Tumor characteristic	N (%)	N (%)	OR (95% CI)[a]
Mean age (±SD)	46.2 (±4.3)	34.7 (±3.8)	
Stage			
Stage I	450 (35.8)	116 (24.9)	1.0
Stage II	572 (45.5)	264 (56.8)	1.81 (1.34, 2.45)
Stage III/IV	234 (18.6)	85 (18.3)	1.32 (0.90, 1.94)
Missing	219	68	
Subtype (IHC/clinically defined)			
Luminal A	492 (51.4)	157 (43.5)	1.0
Luminal B	150 (15.7)	49 (13.6)	1.37 (0.86, 2.17)
HER2	65 (6.8)	29 (8.3)	1.35 (0.77, 2.37)
Triple-negative	250 (26.1)	126 (34.9)	1.56 (1.09, 2.21)
Missing	518	172	
ER status			
Positive	768 (62.1)	227 (52.1)	1.0
Negative	469 (37.9)	209 (47.9)	1.35 (1.03, 1.77)
Missing	238	97	
PR status			
Positive	699 (56.8)	199 (46.2)	1.0
Negative	531 (43.2)	232 (53.8)	1.57 (1.20, 2.06)
Missing	245	102	
HER2 status			
Negative	752 (77.5)	286 (78.1)	1.0
Positive	218 (22.5)	80 (21.9)	1.15 (0.81, 1.65)
Missing	505	167	
Histologic grade			
Low	144 (12.7)	33 (8.4)	1.0
Moderate	373 (33.0)	117 (29.8)	2.00 (1.11, 3.62)
High	613 (54.2)	243 (61.8)	2.17 (1.23, 3.85)
Missing	345	140	
Node status			
Negative	486 (53.8)	175 (51.2)	1.0
Positive	418 (46.2)	167 (48.8)	1.13 (0.87, 1.47)
Missing	571	191	
Tumor size			
<2 cm	477 (40.2)	131 (30.6)	1.0
2-4.9 cm	487 (41.0)	206 (48.1)	1.70 (1.28, 2.26)
≥5 cm	223 (18.8)	91 (21.3)	1.31 (0.83, 2.07)
Missing	288	105	

[a]Adjusted for study site, index year, geographic region, and education status.
[b]Missing data due to ongoing data collection.

Table 4.2. Case-control ORs of breast cancer risk factors by age among premenopausal women in the AMBER Consortium (N=6,736).

	<40 years (N=1,775)			≥40 years (N=4,961)			Test for heterogeneity
Risk factor	Controls N (%)	Cases N (%)	OR (95% CI)[a]	Controls N (%)	Cases N (%)	OR (95% CI)[a]	X^2, df (p)[b]
BMI (kg/m^2)							
<25.0	466 (35.0)	149 (35.1)	1.0	955 (25.4)	283 (24.6)	1.0	0.41, 2 (0.8)
25-29.9	368 (27.7)	119 (28.1)	0.92 (0.66, 1.28)	1,172 (31.2)	376 (32.7)	0.99 (0.81, 1.21)	
≥30.0	497 (37.3)	156 (36.8)	0.71 (0.51, 0.98)	1,628 (43.4)	491 (42.7)	0.91 (0.75, 1.10)	
Trend test			p=0.0005			p=0.2	
Missing	16	4		42	14		
WHR							
<0.77	441 (37.5)	104 (26.3)	1.0	1,166 (34.3)	1,166 (34.3)	1.0	5.70, 2 (0.06)
0.77-0.83	341 (29.0)	128 (32.4)	1.14 (0.81, 1.59)	952 (28.0)	952 (28.0)	1.02 (0.83, 1.25)	
≥0.84	394 (33.5)	163 (41.3)	1.46 (1.04, 2.05)	1,285 (37.8)	1,285 (37.8)	1.11 (0.91, 1.35)	
Trend test			p=0.003			p=0.9	
Missing	171	33		394	77		
Age at menarche							
<13 years	803 (59.8)	251 (58.6)	1.0	2,038 (53.9)	654 (56.4)	1.0	1.19, 1 (0.3)
≥13 years	540 (40.2)	177 (41.4)	0.97 (0.76, 1.24)	1,746 (46.1)	506 (43.6)	0.85 (0.74, 0.98)	
Missing	4	0		13	4		
Parity							
Nulliparous	516 (38.3)	111 (26.9)	0.90 (0.71, 1.14)	872 (23.0)	222 (19.1)	0.91 (0.75, 1.12)	3.04, 2 (0.2)
1-2 births	644 (47.8)	224 (52.3)	1.0	2,081 (54.8)	654 (56.2)	1.0	
≥3 births	187 (13.9)	93 (21.7)	1.06 (0.87, 1.29)	844 (22.2)	288 (24.7)	0.88 (0.73, 1.06)	
Age at first live birth[c]							
<25 years	481 (59.0)	202 (63.9)	1.0	1,633 (56.9)	547 (58.8)	1.0	1.23, 1 (0.3)
≥25 years	334 (41.0)	114 (36.1)	1.03 (0.74, 1.43)	1,238 (43.1)	383 (41.2)	1.18 (0.99, 1.41)	
Missing	16	0		54	12		
Time since last birth[c]							
<10 years	515 (63.6)	202 (64.1)	1.0	521 (18.3)	183 (19.7)	1.0	1.23, 1 (0.3)
≥10 years	295 (36.4)	113 (35.9)	1.14 (0.82, 1.57)	2,323 (81.7)	744 (80.3)	0.86 (0.69, 1.07)	
Missing	21	2		81	15		
Breastfeeding duration[c]							
Parous, never	388 (47.7)	177 (56.2)	1.0	1,444 (50.3)	491 (52.7)	1.0	4.04, 2 (0.1)
<3 months	114 (14.0)	30 (9.5)	0.70 (0.43, 1.16)	315 (11.0)	96 (10.3)	1.08 (0.82, 1.42)	
≥3 months	312 (38.3)	108 (34.3)	0.83 (0.58, 1.17)	1,111 (38.7)	344 (37.0)	0.94 (0.78, 1.13)	
Missing	17	2		55	11		

Oral contraceptive use							
Never	177 (13.2)	74 (17.3)	1.0	539 (14.2)	232 (20.0)	1.0	0.13, 1 (0.7)
Ever	1,167 (86.8)	354 (82.7)	1.22 (0.86, 1.72)	3,254 (85.8)	929 (80.0)	1.18 (0.97, 1.44)	
Missing	3	0		4	3		
Family history of breast cancer							
No	1,276 (94.7)	363 (84.8)	1.0	3,472 (91.4)	996 (85.6)	1.0	7.7, 1 (0.005)
Yes	71 (5.3)	65 (15.2)	3.10 (2.08, 4.63)	325 (8.6)	168 (14.4)	1.57 (1.26, 1.94)	

[a]Adjusted for age, study site, index year, geographic location, education level, and confounders, by model. BMI: WHR, parity; WHR: BMI, parity; parity: age at first live birth; age at last live birth: parity, age at first birth; time since last birth: parity, age at first live birth, age at last live birth; breastfeeding duration: BMI, parity, age at first live birth, age at last live birth; oral contraceptive use: parity, age at first live birth, age at last live birth.

[b]Likelihood ratio tests assessed age-related heterogeneity in risk factor associations by comparing the estimated log-likelihood of adjusted models to that of the adjusted model including a multiplicative interaction term for age and the corresponding risk factor (e.g., BMI*age). Statistically significant heterogeneity by age was defined with $\alpha=0.1$.

[c]Among parous women.

Table 4.3. Case-control ORs of breast cancer risk factors by ER status among young (<40 years) premenopausal women in the AMBER Consortium (N=1,678).

	Controls	ER+		ER-		
Risk factor	N (%)	N (%)	OR (95% CI)[a]	N (%)	OR (95% CI)[a]	*p* for heterogeneity[b]
BMI (kg/m^2)						
<25.0	466 (35.0)	62 (38.3)	1.0	52 (31.0)	1.0	
25-29.9	368 (27.6)	41 (25.3)	0.80 (0.49, 1.29)	49 (20.2)	0.92 (0.56, 1.49)	0.6
≥30.0	497 (37.3)	59 (36.4)	0.61 (0.38, 0.98)	67 (39.9)	0.87 (0.54, 1.40)	0.2
Trend test			*p* = 0.0003		*p* = 0.3	
Missing	16	0		1		
WHR						
<0.77	441 (37.5)	41 (27.0)	1.0	38 (23.6)	1.0	
0.77-0.83	341 (29.0)	48 (31.6)	0.94 (0.57, 1.55)	62 (38.5)	1.56 (0.95, 2.55)	0.1
≥0.84	394 (33.5)	63 (41.4)	1.15 (0.70, 1.89)	61 (37.9)	1.64 (0.98, 2.75)	0.3
Missing	171			8		
Age at menarche						
<13 years	803 (59.8)	90 (55.6)	1.0	102 (60.4)	1.0	
≥13 years	540 (40.2)	72 (44.4)	1.08 (0.76, 1.55)	67 (39.6)	0.86 (0.60, 1.24)	0.1
Missing	4	0		0		
Parity						
Nulliparous	516 (38.3)	42 (25.9)	0.93 (0.56, 1.54)	38 (22.5)	0.79 (0.47, 1.32)	0.7
1-2 births	644 (47.8)	85 (52.5)	1.0	90 (53.3)	1.0	
≥3 births	187 (13.9)	35 (21.6)	0.84 (0.50 1.40)	41 (24.3)	1.04 (0.64, 1.71)	0.9
Breastfeeding duration[c]						
Parous, never	388 (47.7)	59 (49.6)	1.0	79 (60.8)	1.0	
<3 month	114 (14.0)	11 (9.2)	0.84 (0.40, 1.80)	10 (7.7)	0.55 (0.25, 1.19)	0.5
≥3 months	312 (38.3)	49 (41.2)	1.14 (0.68, 1.92)	41 (31.5)	0.88 (0.53, 1.48)	0.7
Missing	17	1		1		
Oral contraceptive use						
Never	177 (13.2)	30 (18.5)	1.0	26 (15.4)	1.0	
Ever	1,167 (86.8)	132 (81.5)	1.47 (0.90, 2.40)	143 (84.6)	1.41 (0.84, 2.37)	1.0
Missing	3	0		0		
Family history of breast cancer						
No	1,276 (94.7)	140 (86.4)	1.0	141 (83.4)	1.0	
Yes	71 (5.3)	22 (13.6)	2.63 (1.47, 4.70)	28 (16.6)	3.32 (1.92, 5.75)	0.4

[a]Adjusted for age, study site, index year, geographic location, education level, and confounders, by model. BMI: WHR, parity; WHR: BMI, parity; parity: age at first live birth; age at last live birth: parity, age at first birth; time since last birth: parity, age at first live birth, age at last live birth; breastfeeding duration: BMI, parity, age at first live birth, age at last live birth; oral contraceptive use: parity, age at first live birth, age at last live birth.
[b]Statistically significant heterogeneity by age was defined with $\alpha=0.05$.
[c]Among parous women.

Figure 4.1. Impact of age cutpoints on risk factor analyses.

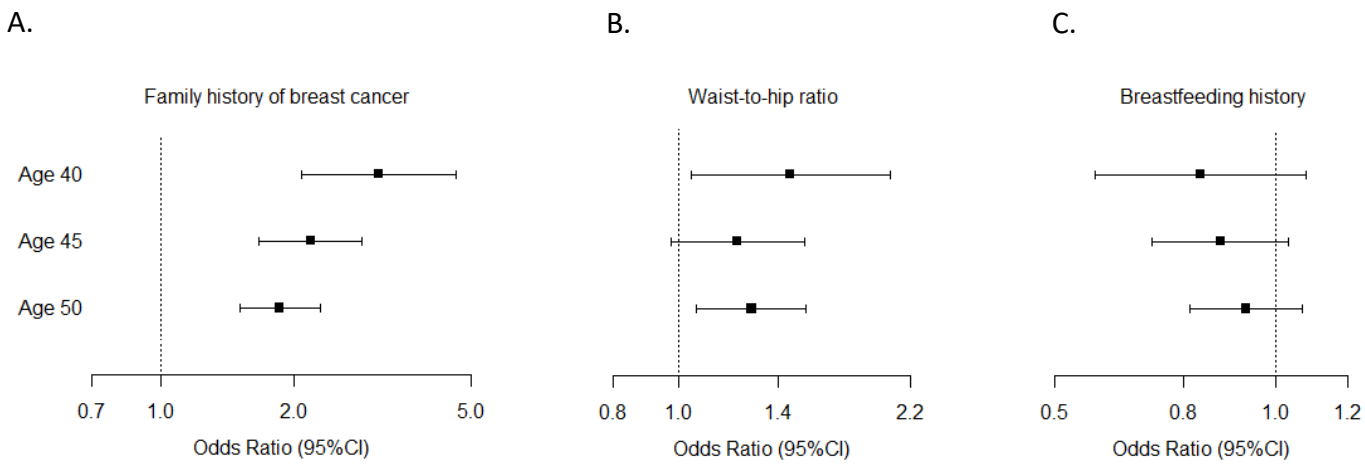


Figure 4.1 legend. Case-control ORs for associations between family history of breast cancer (yes/no; panel A), waist-to-hip ratio (highest/lowest tertile; panel B), and breastfeeding history (ever/never; panel C) and premenopausal young-onset breast cancer. Cutpoints defining “young” varied at <40, <45, or <50 years of age. Error bars represent 95% CIs.

CHAPTER 5: INCREASED DISTANT STAGE BREAST CANCER AMONG YOUNG WOMEN OVER TIME: ETIOLOGY OR TECHNOLOGY?

5.1 Introduction

Recent epidemiologic studies have reported a rise in the incidence of distant metastatic (stage IV) breast cancer among young women <40 years of age in the United States(4, 21). One study conducted using data from the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) program estimated that rates of distant disease nearly doubled among young women from 1976-2009, with the sharpest increases in incidence occurring since 2000(4). Given the poor prognosis associated with distant disease, with a five-year survival rate of only 25%(63), this finding has prompted concern regarding possible causes for the increasing trend. Furthermore, it is well-established that young women less than 40 years have a higher prevalence of aggressive breast cancers compared to older women, including hormone receptor negative, high grade, and larger tumors(22-27, 36, 88). The potential impacts of rising distant disease on the burden of breast cancer in this at-risk demographic are unclear, underscoring a need for longitudinal studies investigating the underlying reasons for shifting incidence among young women.

One hypothesis for the rise in young-onset distant disease suggests a shift to more aggressive breast cancers over time. Breast cancers with poorer prognostic tumor features are faster to metastasize, and a greater representation of aggressive breast cancers among young women could contribute to increasing distant disease in young women. However, recent work conducted in SEER suggests that more favorable tumor characteristics (e.g., estrogen receptor (ER) positivity, low grade, and smaller tumor size) have increased over time among women overall, while the incidence of more aggressive tumor features has potentially declined(21, 64, 121). However, few studies have examined patterns among young women less than 40 years. Furthermore, the presence of substantial missing

data for some tumor characteristics in SEER has complicated the interpretation of longitudinal biological trends(102, 103).

As an alternative to the hypothesized biological shifts, stage migration may have contributed to increases in distant disease. Advances in diagnostic imaging technology may have led to increasing rates of distant breast cancer at diagnosis among young women. For example, studies of lung cancer showed that increased use of imaging technologies (e.g., positron emission tomography (PET), computed tomography (CT), and magnetic resonance imaging (MRI) technologies) identified previously undetectable distant metastases. It would be misleading to interpret stage-specific incidence trends over time as evidence of an etiologic or biological shift, as technology drives changing stage without an underlying change to the actual state of disease(66-69). Studies considering the impact of stage migration on breast cancer incidence patterns have been rare and yielded inconsistent results(4, 122), and the lack of longitudinal administrative healthcare claims data for population-based cancer studies has hampered studies of imaging use patterns.

The present study aimed to evaluate both breast cancer biology and stage migration as factors underlying recent temporal patterns for distant disease in young women (<40 years of age). Using data from the NCI's SEER program, we assessed stage-specific incidence patterns and temporal trends in tumor biological factors associated with young-onset, distant breast cancer from 1992-2011. Additionally, we characterized the extent of missing data over this 20-year time period and explored the impact of missingness on the interpretation of longitudinal incidence trends. To assess changes in diagnostic patterns over the same interval, we used data from the SEER-Medicare linked program to evaluate patterns in the use of imaging technologies associated with breast cancer staging (i.e., PET, CT, MRI, and nuclear medicine bone scans). We hypothesized that these data sources could help to identify the most plausible causes of temporal shifts, including both breast tumor biology and stage migration.

5.2 Methods

5.2.1 Study populations

Our study included female breast cancer cases diagnosed from 1992-2011 in both the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) 13 program(123) and the SEER-Medicare linked database(86). The SEER 13 program comprises cancer registries in Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, Utah, Los Angeles, San Jose-Monterey, Rural Georgia, and the Alaska Native Tumor Registry that have been active since 1992, together representing approximately 15% of the U.S. population. The SEER-Medicare program links SEER cancer incidence data for Medicare-eligible cancer cases to healthcare utilization claims data available in Medicare, including those for diagnostic imaging technology use, through a collaboration between the NCI and the Centers for Medicare and Medicaid Services. We selected the 20-year study period (1992-2011) based on the initiation of the SEER-Medicare program in 1991 and the addition of four racially-diverse SEER registries in 1992 (Los Angeles, San Jose-Monterey, Rural Georgia, and the Alaska Native Tumor Registry) to form the SEER 13 program.

Women residing in the U.S. women aged 20-39 years (SEER) or 65-108 years (SEER-Medicare) were eligible for inclusion in our study if they had a first primary breast cancer diagnosis (ICD-O-3 cancer sites C50.0-50.9) between January 1, 1992 and December 31, 2011 that was reported to a SEER 13 cancer registry. Women with a personal history of cancer prior to the first primary breast cancer diagnosis or a primary breast cancer diagnosed at time of death or autopsy were excluded. To identify imaging use associated with a breast cancer diagnosis, we additionally required continuous fee-for-service Medicare Parts A and B enrollment and no health maintenance organization enrollment (to minimize incomplete claims data) for the two months prior and four months following the month of cancer diagnosis for SEER-Medicare cases. Our final study populations included 30,407 young women <40 years of age within the SEER program and 142,051

women ≥65 years of age in the SEER-Medicare program. This study was approved by the University of North Carolina (UNC) School of Medicine Institutional Review Board.

5.2.2 Data collection

The collection of tumor characteristic data in SEER has been described previously(82, 85). Briefly, stage at diagnosis, estrogen and progesterone receptor (ER/PR) status, tumor grade, and primary tumor size data were collected from pathology reports or medical records. Stage at diagnosis was defined using SEER's historic stage A coding of four summary stage categories, as this coding has remained consistent in SEER throughout the study period, including *in situ* (noninvasive cancer), local (invasive cancer confined to the breast), regional (tumor extension to breast skin, chest wall, and/or regional lymph nodes), and distant (metastasis to non-breast tissues). ER/PR status was recorded as positive, borderline, negative, and unknown. Given that clinical standards defining hormone receptor positivity changed over the study period, varying from 1-10%, we combined positive and borderline cases to approximate more recent guidelines of 1% positivity recommended by the American Society of Clinical Oncology/College of American Pathologists(90). Tumor grade was defined at time of diagnosis by a pathologist as low (well-differentiated), moderate (moderately-differentiated), high (poorly-differentiated) and undifferentiated (anaplastic). Undifferentiated tumors (defined as grade 4) were uncommon (approximately 3% of cases) and were excluded from all grade analyses. Primary tumor size was defined as the single largest tumor dimension prior to any neoadjuvant cancer therapy and dichotomized as ≤2cm vs. >2cm.

We used SEER-Medicare linked data and defined diagnostic imaging technology use as at least one claim in a given calendar year for an imaging technology used to diagnose breast cancer metastases at the time of diagnosis, including CT, PET, MRI, and bone scans. Use of these technologies was defined using relevant Healthcare Common Procedure Coding System (HCPCS) and Current Procedural Terminology (CPT) imaging codes corresponding to these technologies from 1992-2011 (Table 5.5). Breast cancer cases were considered "imaged" if at least one of these four imaging scans was received between 2 months prior and 4 months following the primary diagnosis

date within SEER. This exposure window was selected to identify imaging used incidentally prior to cancer diagnosis and to allow sufficient time for complete disease staging post-diagnosis(68, 98).

5.2.3 Statistical analysis

To examine trends in distant breast cancer among young women <40 years of age during our 20-year study period, we used SEER*Stat software (version 8.3.2(99)) to calculate annual stage-specific breast cancer incidence rates from 1992-2011 in the SEER 13 program, adjusted for age using the 2000 U.S. standard population. Incidence rates were calculated for young women overall and stratified by black and white race to identify whether temporal trends in stage at diagnosis vary according to race. Joinpoint regression(72) (version 4.4.0.0(100)) was used to quantitatively assess temporal trends by fitting up to three joinpoints at statistically significant points at which the slope of the incidence trends changed. Annual percent change (APC) estimates were generated for each trend segment, and weighted average APCs (AAPCs) were calculated to generate a single percent estimate for 1992-2011 and to quantitatively compare trends across analyses(101).

To examine the "biological shift" hypothesis, we calculated annual stage-specific incidence rates and annual incidence rates according to ER/PR status, tumor grade, and primary tumor size, overall and among cases with distant stage disease at diagnosis. All incidence rates were adjusted for age using the 2000 U.S. standard population, and analyses of tumor characteristics were restricted to invasive breast cancer cases (N=26,870). We then used joinpoint regression to calculate AAPCs estimating temporal trends in stage and breast tumor characteristics from 1992-2011 among young women overall and according to race (black vs. white). Substantial missing data has been documented for tumor characteristics reported to SEER, particularly for ER and PR status (approximately 25% in 1992), which may bias longitudinal studies of incidence rates(102, 103). To clarify whether data are missing at random across characteristics, we characterized patients with missing data by estimating descriptive statistics with chi-square significance testing for cases with and without missing data, and used joinpoint regression to examine whether the distribution of breast cancer cases (as measured by percent of cases) varied over time with respect to each tumor characteristic.

To investigate the "stage migration" hypothesis, annual rates of imaging use were calculated overall and by imaging type (CT, PET, MRI, and bone scans) among older breast cancer cases in SEER-Medicare from 1992-2011. Joinpoint regression was used to estimate AAPCs for all longitudinal trends. Sensitivity analyses were conducted excluding elderly women (over age 75) as well as varying the 6-month imaging window to 3 months post-diagnosis; however, these exclusions did not substantially impact results beyond reductions in study power and therefore were not employed in final analyses (results not shown). Unconditional logistic regression was used to estimate age-adjusted odds ratios (ORs) and 95% confidence intervals (CIs) to assess the associations between imaging use and tumor biological characteristics. Logistic regression analyses were performed using SAS software, version 9.4 (SAS Institute, Inc., Cary, North Carolina). Statistical significance for all analyses was defined at an α-level of 0.05.

5.3 Results

5.3.1 Incidence of distant and unstaged breast cancer among young women

In the SEER 13 registries, young (<40 years) breast cancer cases diagnosed from 1992-2011 tended to have local or regional stage at diagnosis (11.6% *in situ*, 42.7% local, 37.6% regional, 6.1% distant, and 2.0% unstaged). Over the 20-year study period, joinpoint regression models revealed little change in the rates of *in situ*, local, and regional stage breast cancer (Figure 5.1). However, distant disease incidence significantly increased by an average of 3.2% annually while the rate of unstaged disease decreased dramatically by 6.8% per year. Comparing trends by race revealed that incidence of *in situ* disease increased for black but not white women (p=0.03; black AAPC=3.0%, 95% CI=0.5, 5.7; white AAPC=0.7, 95% CI=-0.0, 1.5) while regional stage disease showed a significant increase for white but not black women over time (p=0.005; white AAPC=1.1, 95% CI=0.6, 1.5; black AAPC=-0.4, 95% CI=-1.1, 0.3). Distant stage breast cancer increased similarly across race (p=0.5; black AAPC=3.8, 95% CI=2.0, 5.5; white AAPC=3.4, 95% CI=2.5, 4.2) while unstaged disease showed marked declines over time for both races (p=0.4; black AAPC=-8.1, 95% CI=-10.7, -5.5; white AAPC=-5.8, 95% CI=-8.5, -3.1). Local disease incidence did not change significantly

according to race (p=0.06; black AAPC=-0.3, 95% CI=-1.1, 0.4; white AAPC=0.1, 95% CI=-0.3, 0.4).

5.3.2 Temporal incidence trends suggest increasing frequency of favorable tumor biological characteristics

To evaluate the "biological shift" hypothesis, we examined temporal trends in breast cancer incidence according to tumor characteristics and found that breast cancers among young women shifted toward more favorable prognostic features over time (Table 1). Rates of ER and PR positive disease increased significantly by 3.2% and 2.5% per year, respectively, while the incidence of negative hormone receptor status showed a significant decrease (ER status) or small increase (PR status) over time. Similarly, the rate of low grade tumors strongly increased over the study period compared to attenuated positive trends for moderate and high grade disease. Only tumor size showed a shift to less favorable disease features, with incidence of larger tumors >2cm significantly increasing over time by 1.7% per year and no change in smaller tumors. However, ascertainment of all tumor characteristics dramatically improved from 1992-2011, as the rate of missing data for each characteristic decreased with AAPCs of considerably higher magnitude than any of the changes in nonmissing disease categories. Unknown ER and PR status showed the strongest declines, with AAPCs of over 10% over the study period, while rates of unknown grade and tumor size status also decreased substantially by an average of 8.7% and 4.2% per year, respectively. Stratification by race did not reveal differences in incidence trends between white and black women for any of the tumor characteristics, except tumor size which showed a slight but significantly stronger increase in the rate of larger tumors for white compared to black women (p=0.05; white AAPC=1.9, 95% CI=1.5, 2.4; black AAPC=1.1, 95% CI=0.3, 1.8) (results not shown).

We investigated whether observed incidence trends for tumor characteristics differed when restricting to young women with distant disease. As shown in Table 1, strong increases in the incidence rates of both favorable and poor prognostic tumor features were evident for ER and PR status as well as grade, though the rates of positive change for favorable characteristics (ER positive,

PR positive, and low/moderate grade) were substantially larger than for less favorable features. However, the rates of unknown tumor characteristics showed significant declines for all three factors, particularly for ER and PR status which showed negative AAPCs of 9.5% and 11.2%, respectively, while unknown grade declined at a more modest but significant rate of 2.6% per year. In contrast, patterns for tumor size among young women with distant disease showed no temporal changes for smaller tumors ≤2cm or unknown tumor status, although larger tumors >2cm significantly increased by an average of 3.3% annually.

5.3.3 Missing data contribute to biased incidence trends among young women

The dramatic declines in the rates of missing data over time for stage at diagnosis and all tumor characteristics prompted an investigation into the level of potential bias that missing data contributes to observed temporal incidence trends. Women with unstaged breast cancer tended to be black race (p=0.01) and very slightly younger (p<0.0001) than women who were staged at diagnosis (Table 2). Additionally, unstaged women were significantly more likely to have larger tumors (p=0.004) with slightly increased ER positive disease (p=0.08), and were also highly likely to have missing data for other tumor characteristics as well (p<0.0001). However, we examined factors associated with unstaged disease in the earliest (1992-1996) and latest (2007-2011) five years of our study period to assess differences in missingness across time, and found that race was the only factor that significantly differed by stage status for the earliest cases (p=0.003). In contrast, the most recent unstaged cases were more likely to differ based on disease characteristics, including increased ER negative (p=0.003) and low grade disease (p=0.02). With regard to other tumor characteristics, mean age and race were not associated with unknown status for any factors other than tumor size, for which black women were significantly more likely to have missing data (p=0.04; results not shown).

To evaluate how improving ascertainment of tumor characteristics over time impacted temporal incidence trends, we excluded cases with missing data and examined the frequency distribution of young breast cancer cases over time with respect to each tumor characteristic. We found that the percent of breast cancer cases with more favorable tumor characteristics (specifically

ER/PR positive and low grade disease) increased significantly from 1992-2011, while the percent of cases with less favorable characteristics (ER/PR negative and high grade disease) decreased significantly over time (Table 3). Only tumor size showed an increase in less favorable disease over the study period, in that the percent of cases with larger tumor size >2cm showed a small but significant increase in frequency of 0.9% per year while smaller tumors declined significantly by 1.4% per year. Comparing temporal trends in the incidence rate (Table 1) and frequency (Table 3) showed that temporal trends were more pronounced when evaluating rates, and were attenuated or even reversed when considering frequencies.

These analyses were conducted among all women, so we also examined changes among women with distant disease (Table 3). We observed a significant decrease in the percent of cases with ER negative and high grade disease while the percent of cases with ER positive and low/moderate disease significantly increased. PR positive status appeared to increase, though not significantly. The distribution of small and larger tumor sizes did not significantly change over the study period. Figure 5.2 illustrates incidence rates (panel A) and the distribution of ER status (panel B) among cases with distant disease over time, showing equivalence at approximately 50% for both ER positive and negative disease until 2001 when ER negative disease began to decrease, coinciding with lower rates of cases with unknown ER status. In comparison with Figure 5.2A, ER negative breast cancer showed slightly increasing rates over time that reversed to declines when frequency was assessed (Figure 5.2B).

5.3.4 Imaging technology use significantly increased over time and in association with tumor characteristics

Evaluating the "stage migration" hypothesis revealed strongly increased use of imaging technology from 1992-2011 in SEER-Medicare (Figure 5.3). The percent of Medicare women receiving any imaging within 2 months prior and 4 months following their breast cancer diagnosis significantly increased by an average of 1.4% (95% CI=0.5, 2.2) per year over the study period, with the strongest increases occurring after 1998. After stratifying by imaging type, the use of PET and

MRI scans dramatically increased over time, with the percent of cases receiving those imaging technologies increasing by 72.5% (95% CI=56.9, 88.6) and 13.7% (95% CI=11.1, 16.3) per year, respectively. The use of CT scans also significantly increased during this time period, though more gradually at 2.9% (95% CI=2.3, 3.6) per year. In contrast, the use of bone scans significantly decreased from 1992-2011 at a rate of 4.9% (95% CI=-5.9, -3.8) per year.

Logistic regression analyses revealed significant associations between imaging use and all tumor characteristics, with less favorable tumor characteristics being associated with increased imaging use (Table 4). A diagnosis of distant stage disease was most strongly associated with imaging use, with distant stage cases being 14.8 times more likely to receive imaging than women with local disease. Imaged cases were also more likely to have ER and PR negative disease as well as high grade and larger tumors. Additionally, imaged cases were less likely to have missing data for stage at diagnosis and all tumor characteristics.

5.4 Discussion

From 1992-2011 in the SEER 13 registries, the incidence of distant stage breast cancer increased by 3.2% per year among young women <40 years, consistent with another recent report(4). Over this time period, we observed a shift to more favorable tumor characteristics and dramatic declines in the rates of missing data, suggesting that the rise in distant disease may not be attributable to increasing aggressive breast tumor biology. Additionally, investigating temporal diagnostic imaging patterns among older women ≥65 years revealed considerable increases in the use of PET, MRI, and CT scans over time, indicating that stage migration via imaging use may have contributed to shifting incidence trends.

While initial analyses of temporal trends in tumor characteristics suggested significant increases in poor prognostic features among young women with distant disease (including ER and PR negativity, high grade, and larger tumor size >2cm), the substantial reductions in the rates of unknown disease status over time confound interpretation. Indeed, among young women with distant disease, the incidence of all nonmissing tumor categories increased over the study period, indicating

that longitudinal incidence rates were impacted by a redistribution of cases as ascertainment of missing tumor characteristics improved over time. We found that the distribution of young distant stage breast cancer cases with known tumor status trended toward more favorable characteristics over time, with the percent of cases having poor prognostic factors reversing the initial positive trend to show significant declines or no change over time for all characteristics. Others have reported similar patterns in the incidence of hormone receptor status, observing increased ER positive disease and decreased or stable ER negative disease for breast cancer cases of all ages, including young women <40 years(21, 64, 124). Johnson et al.(125) identified stronger increases in ER positive disease over time among young women with distant stage disease, but mixed results for ER negative disease in that ER-/PR+ disease significantly declined while ER-/PR- disease showed a small but significant increase over time. Studies investigating temporal trends in tumor grade and size have been less common, with one study identifying rising incidence rates for both low and high grade tumors among young women with a correspondingly large decrease in missing data(21). Others observed increasing incidence of smaller but not larger tumors among women over age 40 years(121), contrasting with our findings among young women. Taken together, these results provide little evidence for the biological shift hypothesis, suggesting that temporal trends in aggressive breast tumor characteristics do not explain the recent rise in distant breast cancer.

Several studies evaluating incidence trends of breast tumor characteristics have used imputation methods(64, 102, 103, 124-126) to address any bias induced by the substantial changes in missing data over time. These authors have demonstrated that approximately 78% of breast cancer cases with missing data were identified as ER positive after imputation, and imputed incidence trends revealed marked declines in the rate of ER negative compared to ER positive disease that were not previously observed in with non-imputed data(102, 103). Our findings support these conclusions, as the distribution of ER status over time has trended toward ER positivity, and underscore the need for careful interpretation of trends in longitudinal studies of breast cancer incidence.

We observed dramatic shifts in the use of imaging technologies over time among in SEER-Medicare, with the strongest increases occurring for PET and MRI scans. Only bone scans showed a decrease in use from 1992-2011, particularly after 2001 as the rates of other imaging technologies increased. These trends coincide with increased Medicare coverage of PET and advanced imaging for breast cancer staging and highlight the shift to PET, MRI, and CT technologies, rather than bone scans, as the standard of care for identifying advanced stage cancers given improved sensitivity for diagnosing distant metastases(127-129). Increased use of PET imaging has been shown to cause stage migration and inflated distant disease trends in studies of lung cancer incidence patterns(66-69), though limited availability of large administrative healthcare databases linked to cancer populations has hampered evaluating patterns in younger breast cancer populations. Another recent study conducted within SEER-Medicare observed increased imaging among women with early stage breast cancer, but did not consider imaging in relation to distant stage disease(130). Other studies have found mixed evidence of stage migration in breast cancer incidence trends among young women, using quantitative comparisons of stage-specific incidence rate trends to assess whether stage shifts have occurred between regional and distant disease over time(4, 122). Evaluating whether survival has improved over time for cases presenting with distant and regional stage disease may clarify the impact of stage migration on the burden of advanced stage breast cancer among young women.

We identified significant associations between imaging use and distant stage as well as aggressive tumor characteristics, suggesting that clinical recommendations for diagnostic imaging technology are strongly related to disease characteristics at diagnosis. Given that young women <40 have a higher frequency of aggressive breast cancers, it is possible that these associations would be stronger in a younger population and prompt greater use of imaging in this demographic. Two recent studies reported that rates of imaging technology use at the time of breast cancer diagnosis were highest among young women(95, 131), with one estimating that imaging use among young women <40 years of age was approximately double the rate among elderly women eligible for Medicare (21% vs. 10%, respectively)(95). These results suggest that our estimates of imaging technology use

among SEER-Medicare cases are likely attenuated compared to those among younger women, and thus our findings may have underestimated the role of imaging use in shaping stage-specific incidence trends among young women over time.

Our results should be considered in light of some limitations. While we sought to clarify the possible impacts of missing data on stage-specific trends and evaluate factors associated with missingness over time, our assessment of missing data was indirect and there may be an unknown pattern of bias. Additionally, our assessment of imaging use in our SEER-Medicare population may have included imaging that was received for reasons unrelated to breast cancer staging, and it is possible that some imaging data were missing from the Medicare claims. However, expensive healthcare services such as the imaging technologies we evaluated are likely to be well-represented in healthcare utilization data sources, and the lengthy staging process for many cases necessitates an imaging assessment period of at least several months(57, 96, 97). Finally, because we used imaging data from older women, we were unable to directly assess the extent of imaging use in the SEER population of young women <40 years, and we cannot directly determine whether stage migration via imaging use has impacted the rise in distant breast cancer in the SEER 13 program. While our use of SEER-Medicare data enabled the identification of imaging patterns in a national population of breast cancer cases, additional linkages of administrative healthcare databases to longitudinal cancer populations with representation of younger women are needed to assess the role of imaging on stage migration in incidence trends among young women.

5.5 Conclusion

In summary, we found little evidence to suggest that breast cancers among young women <40 years have become more aggressive from 1992-2011, and our findings do not support the hypothesis that shifting breast cancer biology has contributed to the rise in distant stage disease among young U.S. women. Shifting patterns in missing data within SEER have likely biased incidence trends according to stage and tumor characteristics, necessitating caution when interpreting potential causes and impacts of breast cancer incidence trends. Our findings suggest that stage migration via increased

diagnostic imaging use may have contributed to rising rates of distant breast cancer among young women.

Table 5.1. Average annual percent change (AAPC) estimates and 95% CIs for incidence of breast tumor characteristics among young women <40 years overall and with distant stage at diagnosis, 1992-2011.

	All stages	Distant
ER status		
Positive	3.2 (2.8, 3.7)	8.2 (7.0, 9.4)
Negative	-0.9 (-1.8, -0.1)	3.8 (2.3, 5.2)
Unknown	-10.6 (-12.0, -9.1)	-9.5 (-12.0, -6.9)
PR status		
Positive	2.5 (2.0, 3.0)	7.5 (6.0, 9.1)
Negative	1.1 (0.3, 1.9)	5.4 (3.8, 7.1)
Unknown	-10.6 (-11.9, -9.3)	-11.2 (-15.0, -7.1)
Grade		
Low	4.2 (2.9, 5.6)	9.5 (7.5, 11.5)[a]
Moderate	2.1 (1.8, 2.5)	
High	1.5 (0.7, 2.3)	3.7 (2.9, 4.4)
Unknown	-8.5 (-10.1, -6.9)	-2.6 (-4.6, -0.5)
Tumor size		
≤2 cm	-0.3 (-1.2, 0.6)	1.5 (-4.2, 7.5)
>2 cm	1.7 (1.3, 2.1)	3.3 (2.4, 4.1)
Unknown	-4.3 (-5.1, -3.4)	1.5 (-0.2, 3.3)

[a]combined low/moderate grade due to small sample size.

Table 5.2. Descriptive statistics for young women <40 years with unstaged disease, overall and for early and more recent diagnosis years, SEER 13 program, 1992-2011.

	All years			1992-1996			2007-2011		
	Staged	Unstaged	X^2 *p*-value	Staged	Unstaged	X^2 *p*-value	Staged	Unstaged	X^2 *p*-value
Mean age (SE)	35.1 (0.02)	34.1 (0.18)	<0.0001[a]						
Race									
White	21,298 (72.0)	409 (70.6)	0.01	5,387 (74.5)	159 (67.7)	0.003	5,049 (69.3)	54 (70.1)	0.6
Black	4,221 (14.3)	105 (18.1)		1,029 (14.2)	52 (22.1)		1,041 (14.3)	13 (16.9)	
Other	4,072 (13.8)	65 (11.2)		818 (11.3)	24 (10.2)		1,198 (16.4)	*b*	
Unknown	202	35		45	16		78	*b*	
ER status									
Positive	14,443 (62.5)	101 (70.0)	0.08	2,936 (56.8)	28 (60.9)	0.6	4,366 (69.5)	*b*	0.003
Negative	8,661 (37.5)	44 (30.3)		2,234 (43.2)	18 (39.1)		1,913 (30.5)	30 (90.9)	
Unknown	3,152	469		1,308	205		213	52	
PR status									
Positive	12,945 (56.7)	85 (60.3)	0.4	2,818 (55.8)	24 (55.8)	1.0	3,758 (60.0)	22 (68.7)	0.3
Negative	9,878 (43.3)	56 (39.7)		2,235 (44.2)	19 (44.2)		2,503 (40.0)	*b*	
Unknown	3,433	473		1,425	208		231	53	
Grade									
Low	1,698 (7.4)	19 (8.7)	0.5	297 (6.0)	*b*	0.1	516 (8.4)	*b*	0.02
Moderate	7,573 (32.8)	64 (29.4)		1,643 (33.0)	18 (22.5)		2,136 (34.9)	13 (38.2)	
High	13,814 (59.8)	135 (61.9)		3,036 (61.0)	57 (71.3)		3,465 (56.6)	14 (41.2)	
Unknown	2,439	385		1,217	164		309	50	
Tumor size									
≤2 cm	11,307 (45.9)	44 (33.3)	0.004	2,875 (48.3)	23 (39.7)	0.2	2,535 (41.1)	*b*	0.4
>2 cm	13,333 (54.1)	88 (66.7)		3,078 (51.7)	35 (60.3)		3,637 (58.9)	11 (68.8)	
Unknown	1,436	481		485	192		280	69	

[a] *p*-value obtained from linear regression analysis.

b. Cell sizes <11 persons were suppressed to maintain confidentiality.

Table 5.3. Average annual percent change (AAPC) estimates and 95% CIs for the distribution of breast tumor characteristics among young women <40 years, overall and with distant stage at diagnosis, 1992-2011.

	All stages	Distant
ER status		
Positive	1.5 (1.1, 1.9)	1.9 (1.0, 2.7)
Negative	-2.7 (-3.4, -1.9)	-2.2 (-3.8, -0.6)
PR status		
Positive	0.5 (0.1, 1.0)	1.1 (-0.3, 2.5)
Negative	-0.7 (-1.3, -0.1)	-1.0 (-2.2, 0.2)
Grade		
Low	2.5 (1.4, 3.6)	4.4 (2.7, 6.0)[a]
Moderate	0.4 (-0.3, 1.1)	
High	-0.5 (-0.7, -0.3)	-1.1 (-1.7, -0.5)
Tumor size		
≤2 cm	-1.4 (-2.1, -0.7)	-1.2 (-3.7, 1.4)
>2cm	0.9 (0.4, 1.4)	0.2 (-0.7, 1.2)

[a]combined low/moderate grade due to small sample size.

All cases with missing data for a given characteristic are excluded.

Table 5.4. Tumor characteristics associated with imaging technology use among women ≥65 years of age, SEER-Medicare 13 program, 1992-2011.

	Not imaged (N=34,872)	Imaged (N=30,585)	OR (95% CI)[a]
Stage at diagnosis			
In situ	6,929 (20.3)	2,597 (8.6)	0.49 (0.47, 0.52
Localized	22,065 (64.6)	15,833 (52.6)	1.0
Regional	4,965 (14.5)	9,483 (31.5)	2.70 (2.59, 2.81)
Distant	212 (0.6)	2,171 (7.2)	14.8 (12.9, 17.1)
Unstaged	701	501	
ER status			
Positive (ref)	22,067 (85.7)	20,978 (81.9)	1.0
Negative	3,694 (14.3)	4,628 (18.1)	1.30 (1.24, 1.37)
Unknown	9,111	4,979	
PR status			
Positive (ref)	18,519 (53.1)	7,863 (31.1)	1.0
Negative	6,762 (19.4)	17,384 (68.9)	1.24 (1.20, 1.29)
Unknown	9,591	5,338	
Grade			
Low (ref)	7,013 (25.0)	5,281 (20.3)	1.0
Moderate	12,787 (45.6)	11,414 (43.8)	1.19 (1.14, 1.25)
High	8,245 (29.4)	9,377 (36.0)	1.52 (1.45, 1.59)
Unknown	6,827	4,513	
Tumor size			
≤2 cm (ref)	22,963 (73.7)	16,281 (58.4)	1.0
>2 cm	8,213 (26.3)	11,587 (41.6)	2.12 (2.05, 2.20)
Unknown	3,696	2,717	

[a]adjusted for age at diagnosis.

Table 5.5. Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) codes defining diagnostic imaging technology use.

Imaging Technology	Category	CPT/HCPCS Codes
CT	Brain/Head	70450-70492
	Breast/Chest	72125-72133
	Abdomen	71250-71270, 72192-72194, 74150-74170, 74176-74178
	Bone	73200-73202, 73700-73702
	Body	76380, 76497
PET	Body	78608, 78609, 78810-78816, G0125, G0126, G0165, G0210-G0228, G0231-G0235, G0252-G0254, G0296, G0330, G0331
MRI	Brain/Head	70336, 70540, 70542, 70543, 70551-70553, 70557-70559
	Breast/Chest	71550-71552, 75552, 75553, 75557, 75561, 76093, 76094, 77058, 77059, C8903-C8908
	Abdomen	72195-72197, 74181-74183
	Bone	72141, 72142, 72146-72149, 72156-72158, 73218-73223, 73718-73723
	Body	76498
Bone scan (nuclear medicine)	Bone/Body	76400, 78800-78804, 78102-78104, 78300-78320, 78399, 78999

Figure 5.1. Breast cancer incidence among women <40 years by SEER historic stage, SEER 13 program, 1992-2011.

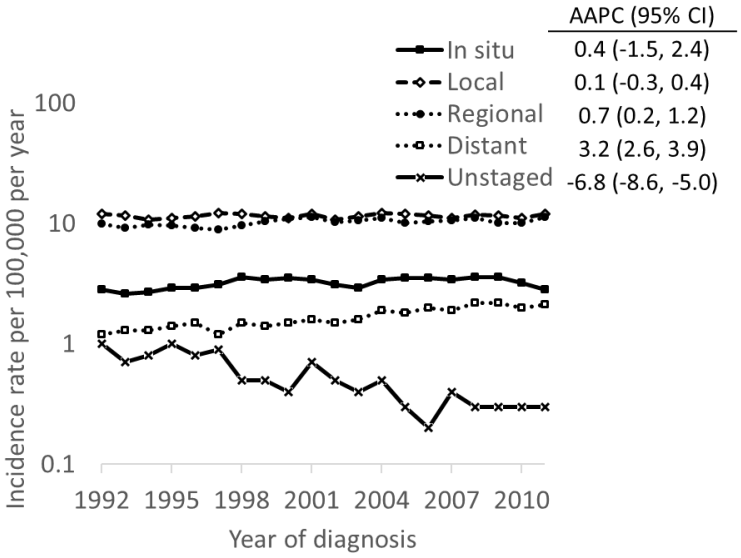


Figure 5.2. Distant breast cancer incidence among women <40 years according to ER status, SEER 13 program, 1992-2011.

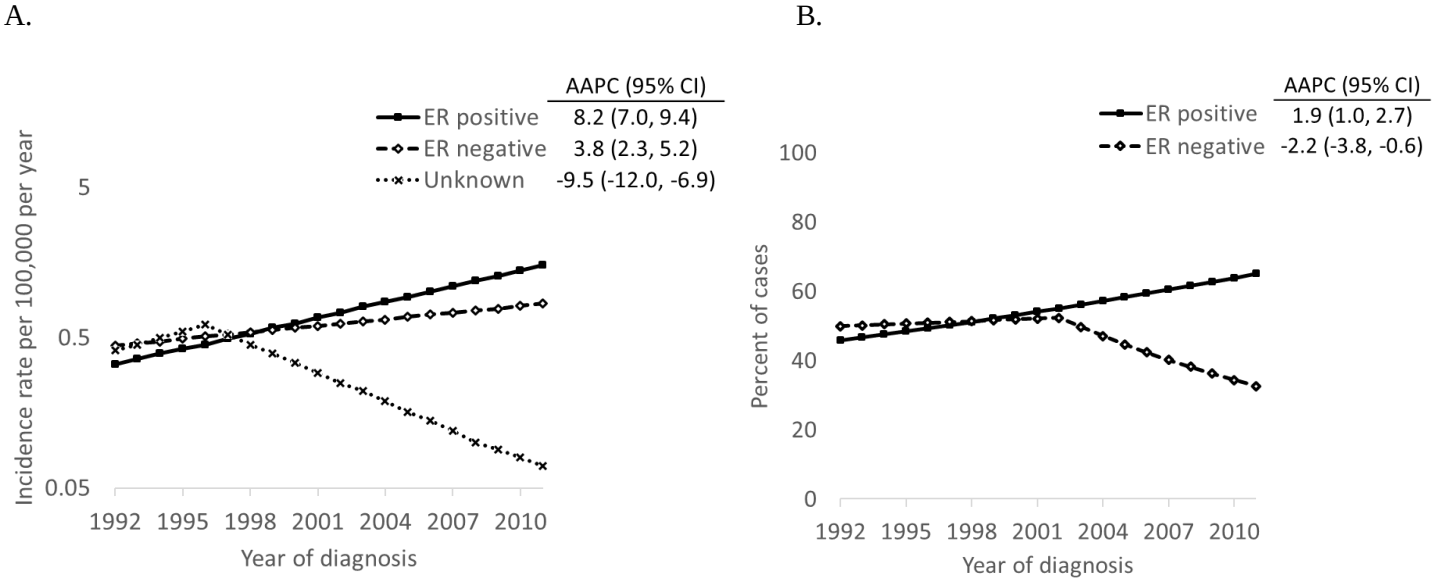
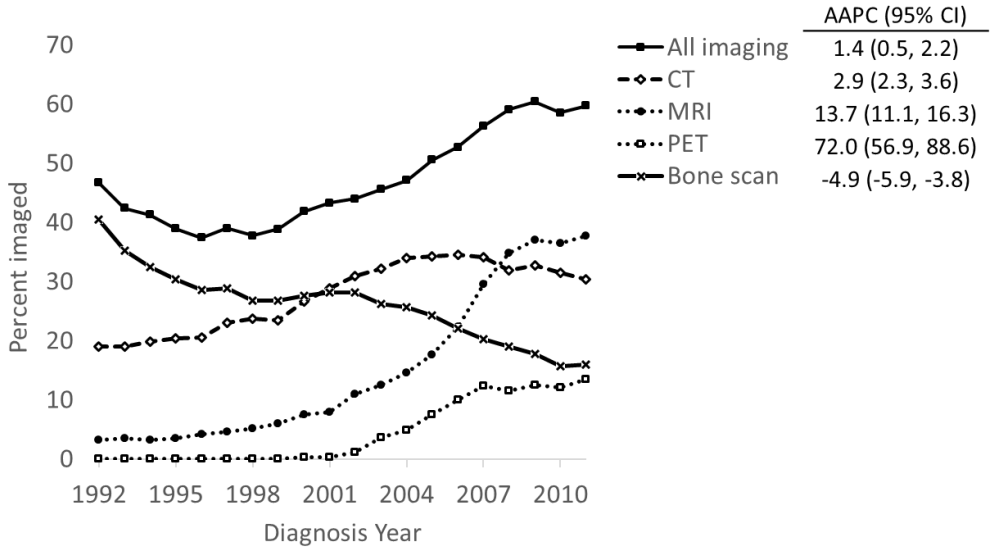


Figure 5.2 legend. Panels A and B show smoothed incidence rate trends estimated by joinpoint regression models. Panel B excludes cases with missing ER status.

Figure 5.3. Diagnostic imaging use among older breast cancer patients ≥65 years in the SEER-Medicare 13 program, 1992-2011.



CHAPTER 6: DISCUSSION

6.1 Summary of findings

The aims of this dissertation were (1) to examine the biology and epidemiology of breast cancers among premenopausal young (<40 years) black women and (2) to evaluate whether shifts in breast cancer biology and/or stage migration due to increased imaging use have contributed to the rise in distant stage breast cancer among young women in the U.S. First, using data from the AMBER Consortium, one of the largest studies of breast cancer epidemiology among black women to date, we found that young-onset breast cancers showed more aggressive tumor characteristic patterns than older-onset tumors. Specifically, young women tended to have increased ER and PR negativity, triple-negative subtype, higher grade, larger tumor size, and more advanced stage at diagnosis than older women, suggesting that young-onset breast cancers have a distinct, poorer-prognostic biology compared to older-onset disease. Additionally, we observed evidence of distinct breast cancer etiology according to age at diagnosis, with young women showing stronger risk associations with family history of breast cancer and waist-to-hip ratio as well as a protective association for breastfeeding compared to older women. These etiologic patterns were strongest when assessing age differences using the youngest cutpoint (<40 years) compared to older cutpoints of 45 and 50 years of age, highlighting that etiologic differences appear to be most pronounced when comparing the youngest women to older women. Increasing breastfeeding and reducing abdominal adiposity and oral contraceptive use were identified as potentially modifiable targets for intervention against young-onset disease among premenopausal black women.

Second, we used SEER and SEER-Medicare linked data to examine trends in young-onset breast cancer incidence, tumor biological characteristics, and diagnostic imaging use from 1992-2011. Among women <40 years of age, distant stage breast cancer incidence rates increased over time while

the rate of unstaged breast cancers markedly declined. Ascertainment of breast tumor characteristics also dramatically improved over time, with significant decreases in the rates of missing data for all tumor characteristics. We found evidence of a shift to more favorable breast tumor characteristics among young women since 1992, with the percent of cases with ER and PR negative as well as low grade disease increasing significantly over time while ER/PR negative and high grade disease significantly decreased. These trends were strongest when restricting to young women with distant stage disease, suggesting that the "biological shift" hypothesis has not contributed to rising distant breast cancer incidence rates. However, we observed significantly increased use of diagnostic imaging from 1992-2011, particularly PET and MRI scans, among Medicare-eligible SEER breast cancer cases, and imaging use was significantly associated with distant stage at diagnosis as well as poorer-prognostic tumor characteristics. Our work suggests that stage migration via increased imaging use may have contributed to rising rates of distant stage breast cancer among young women.

6.2 The roles of biology, etiology, and technology in young women's breast cancer

Among young, premenopausal black women <40 years, breast tumor biology showed a clear pattern of more aggressive disease characteristics compared to older-onset disease. This finding is consistent with other population-based studies of young women's breast cancer(8, 21-23, 27, 48-53, 57, 88, 105), supporting the hypothesis that breast cancers among young women <40 years are biologically distinct from those among older women. Interestingly, our evaluation of temporal trends in breast tumor characteristics in the SEER program did not reveal any evidence of increasing aggressive breast cancer biology from 1992-2011, but rather showed that young women's breast cancer may be shifting toward more favorable disease features that are more common among older women. This result is counter to our "biological shift" hypothesis, and suggests that, while young women appear to have distinct and aggressive tumor biology patterns compared to older women, young-onset disease in the U.S. may be changing in a positive direction over time.

It is possible that the temporal shift to more favorable tumor biological characteristics among young women may be explained in part by changing patterns of breast cancer risk factor exposure. In

the first aim of this dissertation, we showed evidence of different breast cancer etiologic patterns according to age at diagnosis, particularly for body size and reproductive exposures. It is well-established that rates of obesity have increased over the past several decades in the U.S.(132-134), and work conducted using the National Health and Nutrition Examination Survey (NHANES) has identified increasing rates of breastfeeding from 1993 to 2006, particularly among black women(135). We observed that these risk factors were associated with young women's breast cancer in contrasting ways, with adiposity increasing the odds of young-onset disease and breastfeeding having a protective effect. Others have also shown that these factors are differentially associated with aggressive breast cancer subtypes, in that obesity increases risk of basal-like subtype among young women while breastfeeding may be protective(17-19). Given that we did not observe an increase in aggressive disease characteristics among women with distant breast cancer, it is unlikely that changing risk factor patterns are responsible for increasing rates of distant disease at diagnosis. However, temporal shifts in exposures such as adiposity and breastfeeding may be contributing to an increase in more favorable disease characteristics over our study period, highlighting their potential as intervention factors to reduce the burden of young women's breast cancer in the U.S.

We identified that the changing use of diagnostic imaging technologies represents the most plausible explanation for the increasing trend in distant breast cancer incidence rates among young women. The dramatic increase in the percent of breast cancer cases receiving imaging use from 1992-2011 and the strong association between imaging and stage at diagnosis suggest that more women are receiving diagnostic imaging technologies in connection with their breast cancer diagnosis, particularly those with advanced stage disease. While we were unable to directly assess imaging patterns among young women <40 years, other work has reported higher rates of diagnostic imaging among young compared to older women(95, 131). These usage patterns may be a response to the fact that young women <40 years tend to have more advanced stage at diagnosis, as young women are not recommended to receive regular mammographic screening and thereby rely on clinical and self-detection methods(29, 31). Our work reveals that the rise in distant stage breast cancer among young

women may be a beneficial phenomenon, reflecting the use of improved and more accurate staging practices over time. Temporal changes in breast cancer incidence patterns may not be related to etiologic shifts requiring intervention, and cautious interpretation is warranted to prevent undue concern.

6.3 Significance

This dissertation addressed two important gaps in epidemiologic research regarding young women's breast cancer, characterizing the epidemiology of breast cancer among young black women and evaluating hypotheses for the rise in young-onset distant stage breast cancer in the U.S. Using the largest and most comprehensive data programs to date, the AMBER Consortium and the SEER and SEER-Medicare cancer registries, this research provides much-needed insight into the biology and etiology of young black women's breast cancer and clarifies the public health impact of shifting breast cancer incidence trends among young women.

Our findings demonstrate that young black women are significantly more likely to have breast cancers with poorer-prognostic characteristics, underscoring the need for effective interventions to reduce the burden of breast cancer in this demographic. We identified a specific pattern of modifiable factors that could be targeted to decrease the risk of young-onset breast cancers among black women, including promoting breastfeeding and reducing both abdominal adiposity and the use of oral contraceptives. Additionally, these etiologic differences did not differ by ER status, suggesting that interventions targeting young women may protect against young-onset disease across tumor subtypes. This research represents one of the largest epidemiologic studies of breast cancer among young black women to date and provides key insight into the biologic and etiologic features that distinguish young-onset disease from that of older-onset breast cancers.

Additionally, this dissertation is the first investigation of two competing hypotheses for the observed rise in distant breast cancer incidence rates among young women from 1992-2011. By evaluating temporal trends in both breast cancer tumor characteristics and diagnostic imaging use, this research provided a comprehensive analysis of the most plausible contributors to shifting distant

disease incidence trends. Our work identifies that imaging use increased dramatically from 1992-2011 while breast tumor characteristics showed correspondingly small incidence shifts to more favorable disease features, and stage migration due to improved imaging use should be considered as a key contributor to distant breast cancer incidence patterns among young women. This research underscores the importance of evaluating contrasting reasons for shifting temporal incidence trends and asserts that increasing rates of distant breast cancer among young women may be a beneficial result of more accurate staging practices over time.

6.4 Future Directions

In this dissertation, we identified potentially modifiable risk factors associated with breast cancer among young black women. While large epidemiologic studies of breast cancer with sufficient representation of young and black women are rare, increased attention to this at-risk demographic would enable validation of our findings in other study populations. Only one known prospective cohort study to date, the Black Women's Health Study(108), has evaluated risk factors associated with incidence of young-onset disease, and this study was included within the AMBER Consortium. Future prospective epidemiologic work in other study populations would improve understanding regarding the impact of these factors and associated interventions on the risk of young-onset breast cancer.

We found little evidence that the biology of young-onset breast cancers changed over time within the SEER 13 program; however, we posited that temporal trends in risk factor exposure may have shifted over time in ways that influence breast cancer incidence patterns among young women. In our study, abdominal adiposity and oral contraceptive use were associated with increased odds of young-onset breast cancer, while breastfeeding was protective. Increasing rates of obesity, hormone use, and breastfeeding behaviors over time could impact patterns in breast tumorigenesis and progression, resulting in a temporal shift in breast cancer incidence. Furthermore, evidence suggests that these factors are differentially associated with breast tumor subtypes(17-19), thus potentially influencing trends in tumor biological features as well. Due to the nature of longitudinal studies and

the need for existing data extending back to the 1990s, epidemiologic studies of temporal patterns in etiologic factors have been rare. However, future studies examining demographic and risk factor data over time within the NHANES program and other longitudinal data programs may reveal connections between shifting etiologic factors and cancer incidence trends.

Finally, our work identifies that stage migration resulting from increased imaging use may have contributed to shifting distant breast cancer incidence trends among young women. We evaluated this hypothesis by examining longitudinal trends in the use of four major diagnostic imaging technologies used for breast cancer staging (i.e., PET, MRI, CT, and bone scans) and found dramatic changes in imaging use since 1992. Under the stage migration hypothesis, the use of imaging would result in improved identification of distant metastases at time of diagnosis; therefore, young breast cancer cases receiving imaging would be accurately classified as having distant stage rather than regional stage at diagnosis. This reclassification of cases would increase the incidence of distant disease while improving disease prognosis due to the earlier identification of metastases. Future studies examining longitudinal trends in stage-specific breast cancer survival among young women would further clarify the role of stage migration, as survival would be expected to improve among both distant and regional stage cases over time. Additionally, evaluating whether stage-specific incidence rates have changed among young women according to imaging status would enable a direct assessment of the association between imaging use and incidence trends in young-onset distant stage breast cancer.

6.5 Conclusions

In conclusion, this dissertation examined the ways in which breast cancer biology and etiology differ according to age at diagnosis among black women and addressed possible underlying causes for the recent rise in distant breast cancer incidence among young women <40 years of age. We identified potentially modifiable targets for the prevention of breast cancers among young black women, a key demographic at risk for aggressive, advanced stage disease and higher mortality. Additionally, this work included a novel and thorough evaluation of the most credible contributors to

the rise in young-onset distant breast cancer from 1992-2011, considering the roles of shifting breast cancer biology and diagnostic imaging use patterns over time. We clarified the public health impact of observed incidence trends, identifying stage migration and sharply decreasing rates of missing data as key factors that may explain the rise in distant disease among young women. This work highlights the need for cautious interpretation of longitudinal cancer incidence patterns, as shifting trends may not be related to disease etiology.

REFERENCES

1. Desantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. CA Cancer J Clin 2014;64:52-62.

2. Smigal C, Jemal A, Ward E, Cokkinides V, Smith R, Howe HL, et al. Trends in breast cancer by race and ethnicity: update 2006. CA Cancer J Clin 2006;56:168-83.

3. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012;367:1998-2005.

4. Johnson RH, Chien FL, Bleyer A. Incidence of breast cancer with distant involvement among women in the United States, 1976 to 2009. JAMA 2013;309:800-5.

5. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.

6. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98:10869-74.

7. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530-6.

8. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 2006;295:2492-502.

9. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A 2003;100:10393-8.

10. Sandhu R, Parker JS, Jones WD, Livasy CA, Coleman WB. Microarray-Based Gene Expression Profiling for Molecular Classification of Breast Cancer and Identification of New Targets for Therapy. Lab Med 2010;41:364-72.

11. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987;235:177-82.

12. Seshadri R, Firgaira FA, Horsfall DJ, McCaul K, Setlur V, Kitchen P. Clinical significance of HER-2/neu oncogene amplification in primary breast cancer. The South Australian Breast Cancer Study Group. J Clin Oncol 1993;11:1936-42.

13. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, et al. Metastatic behavior of breast cancer subtypes. J Clin Oncol 2010;28:3271-7.

14. Baselga J, Perez EA, Pienkowski T, Bell R. Adjuvant trastuzumab: a milestone in the treatment of HER-2-positive early breast cancer. Oncologist 2006;11 Suppl 1:4-12.

15. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 2002;20:719-26.

16. Dawood S, Broglio K, Buzdar AU, Hortobagyi GN, Giordano SH. Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: an institutional-based review. J Clin Oncol 2010;28:92-8.

17. Millikan RC, Newman B, Tse CK, Moorman PG, Conway K, Dressler LG, et al. Epidemiology of basal-like breast cancer. Breast Cancer Res Treat 2008;109:123-39.

18. Althuis MD, Fergenbaum JH, Garcia-Closas M, Brinton LA, Madigan MP, Sherman ME. Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. Cancer Epidemiol Biomarkers Prev 2004;13:1558-68.

19. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. J Natl Cancer Inst 2011;103:250-63.

20. DeSantis C, Siegel R, Bandi P, Jemal A. Breast cancer statistics, 2011. CA Cancer J Clin 2011;61:409-18.

21. Anderson WF, Chen BE, Brinton LA, Devesa SS. Qualitative age interactions (or effect modification) suggest different cancer pathways for early-onset and late-onset breast cancers. Cancer Causes Control 2007;18:1187-98.

22. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A. Breast cancer before age 40 years. Semin Oncol 2009;36:237-49.

23. Keegan TH, DeRouen MC, Press DJ, Kurian AW, Clarke CA. Occurrence of breast cancer subtypes in adolescent and young adult women. Breast Cancer Res 2012;14:R55.

24. El Saghir NS, Seoud M, Khalil MK, Charafeddine M, Salem ZK, Geara FB, et al. Effects of young age at presentation on survival in breast cancer. BMC Cancer 2006;6:194.

25. Sidoni A, Cavaliere A, Bellezza G, Scheibel M, Bucciarelli E. Breast cancer in young women: clinicopathological features and biological specificity. Breast 2003;12:247-50.

26. Winchester DP, Osteen RT, Menck HR. The National Cancer Data Base report on breast carcinoma characteristics and outcome in relation to age. Cancer 1996;78:1838-43.

27. Anders CK, Hsu DS, Broadwater G, Acharya CR, Foekens JA, Zhang Y, et al. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol 2008;26:3324-30.

28. Fredholm H, Eaker S, Frisell J, Holmberg L, Fredriksson I, Lindman H. Breast cancer in young women: poor survival despite intensive treatment. PLoS One 2009;4:e7695.

29. Gajdos C, Tartter PI, Bleiweiss IJ, Bodian C, Brower ST. Stage 0 to stage III breast cancer in young women. J Am Coll Surg 2000;190:523-9.

30. Maggard MA, O'Connell JB, Lane KE, Liu JH, Etzioni DA, Ko CY. Do young breast cancer patients have worse outcomes? J Surg Res 2003;113:109-13.

31. Ruddy KJ, Gelber S, Tamimi RM, Schapira L, Come SE, Meyer ME, et al. Breast cancer presentation and diagnostic delays in young women. Cancer 2014;120:20-5.

32. Velentgas P, Daling JR. Risk factors for breast cancer in younger women. J Natl Cancer Inst Monogr 1994:15-24.

33. Anderson WF, Matsuno RK, Sherman ME, Lissowska J, Gail MH, Brinton LA, et al. Estimating age-specific breast cancer risks: a descriptive tool to identify age interactions. Cancer Causes Control 2007;18:439-47.

34. White E, Malone KE, Weiss NS, Daling JR. Breast cancer among young U.S. women in relation to oral contraceptive use. J Natl Cancer Inst 1994;86:505-14.

35. Althuis MD, Brogan DD, Coates RJ, Daling JR, Gammon MD, Malone KE, et al. Breast cancers among very young premenopausal women (United States). Cancer Causes Control 2003;14:151-60.

36. Warner ET, Colditz GA, Palmer JR, Partridge AH, Rosner BA, Tamimi RM. Reproductive factors and risk of premenopausal breast cancer by age at diagnosis: are there differences before and after age 40? Breast Cancer Res Treat 2013;142:165-75.

37. Anderson WF, Jatoi I, Devesa SS. Distinct breast cancer incidence and prognostic patterns in the NCI's SEER program: suggesting a possible link between etiology and outcome. Breast Cancer Res Treat 2005;90:127-37.

38. Tavani A, Gallus S, La Vecchia C, Negri E, Montella M, Dal Maso L, et al. Risk factors for breast cancer in women under 40 years. Eur J Cancer 1999;35:1361-7.

39. Hirose K, Tajima K, Hamajima N, Inoue M, Takezaki T, Kuroishi T, et al. A large-scale, hospital-based case-control study of risk factors of breast cancer according to menopausal status. Jpn J Cancer Res 1995;86:146-54.

40. Newcomb PA, Storer BE, Longnecker MP, Mittendorf R, Greenberg ER, Clapp RW, et al. Lactation and a reduced risk of premenopausal breast cancer. N Engl J Med 1994;330:81-7.

41. Borges VF, Schedin PJ. Pregnancy-associated breast cancer: an entity needing refinement of the definition. Cancer 2012;118:3226-8.

42. Lyons TR, Schedin PJ, Borges VF. Pregnancy and breast cancer: when they collide. J Mammary Gland Biol Neoplasia 2009;14:87-98.

43. Pathak DR, Speizer FE, Willett WC, Rosner B, Lipnick RJ. Parity and breast cancer risk: possible effect on age at diagnosis. Int J Cancer 1986;37:21-5.

44. Hunter DJ, Willett WC. Diet, body size, and breast cancer. Epidemiol Rev 1993;15:110-32.

45. Ursin G, Longnecker MP, Haile RW, Greenland S. A meta-analysis of body mass index and risk of premenopausal breast cancer. Epidemiology 1995;6:137-41.

46. Franceschi S, Favero A, La Vecchia C, Baron AE, Negri E, Dal Maso L, et al. Body size indices and breast cancer risk before and after menopause. Int J Cancer 1996;67:181-6.

47. Brinton LA, Swanson CA. Height and weight at various ages and risk of breast cancer. Ann Epidemiol 1992;2:597-609.

48. Anderson WF, Rosenberg PS, Menashe I, Mitani A, Pfeiffer RM. Age-related crossover in breast cancer incidence rates between black and white ethnic groups. J Natl Cancer Inst 2008;100:1804-14.

49. Clarke CA, Keegan TH, Yang J, Press DJ, Kurian AW, Patel AH, et al. Age-specific incidence of breast cancer subtypes: understanding the black-white crossover. J Natl Cancer Inst 2012;104:1094-101.

50. Furberg H, Millikan R, Dressler L, Newman B, Geradts J. Tumor characteristics in African American and white women. Breast Cancer Res Treat 2001;68:33-43.

51. Cunningham JE, Montero AJ, Garrett-Mayer E, Berkel HJ, Ely B. Racial differences in the incidence of breast cancer subtypes defined by combined histologic grade and hormone receptor status. Cancer Causes Control 2010;21:399-409.

52. Hahn KM, Bondy ML, Selvan M, Lund MJ, Liff JM, Flagg EW, et al. Factors associated with advanced disease stage at diagnosis in a population-based study of patients with newly diagnosed breast cancer. Am J Epidemiol 2007;166:1035-44.

53. Jones BA, Kasl SV, Curnen MG, Owens PH, Dubrow R. Can mammography screening explain the race difference in stage at diagnosis of breast cancer? Cancer 1995;75:2103-13.

54. Lannin DR, Mathews HF, Mitchell J, Swanson MS, Swanson FH, Edwards MS. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. JAMA 1998;279:1801-7.

55. Li CI, Malone KE, Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. Arch Intern Med 2003;163:49-56.

56. Moorman PG, Jones BA, Millikan RC, Hall IJ, Newman B. Race, anthropometric factors, and stage at diagnosis of breast cancer. Am J Epidemiol 2001;153:284-91.

57. Richardson JL, Langholz B, Bernstein L, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. Br J Cancer 1992;65:922-6.

58. McGee SA, Durham DD, Tse CK, Millikan RC. Determinants of breast cancer treatment delay differ for African American and White women. Cancer Epidemiol Biomarkers Prev 2013;22:1227-38.

59. Hall IJ, Moorman PG, Millikan RC, Newman B. Comparative analysis of breast cancer risk factors among African-American women and White women. Am J Epidemiol 2005;161:40-51.

60. Mayberry RM. Age-specific patterns of association between breast cancer and risk factors in black women, ages 20 to 39 and 40 to 54. Ann Epidemiol 1994;4:205-13.

61. Jemal A, Ward E, Thun MJ. Recent trends in breast cancer incidence rates by age and tumor characteristics among U.S. women. Breast Cancer Res 2007;9:R28.

62. Brinton LA, Sherman ME, Carreon JD, Anderson WF. Recent trends in breast cancer among younger women in the United States. J Natl Cancer Inst 2008;100:1643-8.

63. Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, et al. SEER Cancer Statistics Review, 1975-2011: National Cancer Institute. Bethesda, MD; 2014. based on November 2013 SEER data submission, posted to the SEER web site, April 4]. Available from: http://seer.cancer.gov/csr/1975_2011/.

64. Hou N, Huo D. A trend analysis of breast cancer incidence rates in the United States from 2000 to 2009 shows a recent increase. Breast Cancer Res Treat 2013;138:633-41.

65. McDougall JA, Li CI. Trends in distant-stage breast, colorectal, and prostate cancer incidence rates from 1992 to 2004: potential influences of screening and hormonal factors. Horm Cancer 2010;1:55-62.

66. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312:1604-8.

67. Morgensztern D, Goodgame B, Baggstrom MQ, Gao F, Govindan R. The effect of FDG-PET on the stage distribution of non-small cell lung cancer. J Thorac Oncol 2008;3:135-9.

68. Dinan MA, Curtis LH, Carpenter WR, Biddle AK, Abernethy AP, Patz EF, Jr., et al. Stage migration, selection bias, and survival associated with the adoption of positron emission tomography among medicare beneficiaries with non-small-cell lung cancer, 1998-2003. J Clin Oncol 2012;30:2725-30.

69. Chee KG, Nguyen DV, Brown M, Gandara DR, Wun T, Lara PN, Jr. Positron emission tomography and improved survival in patients with lung cancer: the Will Rogers phenomenon revisited. Arch Intern Med 2008;168:1541-9.

70. Eheman CR, Shaw KM, Ryerson AB, Miller JW, Ajani UA, White MC. The changing incidence of in situ and invasive ductal and lobular breast carcinomas: United States, 1999-2004. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2009;18:1763-9.

71. Glass AG, Lacey JV, Jr., Carreon JD, Hoover RN. Breast cancer incidence, 1980-2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status. J Natl Cancer Inst 2007;99:1152-61.

72. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335-51.

73. Wo JY, Chen K, Neville BA, Lin NU, Punglia RS. Effect of very small tumor size on cancer-specific mortality in node-positive breast cancer. J Clin Oncol 2011;29:2619-27.

74. Yu KD, Jiang YZ, Chen S, Cao ZG, Wu J, Shen ZZ, et al. Effect of large tumor size on cancer-specific mortality in node-negative breast cancer. Mayo Clin Proc 2012;87:1171-80.

75. Yu KD, Wu J, Shen ZZ, Shao ZM. Hazard of breast cancer-specific mortality among women with estrogen receptor-positive breast cancer after five years from diagnosis: implication for extended endocrine therapy. J Clin Endocrinol Metab 2012;97:E2201-9.

76. Palmer JR, Ambrosone CB, Olshan AF. A collaborative study of the etiology of breast cancer subtypes in African American women: the AMBER consortium. Cancer Causes Control 2014;25:309-19.

77. Newman B, Moorman PG, Millikan R, Qaqish BF, Geradts J, Aldrich TE, et al. The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. Breast Cancer Res Treat 1995;35:51-60.

78. Hair BY, Hayes S, Tse CK, Bell MB, Olshan AF. Racial differences in physical activity among breast cancer survivors: implications for breast cancer care. Cancer 2014;120:2174-82.

79. Rosenberg L, Adams-Campbell L, Palmer JR. The Black Women's Health Study: a follow-up study for causes and preventions of illness. J Am Med Womens Assoc 1995;50:56-8.

80. Rosenberg L, Palmer JR, Bethea TN, Ban Y, Kipping-Ruane K, Adams-Campbell LL. A prospective study of physical activity and breast cancer incidence in African-American women. Cancer Epidemiol Biomarkers Prev 2014;23:2522-31.

81. Ambrosone CB, Zirpoli G, Ruszczyk M, Shankar J, Hong CC, McIlwain D, et al. Parity and breastfeeding among African-American women: differential effects on breast cancer risk by estrogen receptor status in the Women's Circle of Health Study. Cancer Causes Control 2014;25:259-65.

82. National Cancer Institute. Surveillance, Epidemiology, and End Results Program [January 10, 2015]. Available from: seer.cancer.gov.

83. Mohanty S, Bilimoria KY. Comparing national cancer registries: The national cancer data base (NCDB) and the surveillance, epidemiology, and end results (SEER) program. J Surg Oncol 2014.

84. Engels EA, Pfeiffer RM, Ricker W, Wheeler W, Parsons R, Warren JL. Use of surveillance, epidemiology, and end results-medicare data to conduct case-control studies of cancer among the US elderly. Am J Epidemiol 2011;174:860-70.

85. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. Med Care 2002;40:IV-3-18.

86. National Cancer Institute: Applied Research, Cancer Control and Population Sciences. SEER-Medicare linked database.

87. National Heart, Lung, and Blood Institute (NHLBI) Expert Panel on the identification, evaluation, and treatment of overweight and obesity in adults. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report. National Institutes of Health. Obes Res 1998;6 Suppl 2:51S-209S.

88. Chollet-Hinton L, Anders CK, Tse CK, Bell MB, Yang YC, Carey LA, et al. Breast cancer biologic and etiologic heterogeneity by young age and menopausal status in the Carolina Breast Cancer Study: a case-control study. Breast Cancer Res 2016;18:79.

89. Allott EH, Cohen SM, Geradts J, Sun X, Khoury T, Bshara W, et al. Performance of Three-Biomarker Immunohistochemistry for Intrinsic Breast Cancer Subtyping in the AMBER Consortium. Cancer Epidemiol Biomarkers Prev 2016;25:470-8.

90. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 2010;28:2784-95.

91. Bernstein JR. Role of stereotactic breast biopsy. Semin Surg Oncol 1996;12:290-9.

92. Yim JH, Barton P, Weber B, Radford D, Levy J, Monsees B, et al. Mammographically detected breast cancer. Benefits of stereotactic core versus wire localization biopsy. Ann Surg 1996;223:688-97; discussion 97-700.

93. Kaufman CS, Delbecq R, Jacobson L. Excising the reexcision: stereotactic core-needle biopsy decreases need for reexcision of breast cancer. World J Surg 1998;22:1023-7; discussion 8.

94. Morrow M, Venta L, Stinson T, Bennett C. Prospective comparison of stereotactic core biopsy and surgical excision as diagnostic procedures for breast cancer patients. Ann Surg 2001;233:537-41.

95. Gold LS, Buist DS, Loggers ET, Etzioni R, Kessler L, Ramsey SD, et al. Advanced diagnostic breast cancer imaging: variation and patterns of care in washington state. J Oncol Pract 2013;9:e194-202.

96. Gorin SS, Heck JE, Cheng B, Smith SJ. Delays in breast cancer diagnosis and treatment by racial/ethnic group. Arch Intern Med 2006;166:2244-52.

97. Gorey KM, Luginaah IN, Holowaty EJ, Fung KY, Hamm C. Wait times for surgical and adjuvant radiation treatment of breast cancer in Canada and the United States: greater socioeconomic inequity in America. Clin Invest Med 2009;32:E239-49.

98. Dinan MA, Curtis LH, Hammill BG, Patz EF, Jr., Abernethy AP, Shea AM, et al. Changes in the use and costs of diagnostic imaging among Medicare beneficiaries with cancer, 1999-2006. JAMA 2010;303:1625-31.

99. Surveillance Research Program, National Cancer Institute SEER*Stat software (seer.cancer.gov/seerstat) version 8.3.2.

100. Joinpoint Regression Program, Version 4.4.0.0 - January 2017. Statistical Methodology and Applications Branch, Surveillance Research Program, National Cancer Institute.

101. Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK. Estimating average annual per cent change in trend analysis. Stat Med 2009;28:3670-82.

102. Anderson WF, Katki HA, Rosenberg PS. Incidence of breast cancer in the United States: current and future trends. J Natl Cancer Inst 2011;103:1397-402.

103. Howlader N, Noone AM, Yu M, Cronin KA. Use of imputed population-based cancer registry data as a method of accounting for missing information: application to estrogen receptor status for breast cancer. Am J Epidemiol 2012;176:347-56.

104. Gold LS, Klein G, Carr L, Kessler L, Sullivan SD. The emergence of diagnostic imaging technologies in breast cancer: discovery, regulatory approval, reimbursement, and adoption in clinical guidelines. Cancer Imaging 2012;12:13-24.

105. Elledge RM, Clark GM, Chamness GC, Osborne CK. Tumor biologic factors and breast cancer prognosis among white, Hispanic, and black women in the United States. J Natl Cancer Inst 1994;86:705-12.

106. Palmer JR, Boggs DA, Wise LA, Ambrosone CB, Adams-Campbell LL, Rosenberg L. Parity and lactation in relation to estrogen receptor negative breast cancer in African American women. Cancer Epidemiol Biomarkers Prev 2011;20:1883-91.

107. Palmer JR, Viscidi E, Troester MA, Hong CC, Schedin P, Bethea TN, et al. Parity, lactation, and breast cancer subtypes in African American women: results from the AMBER Consortium. J Natl Cancer Inst 2014;106.

108. Bertrand KA, Bethea TN, Adams-Campbell LL, Rosenberg L, Palmer JR. Differential patterns of risk factors for early-onset breast cancer by ER status in African American women. Cancer Epidemiol Biomarkers Prev 2016.

109. Li CI, Beaber EF, Tang MT, Porter PL, Daling JR, Malone KE. Reproductive factors and risk of estrogen receptor positive, triple-negative, and HER2-neu overexpressing breast cancer among women 20-44 years of age. Breast Cancer Res Treat 2013;137:579-87.

110. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000;151:346-57.

111. Connolly BS, Barnett C, Vogt KN, Li T, Stone J, Boyd NF. A meta-analysis of published literature on waist-to-hip ratio and risk of breast cancer. Nutr Cancer 2002;44:127-38.

112. Mannisto S, Pietinen P, Pyy M, Palmgren J, Eskelinen M, Uusitupa M. Body-size indicators and risk of breast cancer according to menopause and estrogen-receptor status. Int J Cancer 1996;68:8-13.

113. Huang Z, Willett WC, Colditz GA, Hunter DJ, Manson JE, Rosner B, et al. Waist circumference, waist:hip ratio, and risk of breast cancer in the Nurses' Health Study. Am J Epidemiol 1999;150:1316-24.

114. Li R, Darling N, Maurice E, Barker L, Grummer-Strawn LM. Breastfeeding rates in the United States by characteristics of the child, mother, or family: the 2002 National Immunization Survey. Pediatrics 2005;115:e31-7.

115. Dolle JM, Daling JR, White E, Brinton LA, Doody DR, Porter PL, et al. Risk factors for triple-negative breast cancer in women under the age of 45 years. Cancer Epidemiol Biomarkers Prev 2009;18:1157-66.

116. Bethea TN, Rosenberg L, Hong CC, Troester MA, Lunetta KL, Bandera EV, et al. A case-control analysis of oral contraceptive use and breast cancer subtypes in the African American Breast Cancer Epidemiology and Risk Consortium. Breast Cancer Res 2015;17:22.

117. Brinton LA, Daling JR, Liff JM, Schoenberg JB, Malone KE, Stanford JL, et al. Oral contraceptives and breast cancer risk among younger women. J Natl Cancer Inst 1995;87:827-35.

118. Colditz GA, Kaphingst KA, Hankinson SE, Rosner B. Family history and risk of breast cancer: nurses' health study. Breast Cancer Res Treat 2012;133:1097-104.

119. Collaborative Group on Hormonal Factors in Breast C. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet 2001;358:1389-99.

120. Gilliland FD, Joste N, Stauber PM, Hunt WC, Rosenberg R, Redlich G, et al. Biologic characteristics of interval and screen-detected breast cancers. J Natl Cancer Inst 2000;92:743-9.

121. Elkin EB, Hudis C, Begg CB, Schrag D. The effect of changes in tumor size on breast carcinoma survival in the U.S.: 1975-1999. Cancer 2005;104:1149-57.

122. Polednak AP. Increase in Distant Stage Breast Cancer Incidence Rates in US Women Aged 25-49 Years, 2000-2011: The Stage Migration Hypothesis. J Cancer Epidemiol 2015;2015:710106.

123. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 13 Regs Research Data, Nov 2015 Sub (1992-2011) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1992-2011 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2016, based on the November 2015 submission.

124. DeSantis C, Howlader N, Cronin KA, Jemal A. Breast cancer incidence rates in U.S. women are no longer declining. Cancer Epidemiol Biomarkers Prev 2011;20:733-9.

125. Johnson RH, Chien FL, Bleyer A. Incidence rate of breast cancer in young women--reply. JAMA 2013;309:2435-6.

126. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger PA. A multivariate technique for multiply imputing missing values using a sequence of regression models. Surv Methodol 2001;27:85-95.

127. McCartan DP, Prichard RS, MacDermott RJ, Rothwell J, Geraghty J, Evoy D, et al. Role of bone scan in addition to CT in patients with breast cancer selected for systemic staging. Br J Surg 2016;103:839-44.

128. Niikura N, Costelloe CM, Madewell JE, Hayashi N, Yu TK, Liu J, et al. FDG-PET/CT compared with conventional imaging in the detection of distant metastases of primary breast cancer. Oncologist 2011;16:1111-9.

129. Mahner S, Schirrmacher S, Brenner W, Jenicke L, Habermann CR, Avril N, et al. Comparison between positron emission tomography using 2-[fluorine-18]fluoro-2-deoxy-D-glucose, conventional imaging and computed tomography for staging of breast cancer. Ann Oncol 2008;19:1249-54.

130. Crivello ML, Ruth K, Sigurdson ER, Egleston BL, Evers K, Wong YN, et al. Advanced imaging modalities in early stage breast cancer: preoperative use in the United States Medicare population. Ann Surg Oncol 2013;20:102-10.

131. Stout NK, Nekhlyudov L, Li L, Malin ES, Ross-Degnan D, Buist DS, et al. Rapid increase in breast magnetic resonance imaging use: trends from 2000 to 2011. JAMA Intern Med 2014;174:114-21.

132. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. Lancet 2011;377:557-67.

133. Sturm R, Hattori A. Morbid obesity rates continue to rise rapidly in the United States. Int J Obes (Lond) 2013;37:889-91.

134. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA 2012;307:491-7.

135. McDowell MM, Wang CY, Kennedy-Stephenson J. Breastfeeding in the United States: findings from the national health and nutrition examination surveys, 1999-2006. NCHS Data Brief 2008:1-8.