THE EFFECTS OF ANTHROPOMETRY AND ANGIOGENESIS ON BREAST CANCER ETIOLOGY

by

Katherine Whitney Reeves

B.S., Yale University, 2002

M.P.H., The Ohio State University, 2003

Submitted to the Graduate Faculty of

Department of Epidemiology

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Katherine W. Reeves

It was defended on

January 11, 2008

and approved by

Dissertation Advisor:
Jane A. Cauley, DrPH, MPH
Professor of Epidemiology
Associate Professor of Nursing
Graduate School of Public Health
School of Nursing
University of Pittsburgh

Roberta B. Ness, MD, MPH
Professor and Chair of Epidemiology
Graduate School of Public Health
University of Pittsburgh

Roslyn A. Stone, PhD
Associate Professor of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Victor G. Vogel III, MD, MHS
Professor of Medicine and Epidemiology
Graduate School of Public Health
School of Medicine
University of Pittsburgh

Joel L. Weissfeld, MD, MPH
Associate Professor of Epidemiology
Assistant Professor of Medicine
Graduate School of Public Health
School of Medicine
University of Pittsburgh

Jane A. Cauley, DrPH, MPH

THE EFFECTS OF ANTHROPOMETRY AND ANGIOGENESIS ON BREAST CANCER ETIOLOGY

Katherine W. Reeves, PhD

University of Pittsburgh, 2008

Factors such as mammographic breast density and angiogenesis may be related to breast cancer development, though numerous questions about the etiologic mechanisms remain. Percent density is positively associated with breast cancer risk, yet is negatively associated with another breast cancer risk factor, body mass index (BMI). Vascular endothelial growth factor (VEGF) is a primary regulator of angiogenesis, yet its relationship to breast cancer risk is unclear. We evaluated the longitudinal association between BMI and breast density in the Study of Women's Health Across the Nation (SWAN) Mammographic Density Substudy (N=834). Using adjusted random intercept models, changes in BMI were not associated with changes in dense breast area (β=-0.0105, p=0.34), but were strongly negatively associated with changes in percent density (β=-1.18, p<0.001). Thus, effects of changes in anthropometry on percent breast density may reflect effects on non-dense tissue, rather than on the dense tissue where cancers arise. Breast density was measured from routine screening mammograms which were not timed with SWAN visits. We developed a method to align the off-schedule mammogram data to the study visit times using linear interpolation with multiple imputation. Our method was shown to be valid, with an average bias for dense breast area of 0.11 cm^2. In the random intercept models, use of a

simple matching algorithm to estimate breast density produced different (β=-0.0155, p=0.04), and likely incorrect, results. Our linear interpolation with multiple imputations method may be applicable to other longitudinal datasets with important data collected off-schedule. In a separate case-control study, the Mammograms and Masses Study (MAMS), we evaluated the association between serum VEGF levels and breast cancer (N=407). Geometric mean VEGF levels were higher among cases (331.4 pg/mL) than controls (291.4 pg/mL; p=0.21). In a multivariable logistic regression model, VEGF ≥314.2 pg/mL was positively associated with breast cancer (odds ratio 1.37, 95% confidence interval 0.88 – 2.12), albeit non-significantly. Higher levels of VEGF may increase breast cancer risk. We have identified roles for anthropometry and angiogenesis in breast carcinogenesis. Enhancing knowledge of breast cancer etiology is a significant contribution to public health and may lead to improved opportunities for prevention or early detection.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to take this opportunity to acknowledge the many individuals who have contributed both to this work and to my development as a cancer epidemiologist. I am extremely fortunate to have been surrounded by people who supported my education and who gave me the opportunity to pursue this line of research.

First and foremost, none of this work would have been possible without the generous participation of the women who joined the MAMS and SWAN studies. I am extremely grateful for the contributions of these women. I am also appreciative of the funding support from the Department of Epidemiology Small Grants Program and the National Cancer Institute's Predoctoral Cancer Epidemiology Training Grant.

The members of my committee have been an indispensable resource throughout the dissertation process. Dr. Jane Cauley mentored me and helped me to persevere through what could have been a difficult transition. That I was able to remain on schedule for completion of my degree is a credit to her exceptional advising. Dr. Cauley allowed me to work independently while always being accessible to discuss my research. Dr. Roberta Ness has challenged me to have a critical eye towards my own work. Her questions and critiques have improved this research and have made me a better epidemiologist. Dr. Roslyn Stone has spent countless hours with me working through the challenging statistical problems posed by my research questions. She has greatly enhanced my knowledge of biostatistics, and I sincerely appreciate her

dedication to my work. Dr. Joel Weissfeld has fostered my interest in cancer epidemiology through his outstanding ability to teach. His thoughtful critiques and willingness to review my analyses have significantly enhanced my research skills. Dr. Victor Vogel has given me the opportunity to see the clinical side of breast cancer. He has reminded me to think beyond the statistics and to always remain cognizant of the clinical relevance of my findings.

I thank Dr. Francesmary Modugno for mentoring me during my first year as a doctoral student. I am also appreciative of the individuals who, along with Dr. Modugno, established the MAMS study. My fellow graduate students have supported my work by always being available to help me think through problems with my research or just to commiserate with the troubles that are typical of doctoral students. I am grateful for the friendship of Jessica Albano, Kwame Appenteng, Cher Dallal, Gretchen Gierach, Alana Hudson, Claudia Leiras, and Alicia McDonald.

I would also like to thank the investigators in SWAN and the SWAN Mammographic Density Study for their efforts in establishing this unique study. I am especially indebted to Drs. Laurel Habel and Marike Vuga for their support of my work.

I thank Dr. Electra Paskett for being a wonderful mentor to me, beginning when I was first learning about cancer epidemiology as a master's student. Her encouragement led me to pursue a doctorate, and she has always been supportive of my career. The friends, family members, and former teachers and coaches who have contributed to my development as a student and a person are too numerous to name. I am grateful for the lasting impact they have had on my life.

I thank my brothers for leading by example. Payson and Michael's accomplishments have always set a high bar for my own work. They have taught me to be strong and to not take myself too seriously—lessons that have served me well throughout this process.

I credit my husband with preserving my sanity throughout graduate school. Mike's encouragement and ability to listen have been indispensable. His unfailing belief in my abilities has helped me to persevere at times when I was ready to quit. I could not have accomplished this goal without Mike's love and support.

Lastly, I thank my parents for instilling in me a strong work ethic, a love of learning, and the belief that I could achieve anything I put my mind to. The sacrifices they made for my education have not gone unnoticed; I only hope that I have made them proud. Mom and Dad, I dedicate this work to you.

1.0 INTRODUCTION

Breast cancer is a significant public health problem in the United States. Despite decades of promising research advances, the incidence of breast cancer continues to rise in the U.S. Although much is understood about the etiology of breast cancer, opportunities for its prevention are limited, and most of the decrease in breast cancer mortality has resulted from earlier detection and improved treatment of the disease. One reason for the lack of preventive options for breast cancer is that research aimed at testing the efficacy of preventive interventions requires substantial commitments of time, participants, and monetary resources. As an alternative to following subjects for an outcome of incident cancer, many cancer epidemiology studies employ a surrogate endpoint instead, thus allowing for the study to be performed over a shorter period of time and with fewer participants. Mammographic breast density is determined by the relative proportions of fat and structural tissues in the breast and has been proposed for use as a surrogate endpoint in breast cancer prevention trials. Breast density is highly related to the amount of fat in a woman's breast, and how changes in weight relate to changes in breast density has not been documented. This issue must be understood prior to widespread use of breast density change as an endpoint in longitudinal studies.

Numerous biological processes such as angiogenesis are believed to play a role in breast cancer etiology. Angiogenesis is known to be necessary for the growth of tumors beyond a few millimeters in size, and vascular endothelial growth factor (VEGF) is believed to be one of the most important angiogenic factors. While some studies have reported that serum and plasma VEGF levels are higher among breast cancer cases than among controls, these studies have had

significant limitations including small sample size and failure to control for menopausal status. Thus, the role of VEGF in breast cancer etiology remains unclear.

For these reasons, the purpose of the present research is as follows: 1) to evaluate longitudinally how anthropometry is associated with breast density, 2) to describe an approach for estimating data collected off-schedule from planned study visits, and 3) to evaluate how a biomarker of angiogenesis is associated with breast cancer. The following literature review presents an overview of breast cancer epidemiology and known risk factors for breast cancer. A more detailed background on mammographic breast density and angiogenesis as they relate to breast cancer is also provided.

2.0 LITERATURE REVIEW

2.1 BREAST CANCER EPIDEMIOLOGY

It is predicted that 178,480 American women will have been diagnosed with invasive breast cancer in 2007 alone.[1] Breast cancer is the most commonly diagnosed cancer among women in the United States and accounts for 31% of all female cancers (excluding non-melanoma skin cancers and in situ cancers).[1] Though breast cancer may occur in men, this is a rare event; the American Cancer Society estimates 2,030 new cases of male breast cancer will have been diagnosed in 2007.[1] Among U.S. females breast cancer ranks second to lung cancer in terms of cancer morality, with 40,460 female breast cancer deaths predicted for 2007.[1] Breast cancer deaths account for 15% of the burden of cancer mortality among female Americans.[2] Mortality from breast cancer has decreased in recent years due to early detection and improved treatment of the disease. Data from the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) program show that the breast cancer mortality rate declined 2.3% each year between 1990 and 2003.[3] The percentage of women surviving at least five years after diagnosis has risen to 88%, and 5-year survival is 98% for women diagnosed with localized disease.[3]

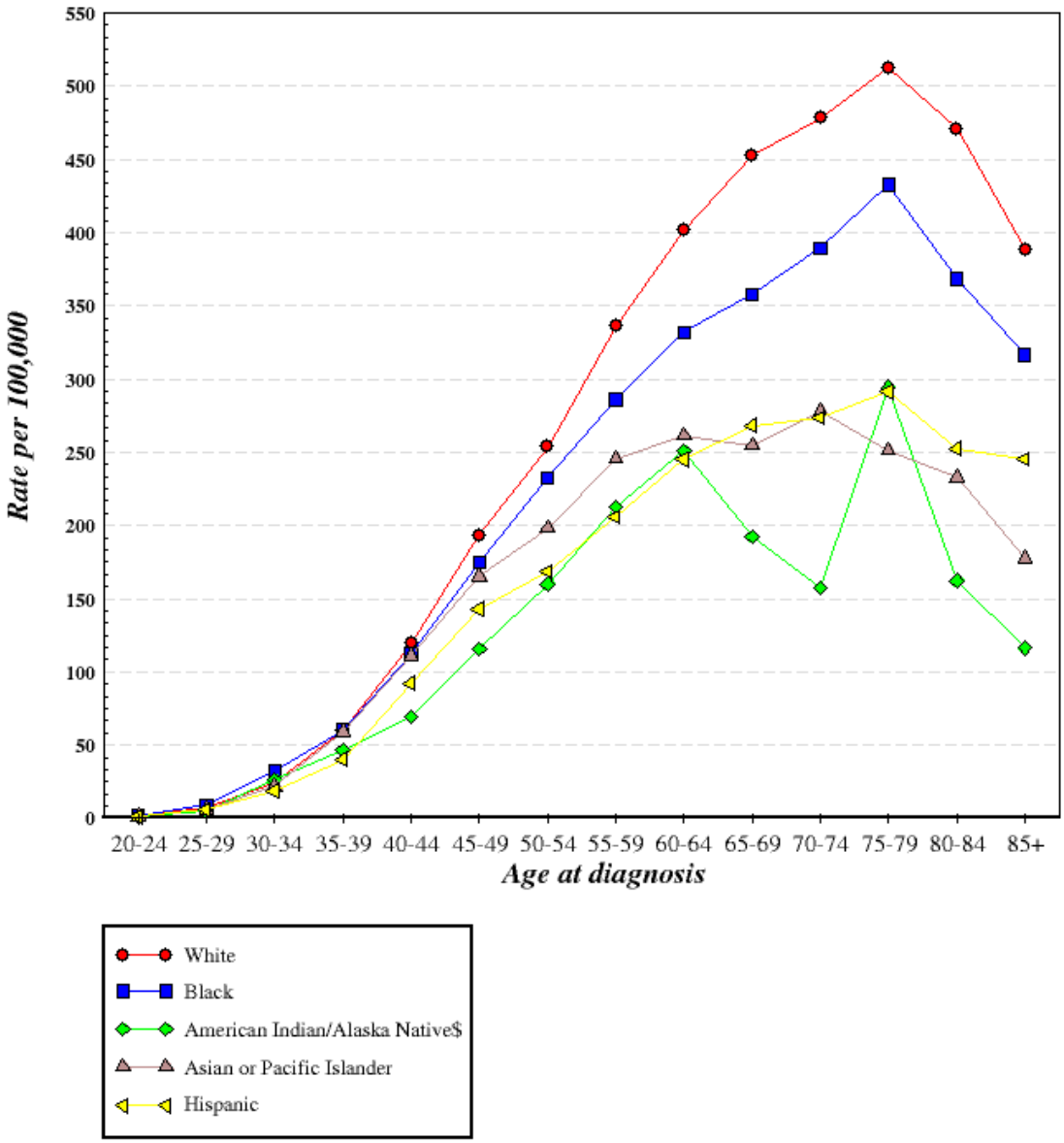


Figure 2.1 Age-specific breast cancer incidence rates among U.S. women by race, 1999-2003.

Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence-SEER 9 Regs Public-Use, Nov 2005 Sub (1973-2003), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2006, based on the November 2005 submission.

2.1.1 Menopausal status and breast cancer

Many oncologists and cancer epidemiologists consider pre- and postmenopausal breast cancer to represent two different etiologic forms of breast cancer. The overall incidence rate of breast cancer is low at younger ages (e.g. 1.4 per 100,000 among women ages 20-24).[4] As women begin to transition through menopause the rates of breast cancer increase substantially; data from SEER (Figure 2.1) show that between 1999 and 2003 the incidence rate of breast cancer was 119.3 per 100,000 for women ages 40-44, 249.0 per 100,000 for women ages 50-54, and 388.3 for women ages 60-64. The highest rate of breast cancer was observed among women ages 75-79, in whom 490.4 incident cases of breast cancer were diagnosed for every 100,000 women in this age group.[4] Further, since 1987 the incidence rate of breast cancer has been unchanged among women under age 40 and has decreased slightly among women age 40-49. The incidence of breast cancer among women over 50, however, has exhibited a slight increase during this time period.[3] Disease occurring among younger, premenopausal women tends to be more aggressive, and survival is lower among younger breast cancer patients. Often when menopausal status is unknown in research studies, the woman's age is used as a proxy with age 50 as the cutpoint to define pre- versus postmenopausal women. SEER statistics show that 5-year survival from breast cancer, among cases diagnosed between 1996 and 2002, was 86.7% in women under age 50 compared to 90.0% among women age $\geq$50.[5]

Indeed, some risk factors for breast cancer, such as family history and obesity, differ in their effect on breast cancer risk depending on the woman's age or menopausal status (see section 2.3.2). Mammographic breast density is a significant risk factor for breast cancer in both pre- and postmenopausal women, although, as reviewed by Boyd et al., there is some evidence

that it may be a stronger risk factor among postmenopausal women.[6] Due to the significant differences in breast cancer incidence, effects of risk factors, and endogenous hormonal environments between pre- and postmenopausal women, it is important that studies of breast cancer epidemiology consider that associations may vary by menopausal status.

2.1.2 Race/ethnicity and breast cancer

Breast cancer rates also differ by race and ethnicity. Although African American women have a lower overall incidence of breast cancer compared to Caucasian women (rates of 118.9 per 100,000 vs. 137.6 per 100,000, respectively for 1999-2003),[4] African Americans have a higher incidence of breast cancer before age 35.[3] Breast cancer mortality was substantially greater at all ages among African Americans (34.4 deaths per 100,000) than it was among Caucasians (25.4 deaths per 100,000) for the period 1999-2003.[7]

The reasons for these disparities are not well-understood, although a number of possible explanations have been suggested and investigated: 1) differential utilization of mammographic screening and stage at diagnosis, 2) differential effect and/or distribution of breast cancer risk factors, 3) differences in inherent genetic susceptibility, 4) differences in tumor characteristics, 5) differential access to treatment, and 6) differences in prevalence of comorbidities among women diagnosed with breast cancer. Each of the six hypothesized explanations for the racial disparities in breast cancer incidence and mortality has some merit. Mammography use is generally similar between African Americans and Caucasians,[3] and differences in stage at diagnosis can be at least partially explained by differences in obesity.[8] Therefore, screening and differences in stage at diagnosis are likely not the most important factors. Likewise, access to treatment is important, but disparities in mortality exist even in systems in which African

Americans and Caucasians have equal access.[9-11] Genetic susceptibility may also be important, yet the impact of such factors is not well-understood.[12-15] It appears that the differential distribution of risk factors, especially obesity,[16-18] differences in tumor characteristics,[19-27] and differences in comorbid conditions[28, 29] between African American and Caucasian breast cancer patients are largely responsible for the racial disparity observed in breast cancer.

Breast cancer incidence and mortality are higher among African Americans and Caucasians than among other races and ethnicities. According to SEER data, breast cancer incidence among Asians and Pacific Islanders in 1999-2003 was 93.5 per 100,000. Incidence among Hispanics for the same time period was 87.1 per 100,000, while incidence among Native Americans and Alaskan Natives was 74.4 for 1999-2002.[4] Mortality rates (per 100,000) for these time periods were 12.6 for Asians and Pacific Islanders, 16.3 for Hispanics, and 13.8 for Native Americans and Alaskan Natives.[7] A variety of genetic, environmental, and behavioral factors may explain these racial differences. Migration studies have documented that breast cancer risk is higher among Asian-American women born in the West compared to those born in the East (RR 1.6, 95% CI 1.2-2.1).[30] Breast cancer risk is further increased along with the number of the woman's grandparents born in the West (RR 1.89, 95% CI 1.2-3.0 for woman and 1-2 of her grandparents born in the West compared to woman and all grandparents born in the East), and risk is decreased among more recent immigrants (RR 0.32, 95% CI 0.18-0.57 for 2-4 years lived in the West) compared to those who have lived in the West their entire lives.[30] A recent study also reported that breast cancer rates among Asian Americans in Los Angeles County rose substantially between 1993-1997.[31] Most notably, the breast cancer rate among Japanese American women in this county was rapidly approaching that of non-Hispanic white

women.[31] Factors such as acculturation and adoption of a Western diet may at least partially explain these recent trends.[31]

2.2 BREAST CANCER RISK FACTORS

Besides age and race, a number of risk factors for breast cancer have been identified. Table 2.1 summarizes many of the known risk and protective factors for breast cancer. Women having a first-degree relative with a history of breast cancer are at increased risk of the disease themselves.[32, 33] The risk conferred by family history is further increased if the affected family member was diagnosed with the disease at a younger age. For example, a woman with a first-degree relative diagnosed with breast cancer before age 40 has a 5.7 times increased risk (99% CI 2.7-11.8) of being diagnosed with breast cancer before she is 40 compared to a woman of the same age but without a family history of breast cancer.[34] Two genes, BRCA1 and BRCA2, have been implicated in familial breast cancer, but account for less than 10% of all breast cancer cases.[49] BRCA mutations are most strongly related to breast cancer occurring in younger, premenopausal women. Among women diagnosed with breast cancer before age 40, 9% have a BRCA mutation, compared to only 2% of women of any age diagnosed with breast cancer.[37] Modifiable risk factors for breast cancer, such as physical activity and alcohol intake, also have been documented. Increased amounts of physical activity have been reported to result in slight decreases in the risk of breast cancer.[47, 50, 51] Alcohol is also a risk factor for breast cancer, with a 21% increase in breast cancer risk for women who consume two alcoholic drinks each day.[44] Despite its strong causal relationship with many other cancers, smoking does not appear to increase the risk of breast cancer. A meta-analysis of 53 studies with over 22,000 breast cancer

Table 2.1 Summary of known risk and protective factors for breast cancer*

Characteristic	Estimate of Effect	Study Design	Reference
Risk Factors			
Older age	IRR 2.09 age 50-54 versus 40-44; IRR 4.11 age 75-79 versus 40-44	Surveillance Data	Calculated from SEER incidence rates[4]
Caucasian race	IRR 1.16 for Caucasian versus African American; IRR 1.42 for Caucasian versus Asian/Pacific Islanders; IRR 1.57 for Caucasian versus Hispanic	Surveillance Data	Calculated from SEER incidence rates[4]
First degree relative with breast cancer	RR 2.1, 95% CI 1.6-2.8 for mother diagnosed before age 40 versus mother not diagnosed with breast cancer; RR 1.5, 95% CI 1.1-2.2 for mother diagnosed after age 70 versus mother not diagnosed with breast cancer; RR 2.3, 95% CI 1.6-3.4 for one sister with breast cancer versus no sisters with breast cancer	Prospective Cohort	Colditz (1993)[35]
Personal history of benign breast disease	RR 1.6, 95% CI 1.0-2.5 for proliferative disease without atypia; RR 3.7, 95% CI 2.1-6.8 for proliferative disease with atypical hyperplasia	Prospective Cohort	London (1992)[36]
Presence of BRCA1 or BRCA2 mutation	BRCA1: lifetime risk 50-73% by age 50; 65-87% by age 70 BRCA2: lifetime risk 59% by age 50; 82% by age 70	Review	National Cancer Institute (2005)[37]
Early age at menarche	RR 1.30 for age 12 at menarche versus age ≥15 at menarche, p trend <0.01	Case-control	Brinton (1988)[38]

Table 2.1 (continued)

Late age at menopause	RR 1.22 for age ≥55 at menopause versus age <45 at menopause, p trend =0.04 in analyses adjusted for interval between menopause and breast cancer diagnosis	Case-control	Brinton (1988)[38]
Later age at first birth	OR 1.65, 95% CI 1.40-1.93 for age ≥31 at first birth versus <18 at first birth; OR 1.20, 95% CI 1.16-1.24 per 5 year increase in age at first birth	Nested case-control	Lambe (1996)[39]
Oral contraceptive use	RR 1.24, 95% CI 1.15-1.33 for current users versus never users; RR 1.01, 95% CI 0.96-1.05 for 10 years since last use versus never users	Meta-analysis	Collaborative Group on Hormonal Factors in Breast Cancer (1996)[40]
Postmenopausal hormone therapy use	HR 1.24, 95% CI 1.01-1.54 for estrogen + progestin users versus placebo; HR 0.80, 95% CI 0.62-1.04 for unopposed estrogen users versus placebo	Randomized controlled trials	Chlebowski (2003);[41] Stefanick (2006)[42]
Postmenopausal obesity	RR 1.07, 95% CI 1.02-1.11 per 4 kg/m^2 increase in BMI	Meta-analysis	van den Brandt (2000)[43]
Alcohol use	RR 1.06, 95% CI 1.00-1.11 for alcohol intake >12g/day versus never drinkers	Meta-analysis	Ellison (2001)[44]
Mammographically dense breasts	RR 4.64, 95% CI 3.64-5.91 for ≥75% density versus <5% density	Meta-analysis	McCormack (2006)[45]

Table 2.1 (continued)

Protective Factors			
Increased parity	OR 0.91, 95% CI 0.85-0.97 for 2 live births versus nulliparous; OR 0.67, 95% CI 0.57-0.80 for 4 live birth versus nulliparous; OR 0.90, 95% CI 0.88-0.91 per each additional live birth among parous women	Nested case-control	Lambe (1996)[39]
History of breastfeeding	Premenopausal: RR 0.78, 95% CI 0.66-0.91 for ever breastfed versus parous but never breastfed Postmenopausal: RR 1.04, 95% CI 0.95-1.14 for ever breastfed versus parous but never breastfed	Case-control	Newcomb (1994)[46]
Current physical activity	RR 0.78, 95% CI 0.62-1.00 for >40 MET-hours/week current physical activity versus none; p trend =0.03 for increasing MET-hours/week current physical activity	Prospective Cohort	McTiernan (2003)[47]
NSAID use	OR 0.80, 95% CI 0.73-0.87 for women using any NSAID vs. non-users	Meta-analysis	Khuder and Mutgi (2001)[48]

*Abbreviations used are incidence rate ratio, IRR; confidence interval, CI; hazard ratio, HR; relative risk, RR; MET, metabolic equivalent; NSAID, non-steroidal anti-inflammatory drugs

cases reported that the risk of breast cancer was not increased among ever smokers as compared to never smokers (RR 1.03, standard error 0.02).[52]

The current knowledge about risk factors for breast cancer has led to the development of models which are useful for predicting a woman's risk of the disease. The Gail Model is a statistical model that is used to predict a woman's 5-year and lifetime risk of breast cancer.[53-55] The model was developed using case-control data from the Breast Cancer Detection Demonstration Project. Breast cancer risk factors included in the model are: current age, age at menarche, age at first live birth, number of previous breast biopsies, previous pathological finding of atypical hyperplasia on biopsy, and number of first degree relatives with breast cancer.[54] This model allows for the prediction of individual probabilities for being diagnosed with breast cancer.[55] The Gail Model has been validated and shown to be useful among women receiving annual screening, although limitations of the model have been noted.[54, 56]

2.2.1 Endogenous and exogenous estrogen and breast cancer

Exposure to estrogen is believed to be the underlying cause of breast cancer. Estrogen is a female sex hormone which is required for a number of processes in the body, including normal breast development. Markers of exposure to endogenous hormones such as early age at menarche and late age at menopause have been found to increase breast cancer risk, while breastfeeding and increased parity have been consistently shown to decrease risk.[38, 39, 46, 49, 57] Increased levels of endogenous hormones have been implicated in breast cancer. Numerous studies have consistently demonstrated that increased levels of endogenous estrogen are related to increased risk of breast cancer in postmenopausal women.[58-62] For example, a meta-analysis

of nine prospective studies examining hormone levels in relation to postmenopausal breast cancer reported a two-fold increase (RR 2.00, 95% CI 1.47-2.71, p trend<0.001) in risk of breast cancer for women in the highest quintile of estradiol (E2) compared to those in the lowest quintile.[59] The association between E2 and premenopausal breast cancer, however, is far less clear. Estradiol levels fluctuate throughout the menstrual cycle, with peaks occurring towards the ends of both the follicular and luteal phases.[63] Some studies have reported similar positive associations between E2 and breast cancer among premenopausal women,[62, 64-66] while others have reported no association.[67-73] Studies of E2 and premenopausal breast cancer have been limited by a number of factors, however, including small numbers,[62, 65, 67, 70, 71, 73] failure to control for phase of the menstrual cycle,[65, 72] and inclusion of cases that were premenopausal at the time of the blood sample but not at the time of breast cancer diagnosis.[65, 68] The largest and most recent study, a nested case-control study of 197 cases and 394 matched controls from the Nurses' Health Study II, reported that free E2 (RR 2.4, 95% CI 1.3-4.5 for 4^{th} vs 1^{st} quartile) and total E2 (RR 2.1, 95% CI 1.1-4.1 for 4^{th} vs 1^{st} quartile) levels during the follicular phase were positively associated with breast cancer, while free and total E2 levels during the luteal phase were not.[64] Though this study was prospective, carefully controlled for phase of the menstrual cycle, and employed large numbers, the menopausal status of the cases at the time of diagnosis was unclear.

Exposure to exogenous estrogen has been related to breast cancer risk. Use of oral contraceptives has been shown to slightly increase the risk of breast cancer, primarily among women who are current or recent users.[40, 49] Use of postmenopausal hormone therapy (HT) has been shown to increase risk of breast cancer by 10-80% depending on the duration of use.[41, 74, 75] The results of the Women's Health Initiative, however, showed that this increased risk may

occur only among users of combined estrogen and progestin regimens[41, 75] and not among women using unopposed estrogen.[42, 76] Women randomized to take a combined estrogen and progestin pill had a 24% increase in risk of invasive breast cancer compared to those randomized to placebo (HR 1.24, 95% CI 1.01-1.54),[41] while, in a separate study, women randomized to an unopposed estrogen pill had a similar risk of invasive breast cancer as women randomized to placebo (HR 0.80, 95% CI 0.62-1.04).[42]

2.2.2 Obesity and breast cancer

Obesity has emerged as a significant risk factor for postmenopausal breast cancer and is possibly a protective factor for premenopausal breast cancer. Further, adjustment for measures of obesity attenuates, but does not eliminate, the racial difference in stage at breast cancer diagnosis.[77, 78] The most frequently used measure of obesity is the body mass index (BMI). BMI is a measure of weight for height and is calculated as weight in kilograms divided by the square of height in meters. The World Health Organization has developed the following BMI categories: underweight ($<18.5 kg/m^2$), normal weight (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2), and obese ($\geq 30.0 kg/m^2$).[79]

2.2.2.1 Obesity and postmenopausal breast cancer

A striking difference in the effect of obesity on breast cancer risk appears when analyses are conducted separately among pre- and postmenopausal women. Among postmenopausal women, some studies report either no association or only a weak association between BMI and breast cancer risk,[17, 80-85] while the vast majority report that increased BMI significantly raises the risk of breast cancer.[43, 85-98] For example, a large case-control study reported a 4% (OR 1.04, 95% CI

1.03-1.04) increase in the odds of postmenopausal breast cancer for every 1 kg/m^2 increase in current BMI.[95] A meta-analysis of prospective cohort studies found that the risk of breast cancer increased 7% with each 4 kg/m^2 increase in BMI among postmenopausal women (RR 1.07, 95% CI 1.02-1.11).[43] Some studies have reported that the positive association between BMI and postmenopausal breast cancer risk occurs only or more strongly among women with certain other risk factors, such as a family history of breast cancer[97] or older age.[91, 99] A consistent finding is that elevated BMI increases the risk of postmenopausal breast cancer only among women who have never used HT.[90, 92, 100-103] For example, a study of postmenopausal women enrolled in the Women's Health Initiative Observational Study reported no association with BMI among HT users (ever or current), but that the risk of breast cancer was more than doubled among obese women who had never used HT (RR 2.52, 95% CI 1.62-3.93 for BMI ≥31.1 vs. ≤22.6 kg/m^2).[102]

Studies have also considered other measures of anthropometry in relation to risk of postmenopausal breast cancer. Increased weight was positively associated with postmenopausal breast cancer in some studies[43, 92, 95, 96, 104, 105] but was not associated with breast cancer in others.[81-85] This association between increased weight and postmenopausal breast cancer risk may only occur among women who have never used HT.[102] Central adiposity, commonly measured by waist circumference or waist-to-hip ratio, has been positively associated with postmenopausal breast cancer,[106, 107] and one study reported that this effect was stronger in women who never used HT.[106] Finally, multiple studies have reported that weight gain during adulthood increases postmenopausal breast cancer risk[84, 92, 94, 95, 97, 103, 108-110] while weight loss can reduce this risk.[110, 111] Thus it is well-documented that obesity increases breast cancer risk among postmenopausal women.

2.2.2.2 Obesity and premenopausal breast cancer

On the contrary, obesity appears to have the exact opposite effect on breast cancer risk among premenopausal women. Few studies report either a positive association[87] or no association[83, 86, 89, 96, 98, 112] between BMI and premenopausal breast cancer. Many studies, however, have reported that BMI is inversely associated with premenopausal or early-age breast cancer risk.[43, 80, 84, 85, 93, 103, 113] For example, the same meta-analysis that reported a positive association between BMI and postmenopausal breast cancer risk reported a significant negative association between BMI and premenopausal breast cancer risk, with an 11% reduction in risk for every 4 kg/m^2 increase in BMI (RR 0.89, 95% CI 0.81-0.97).[43] The effect of BMI on premenopausal breast cancer risk may vary by race, with one study reporting a negative association among Caucasian women but no association among African American women.[17]

Similar relationships between obesity and premenopausal breast cancer risk are observed when other anthropometric measures are considered. Weight has been reported to either be negatively associated[43, 84, 85, 112-114] or not associated[83, 96, 104, 105] with premenopausal breast cancer. One study reported a positive association between waist-to-hip ratio and risk of premenopausal breast cancer,[107] while another reported no association.[106] The effect of weight gain on premenopausal breast cancer may also vary by race, with studies of Caucasian women reporting either no[103, 108] or a negative association,[112] while a study of Hispanic women reported a non-significant positive association.[108] Overall, the totality of the current evidence suggests that obesity reduces the risk of premenopausal breast cancer.

2.2.2.3 Mechanisms relating obesity to breast cancer risk

At first consideration, it appears counterintuitive that a single risk factor could impart such opposite effects on the risk of one disease depending on the menopausal status of the woman.

Consideration of the biological mechanisms which may explain the associations between obesity and pre- and postmenopausal breast cancer, however, clarifies why such a paradoxical effect of obesity exists. Both effects appear to be related to changes in endogenous hormone exposure which occur among obese pre- and postmenopausal women. As noted in a review by Key et al.,[115] in premenopausal women, obesity increases the number of anovulatory menstrual cycles. While this may have only a slight effect on altering estrogen levels due to homeostatic control mechanisms, progesterone levels are substantially decreased in anovulatory cycles.[115] Progesterone has been implicated in cancer risk, and decreased progesterone may thus decrease the risk of breast cancer among premenopausal obese women.[115] Further, in postmenopausal women the ovaries no longer produce estrogen or progesterone, and levels of these sex hormones are significantly reduced relative to levels observed in premenopausal women.[115] The primary source of estrogen in postmenopausal women is through the conversion of androstenedione to estrone catalyzed by aromatase which occurs in the adipose tissue, including the adipose tissue of the breast.[115, 116] Estrogen levels among postmenopausal women have a strong, positive relationship with obesity,[116] and this increased estrogen may promote carcinogenesis among postmenopausal women.

In both pre- and postmenopausal women, sex hormone binding globulin (SHBG) is inversely associated with BMI.[117] SHBG can bind estrogen and thereby reduce the pool of bioavailable estrogen. In premenopausal women the body manages to maintain homeostatic control of estrogen levels such that the impact of decreased SHBG is minimal, while in postmenopausal women these controls do not exist and decreased levels of SHBG result in substantially increased levels of bioavailable estrogen.[115] Thus this increased endogenous estrogen exposure among obese postmenopausal women may confer an increased risk of breast

cancer. Other possible mechanisms to explain associations between obesity and breast cancer risk include effects of hyperinsulinemia as well as mechanisms involving leptin and adiponectin.[116] These mechanisms are currently less well-understood in relation to breast cancer etiology and do not appear to account for the divergent effects of obesity on the risk of pre- and postmenopausal breast cancer.

2.3 MAMMOGRAPHIC BREAST DENSITY

It is generally accepted that mammography represents the best opportunity for early detection of breast cancer. Current American Cancer Society guidelines state that women over the age of 40 should receive annual mammograms as a means of screening for breast cancer.[2] Although mammography is most often used to look for signs of tumors or other breast abnormalities, it is believed that the composition of the breast may also yield information about a woman's risk of breast cancer. The breast is composed of different types of tissues, and the composition varies from woman to woman. The primary functional units of the breast are the lobules, which are connected to the nipple through a series of ducts; these structural and functional units are composed of epithelial tissue. The majority of the breast is composed of fat tissue, except during lactation when the breast consists primarily of glandular units. After menopause the number of lobules substantially decreases and the amount of fat tissue in the breast increases.[118] Fat appears dark on a mammogram because it is radiologically lucent, while epithelial and connective tissues appear bright because they are radiologically dense.

2.3.1 Methods of measuring mammographic breast density

There are a variety of methods that have been developed to measure mammographic breast density. These methods can be grouped according to whether they represent qualitative or quantitative measurements.

2.3.1.1 Qualitative methods

Wolfe first hypothesized that the distribution of different types of tissue in the breast may be related to risk of breast cancer.[119, 120] He proposed a classification system of breast density in which the parenchymal pattern was categorized in order of risk of breast cancer as N1 (primarily fat tissue), P1 (mostly fat tissue but with some dense areas of less than 25% of the total breast), P2 (more than 25% of the breast composed of dense tissue along with a noticeable ductal pattern), and DY (primarily homogeneous dense tissue and no conspicuous ductal pattern).[120, 121] Although these "Wolfe patterns," as they are now called, are a qualitative measure of breast density and are, therefore, somewhat subjective, the P2 and DY patterns have been repeatedly linked to an increased risk of breast cancer as compared to the N1 and P1 patterns.[120-123]

Subsequent to the development of the Wolfe patterns, other qualitative methods have been proposed. The Tabar classification describes five patterns of breast density "based on anatomic-mammographic correlation using three-dimensional, sub-gross (thick-slice) technique."[124] Pattern I is considered low risk and is characterized by scalloped contours and Cooper's ligaments, terminal ductal lobular units that are evenly scattered, and oval-shaped areas where fatty replacement of tissue has occurred. Pattern II demonstrates complete fatty replacement and Pattern III shows a prominent duct pattern in the retroareolar area; both Patterns II and III also are considered low risk patterns. Patterns IV and V are considered to be high risk;

Pattern IV is characterized by nodular and linear densities throughout the breast, and Pattern V is described as having extensive fibrosis with an appearance similar to ground glass.[124] The first paper reporting the Tabar classification showed that agreement between the Wolfe method and Tabar method in terms of classifying high-risk versus low-risk mammograms was poor.[124] This appeared to be largely because a large proportion (45.6%) of the evaluated mammograms were classified as Wolfe DY (high risk) yet Tabar Pattern I (low risk).[124] The creator of the Tabar method performed both the Wolfe and Tabar assessments in this study, however, and this may have led to bias in the measurements.

A third commonly used qualitative method for assessing breast density is the Breast Imaging Reporting and Data System (BI-RADS). This method is the most commonly used method of assessing density by radiologists. BI-RADS recognizes four class of density: 1) breast is composed almost entirely of fat, 2) breast contains scattered fibroglandular densities, 3) breast is heterogeneously dense, and 4) breast is extremely dense.[125] The BI-RADS score is typically reported by the radiologist when reviewing mammograms.[45, 125]

Although each method has been well-described, the Wolfe, Tabar, and BI-RADS systems are all qualitative methods and are therefore prone to a high-degree of subjectivity and potentially lower reproducibility. In fact, studies evaluating the reproducibility of these three methods between raters have reported only moderate reproducibility.[126-128] For example, one study of Tabar patterns reported an intra-rater reliability of κ=0.65.[127] A study of the reliability of Wolfe pattern assessment reported an intra-observer intraclass correlation coeffiecient (ICC) of 0.68 and an inter-observer ICC of 0.65,[126] indicating good reliability. Finally, a study of inter-observer agreement of BI-RADS density assessments reported an overall reliability of κ=0.43,

with poor agreement for the "extremely dense" category (κ=0.17) and highest agreement for the "fatty" category (κ=0.76).[128]

2.3.1.2 Quantitative methods

More recently, methods have been developed which allow for quantitative measurement of breast density. These studies report percent density, which is calculated as the percentage of the breast comprised of dense areas. One method of quantitative assessment is manual planimetry. This method is performed by using a wax pencil to trace the total area of the breast and all areas of density (excluding biopsy scars, Cooper's ligaments, and breast masses) onto a clear acetate sheet placed over the mammogram. A compensating polar planimeter is then used to measure the total area of the breast and the area of breast density.[129-131] Computerized planimetry can also be used in a similar manner. This method uses either digital mammograms or film mammograms that have been digitized. The total breast area and areas of density are outlined on the computer using a mouse or other tracing device, and then the respective areas are calculated by the computer.[132, 133] Intra-reader reliability is reported to be high, with intra-class correlation coeffiecients of 0.93 for percent density, 0.82 for absolute dense area, and 0.97 for non-dense area.[133]

A third method of quantitative density assessment is that of computerized thresholding. This technique can be used with either digital mammograms or film mammograms that have been digitized. The mammogram is viewed on the computer and the reader first selects a "threshold brightness" that distinguishes the breast tissue from the background of the mammogram. Next the reader chooses another "threshold brightness" which differentiates between dense and non-dense tissue. The computer uses these thresholds to identify both the total area of the breast and areas of density, and the number of pixels within these areas are

summed to give a measure of the total breast area and the dense area.[134-138] The intra- and inter-reader reliability of density measurements using the computerized thresholding technique have been shown to be very high, both having ICCs >0.9.[138]

2.3.2 Mammographic breast density and risk of breast cancer

Numerous studies have investigated associations between breast density and breast cancer since Wolfe hypothesized such a relationship over thirty years ago. Recently, McCormack and dos Santos Silva performed a meta-analysis of such studies.[45] These authors conducted a systematic review with well-defined search criteria to identify all published studies examining associations between mammographic breast density and breast cancer risk, including various methods of density measurement. Their search strategy and eligibility criteria resulted in the identification of 42 articles which were further grouped into incidence studies, prevalence studies, and studies in symptomatic populations and by the type of density measurement used. Overall, the results of the meta-analysis showed a high degree of consistency among the identified studies. The combined relative risk from incidence studies of the general population using Wolfe patterns was 1.76 (95% CI 1.41-2.19) for P1 versus N1, 3.05 (95% CI 2.54-3.66) for P2 versus N1, and 3.98 (95% CI 2.53-6.27) for DY versus N1. These point estimates are slightly higher than those calculated using prevalence studies of the general population: 1.25 (95% CI 1.02-1.54) for P1 versus N1, 1.97 (95% CI 1.29-3.00) for P2 versus N1, and 2.92 (95% CI 1.98-2.97) for DY versus N1. Combined relative risk estimates of the two studies using the BIRADS classification with fatty breast as the referent group were 2.04 (95% CI 1.56-2.67) for scattered density, 2.81 (95% CI 2.13-3.71) for heterogeneously dense, and 4.08 (95% CI 2.96-5.63) for extremely dense. Only one study used the Tabar classification, thus a combined relative risk was not

provided;[45] this study, however, reported that the odds of breast cancer were substantially increased among women with Tabar pattern IV versus Tabar pattern I (adjusted OR 2.42, 95% CI 0.98-5.97).[139]

Similar combined estimates of relative risks using quantitative percent density assessments were also reported. Compared to having less than 5% breast density, incidence studies had combined relative risks of 1.79 (95% CI 1.48-2.16) for 5-24% density, 2.11 (95% CI 1.70-2.63) for 25-49% density, 2.92 (95% CI 2.49-3.42) for 50-74% density, and 4.64 (95% CI 3.64-5.91) for ≥75% density. The combined relative risk estimates for prevalence studies were similar but slightly lower: 1.39 (95% CI 1.10-1.76) for 5-24% density, 2.22 (95% CI 1.75-2.81) for 25-49% density, 2.93 (95% CI 2.27-3.79) for 50-74% density, and 3.67 (95% CI 2.72-4.96) for ≥75% density versus <5% density.[45]

The results of this meta-analysis demonstrate a number of important findings regarding breast density. First, there is clearly strong evidence of an association between breast density and risk of breast cancer. This finding is consistent regardless of the methods employed to measure breast density and whether the studies are performed using incident or prevalent cases. Second, the results show that masking bias, in which cancers are "masked" by high breast density, is relevant to the study of breast density and breast cancer. Presence of masking bias would result in underestimated relative risks reported by prevalence studies, in which cancers were detected at the time of screening, and overestimated relative risks reported by incidence studies; this is consistent with the results of the present meta-analysis. Relative risk estimates were also lower after excluding cancers diagnosed in the year after the density measurement. Comparing women with ≥75% density to those with <5% density, the meta-analysis reported combined relative risks of 4.64 (95% CI 3.64-5.91) for all eligible studies, 4.52 (95% CI 3.54-

5.78) among studies excluding cancers diagnosed in the first year, and 13.38 (95% CI 2.73-66.6) among studies including cancers diagnosed in the first year.[45] Third, the review and meta-analysis reported that breast density remains associated with breast cancer risk regardless of age, menopausal status, or race.[45]

This meta-analysis included studies indexed in the Medline, EMBASE, and Pubmed databases on November 18, 2005. Therefore, the same search strategy employed in this meta-analysis was used to identify additional relevant articles published after this date through January 31, 2007. The search strategy used is well-described by McCormack and dos Santos Silva.[45] Briefly, the Medline, EMBASE, and PubMed databases were searched using keywords related to cancer, mammography, and breast density and results were limited to English language journal articles. To update the literature review, the search was further restricted to articles published in 2005-2007. This search resulted in the identification of one additional article evaluating breast density in relation to breast cancer risk.[140] Mitchell et al. reported that higher percent breast density remains a strong risk factor for breast cancer among women with known BRCA1/BRCA2 mutations. The odds of breast cancer among mutation carriers with density greater than or equal to 50% were twice that of mutation carriers with less than 50% density (OR 2.29, 95% CI 1.23-4.26).[140]

As noted previously, mammographic breast density may be a stronger risk factor for postmenopausal breast cancer than for premenopausal breast cancer.[6] For example, a nested case-control study of 1,717 pre- and 1,208 postmenopausal women reported that breast cancer risk increased with increasing percent breast density, but that the point estimates were higher among postmenopausal women. Compared to a premenopausal woman with 0% breast density, a premenopausal woman with 1-24% breast density had an odds ratio of 1.47 (95% CI 0.95-2.3)

and one with ≥75% density had an odds ratio of 3.79 (95% CI 2.3-6.2), while a postmenopausal woman had odds ratios of 1.79 (95% CI 1.3-2.5) for 1-24% density and 5.82 (95% CI 3.0-11.3) for ≥75% density compared to a postmenopausal woman with 0% density.[141]

Changes in breast density have also been related to subsequent changes in risk of breast cancer. One study reported that women who consistently had high-risk Wolfe patterns (P2 or DY) had over twice the risk of breast cancer (RR 2.2, 95% CI 1.2-3.9) as compared to women who consistently had low-risk Wolfe patterns (N1 or P1). Those women whose patterns on the first mammogram were either P2 or DY but then had a low-risk pattern on a subsequent mammogram had similar risk to women with consistently low-risk Wolfe patterns (RR 1.2, 95% CI 0.5-2.8).[142] It appears that breast density may be useful as a biomarker of breast cancer risk. In fact, breast density has been used as an intermediate or surrogate endpoint in at least two intervention trials investigating the effects of diet on breast cancer risk.[136, 143] Though the ability of breast density to change in response to known risk factors for breast cancer, such as use of postmenopausal hormone therapy, has been established (see section 2.4.3.2), it is unclear what these changes in breast density mean in terms of altering breast cancer risk.[144] To date few studies have reported on how changes in breast density relate to changes in breast cancer risk. One case-control study that examined this issue reported no statistically significant associations between change in percent breast density and breast cancer risk.[145] Further studies are reported to be currently in progress.[144]

2.3.3 Mammographic breast density and breast cancer risk factors

In general, risk factors for breast cancer also increase mammographic breast density. For example nulliparity and later age at first birth have been associated with increased density.[146] On

the other hand, breast density has been shown to decrease with increasing age,[146-149] although increased age is a risk factor for breast cancer. This apparent contradiction has been explained by noting that breast density may be related to the rate of change in breast cancer incidence rather than the incidence of breast cancer itself.[6] The menopausal transition may more strongly influence changes in breast density than age, however. A study of women who were premenopausal at a baseline mammogram and then postmenopausal at a subsequent mammogram and were matched on age to women who remained premenopausal at both mammograms showed that percent density decreased more among the women who transitioned through menopause than those who did not.[150] Age may not be related to breast density among women over age 70, however.[151] A study of 239 participants from the Study of Osteoporotic Fractures (SOF) reported that only BMI, parity, surgical menopause, and current smoking status were significantly associated with mammographic breast density in multivariable analyses. The mean age of these women at the time of consent for obtaining the most recent mammogram was 78.6 (SD 3.8), thus indicating that the factors associated with breast density among older women may differ from premenopausal and younger postmenopausal women.[151]

It appears that there is a genetic component to mammographic breast density. Boyd et al. reported the results of two twin studies, showing correlation coefficients for percent breast density of 0.61 and 0.67 for monozygotic twins and correlation coefficients of 0.25 and 0.27 for dizygotic twins.[152] In these studies genetic factors explained 60-75% of the variability in percent breast density.[152] Specific genes responsible for differences in breast density have yet to be conclusively identified, however.

Despite the relationships between breast cancer risk factors and breast density, presence of other breast cancer risk factors does not fully account for the association between increased

breast density and risk of breast cancer.[6] Mammographic breast density is indeed an independent risk factor for breast cancer.

2.3.3.1 Endogenous hormones and mammographic density

In general, studies have not shown statistically significant associations between levels of endogenous hormones and mammographic breast density. The previously described breast density study from SOF reported no associations between percent breast density and total estradiol, estrone, free testosterone, total testosterone, dehydroepiandrosterone (DHEA) sulfate, or SHBG.[151] Another study of postmenopausal women reported no associations between percent breast density and estrone, estradiol, free estradiol, testosterone, free testosterone, androstenedione, DHEA sulfate, SHBG, or follicle stimulating hormone (FSH) in women who had never used HT.[153] This study did, however, report moderate negative associations between estrone, estradiol, free estradiol, testosterone, free testosterone, androstenedione, DHEA, and FSH among women who were former users of HT and had used HT within the previous 5 years.[153] It is important to note that this study population consisted of overweight, nonsmoking, postmenopausal women who were currently non-HT users enrolled in a separate randomized controlled trial, and who had a mammogram within 12 months prior through 1 month after enrollment; thus, this highly selected population may not be generalizable to other populations of women. Further, only 88 women comprised the study population.[153] Boyd et al.[154] reported slight, negative associations between free estradiol and both percent density and dense area in postmenopausal women after adjustment for age and waist circumference. They also reported a very small positive association between SHBG and percent density and dense area among these women that remained after adjustment for age and waist circumference. Total estradiol and progesterone were not significantly associated with either percent density or dense area after

adjustment.[154] An important limitation of this study, however, is that the blood samples were not taken at the time of mammography, but were 32 weeks after the date of mammography on average.[154]

A similar lack of associations has been observed in premenopausal women as well. In the above-referenced study, Boyd et al. also demonstrated that no association between free estradiol, SHBG, or progesterone and either percent density or dense area was apparent in premenopausal women after adjustment for age and waist circumference.[154] Another study of premenopausal women reported that progesterone, SHBG, estrone, total estradiol, and free estradiol were not associated with either percent breast density or dense area after adjustment for age, body weight, height, ethnicity, age at menarche, parity, and age at first birth.[155] The timing of the blood draw in relation to the mammogram was unclear in this study, although care was taken to standardize the blood collection by menstrual cycle phase.[155]

2.3.3.2 Exogenous hormones and mammographic density

Studies have repeatedly shown that increased breast density is related to use of postmenopausal hormone therapy (HT).[156-172] The percent of women whose density changes after initiating HT varies by type of HT used, however, with increased density occurring more often in estrogen plus progestin regimens than with estrogen alone regimens.[156-158, 162, 163, 165, 166, 168-170] In a sub-study of the WHI, investigators reported that 75% of women on active treatment experienced an increase in breast density after 1 year. The mean change in percent density from baseline to year 1 was 6.0% (95% CI 4.6 – 7.5) in the treatment group compared to -0.9% (95% CI -1.5 – -0.2) in the placebo group.[172] Studies have also reported that short-term cessation of HT use prior to mammography results in a decrease in breast density[173] or less frequent increase in density

compared to women who continue to take HT,[165] although even after more than 2 months of cessation, there appear to be residual effects of HT on density.[174]

Data on the effect of oral contraceptive use on breast density are limited, likely because the majority of women for whom screening mammography is recommended (age ≥40) are postmenopausal and would not be currently using oral contraceptives. One study has reported, however, that use of oral contraceptives prior to the first birth was not related to breast density later in life.[175]

2.3.4 Anthropometry and mammographic density

As previously noted, increased BMI results in an increased risk of breast cancer among postmenopausal women (see section 2.3.2.1). Studies of mammographic breast density, however, consistently report that increased weight or BMI is associated with lower percent breast density.[127, 133, 175-181] For example, a study of pre- and postmenopausal women reported that the difference in percent density between the 3rd and 1st quartiles of BMI was a 5.2 percentage-point decrease among premenopausal women and a 4.7 percentage-point decrease among postmenopausal women.[178]

While one study evaluating the possibility of interactions between weight or BMI and breast density on risk of breast cancer found these interactions to be non-significant,[180] others have noted significant effect modification. Ursin et al. reported a U-shaped relationship, with women having the lowest and the highest BMI demonstrating the strongest association between breast density and risk of breast cancer.[182] Duffy et al. also reported that when evaluating the relationship between high-risk Tabar patterns and breast cancer risk, only those women who were both overweight and had dense breasts showed an increased risk of breast cancer (OR 2.30,

95% CI 0.98-5.40 for women with BMI >25 kg/m^2 and dense breasts compared to those with BMI <25 kg/m^2 and non-dense breasts).[183] Most recently, Boyd et al. investigated associations between anthropometry, breast density and breast cancer risk in a case-control study of pre- and postmenopausal women.[184] Their results showed that BMI was not significantly associated with breast cancer in either pre- or postmenopausal women prior to adjustment for breast density. Additional control for percent breast density resulted in increased and statistically significant associations with breast cancer overall and for postmenopausal women.[184] Further, the association between BMI and breast cancer among premenopausal women was strengthened and made positive with adjustment for percent density, yet still remained non-significant.[184] Analyses relating percent breast density to breast cancer risk showed increased odds ratios for this relationship after additional adjustment for BMI. These findings led the authors to conclude that anthropometry and breast density are confounders of one another in relation to breast cancer risk. Further, the authors suggested that failure to adjust for breast density in previous studies may explain the negative associations between BMI and breast cancer that are commonly reported among premenopausal women.[184]

Observed differences in mammographic density by race/ethnicity may also be explained by racial and ethnic differences in body size.[134] Indeed, body size does confound the relationship between breast density and breast cancer risk and therefore anthropometric measures must be taken into account in studies of breast density.[176, 177, 180]

The amount of fat present in the breast is strongly related to BMI.[185] Recent studies have therefore evaluated the effects of anthropometric measures separately on the size of the dense and non-dense areas of the breast.[133, 179, 186] These studies have all shown that weight and BMI are positively associated with the size of the non-dense area.[133, 179, 186] Two studies report

negative correlations between weight or BMI and the size of the dense area,[179, 186] while the third reports a positive, yet non-significant, association.[133] Boyd et al. reported the correlation between BMI and the size of the dense area to be -0.191 (p=0.002). These authors also noted that weight and height differences explained significant amounts of the variance in size of the non-dense area but not of the size of the dense area.[179] After controlling for non-dense area, both BMI and weight had weakly positive associations with dense breast area (both p=0.01).[179] Haars et al. reported that BMI alone explains 40% of the variance in size of non-dense tissue, yet only 17% of the variance in size of dense tissue.[133] Both groups of researchers support the use of absolute measures of density, rather than a relative measure such as percent density, when studying mammographic breast density as an indicator of breast cancer risk.[133, 179] No prospective study, however, has been specifically designed to evaluate the effect of weight change on the size of the dense and non-dense tissues of the breast.

The majority of studies that have investigated associations between anthropometry and mammographic breast density have employed a cross-sectional design. Therefore causality could not be established. A few studies have investigated how *change* in anthropometric measures affects breast density (Table 2.2). One study reported that the size of the dense area decreased among women who lost weight, although the average weight loss was quite small (0.3 kg).[136] A separate study reported that women who had gained weight in recent years had an increased risk of having a high-risk Wolfe pattern compared to women with a consistently elevated BMI throughout their lifetime.[181] Finally, a recent study found that percent density decreased more slowly among overweight and obese women compared to normal weight women, with overweight women having a 1.9 unit decrease and obese women having a 3.6 unit decrease in percent density per decade.[187] All three of these studies have significant

Table 2.2 Summary of studies investigating the impact of change in anthropometry on mammographic breast density

Author	Population	Study Design	Results	Other Comments
Boyd (1997)[136]	N=817 Mean age=46 Mean BMI=23kg/m^2	Randomized controlled trial of low fat, high carbohydrate diet vs. usual diet	Dense area decreases among women who lost weight (weight change β=98.8, p=0.04)	Change only over 2 years; very small change in weight (-0.3kg in intervention, +0.9kg in control)
McCormack (2003)[181]	N=1,298 Mean age=51 Middle quintile of BMI 25.16-27.57kg/m^2	Retrospective cohort	High-risk Wolfe patterns increased among women with larger increases in BMI between ages 43-53 (OR 1.46, 95% CI 1.27-1.68)	Used Wolfe patterns rather than a quantitative measure, only assessed one mammogram and closest BMI measure was an average of 2 years after the mammogram
Maskarinec (2006)[187]	N=1,274 Mean age=58.7 Mean BMI=24.9kg/m^2	Nested case-control	Overweight and obese women have more gradual decline in percent density per decade (-1.9% p=0.04, and -3.6% p=0.01, respectively) compared to normal weight women	Self-reported height and weight

methodological limitations, including limited weight change, self-reported anthropometric measurements, or use of a non-quantitative measure of breast density. Clearly the impact of changes in anthropometry must be understood, as failure to properly account for their influence on breast density could lead to incorrect analyses and conclusions of studies employing mammographic breast density as a surrogate endpoint. Prospective studies examining longitudinal associations between anthropometric measures and measures of mammographic breast density, however, have not been conducted.

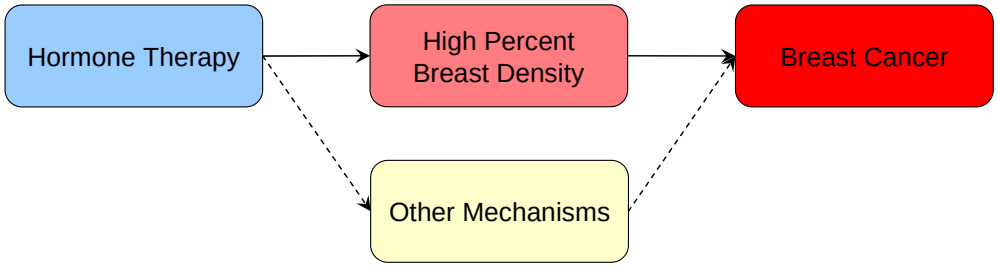
2.3.5 Mammographic density as a surrogate endpoint for breast cancer

A surrogate endpoint is a factor which may be measured and used as a substitute for the occurrence of a true endpoint, such as disease incidence, in observational and experimental studies.[188] Such surrogate endpoints have numerous advantages, including allowing for studies to be conducted with fewer subjects and at a lower cost and also aiding in the understanding of the mechanisms of cancer development.[188] Prentice has defined a surrogate endpoint as a factor for which testing the null hypothesis that an exposure is not related to the factor is equivalent to testing the null hypothesis that the exposure is not related to disease.[188] Schatzkin and Gail outlined the following criteria for evaluating the validity of a factor as a surrogate endpoint: 1) the surrogate measure must be associated with the true outcome, 2) the exposure of interest must also be associated with the surrogate measure, and 3) the entire effect of the exposure on the true outcome must be mediated by the surrogate measure.[189] Some have stated that these criteria are too stringent and have suggested modifications.[188] Most notably, it is rare that criterion #3 is perfectly met, as few surrogate endpoints are able to capture the entire effect of an exposure on a true outcome.[188, 189]

Mammographic breast density may be useful as a surrogate endpoint in studies of breast cancer. Indeed, some clinical trials have utilized changes in breast density as a surrogate endpoint rather than waiting to observe the true outcome of breast cancer.[136, 143] Breast density does in fact satisfy criterion #1, in that higher breast density has been consistently shown to be strongly associated with risk of breast cancer (see section 2.4.2). Whether or not the final two criteria are met for breast density, however, depends on the exposure being studied. It is possible that a surrogate endpoint may be valid for some exposures yet not for others, and therefore the endpoint must be validated separately for each exposure.[188, 189] As an illustration of this point, Figure 2.2 provides examples of exposures for which percent breast density may or may not be valid surrogate markers for breast cancer. Use of postmenopausal hormone therapy (Figure 2.2.A) is known to result in both increased breast density (see section 2.4.3.2) and increased breast cancer risk (see section 2.3.1); thus it satisfies both criteria #1 and #2. It is unlikely that breast density mediates the entire effect of hormone therapy on breast cancer (criterion #3), however, and other mechanisms are likely involved as well. It is rare that a surrogate endpoint fully satisfies criterion #3, however.[188, 189]

Boyd et al. recently evaluated whether breast density was a valid surrogate endpoint in studies of hormone therapy as a risk factor for breast cancer.[190] Their results showed that the association between hormone therapy use and breast cancer was only slightly attenuated after adjustment for breast density, thus failing to meet criterion #3. The authors concluded that breast density is not a valid surrogate endpoint in this case.[190] It should be noted, however, that hormone therapy use was not associated with breast cancer risk prior to adjustment for breast density either, although this association has been well-established by numerous previous studies.

A. Example of an exposure for which breast density may be valid as a surrogate end-point: Hormone therapy



B. Example of an exposure for which breast density is not a valid surrogate end-point: Postmenopausal obesity

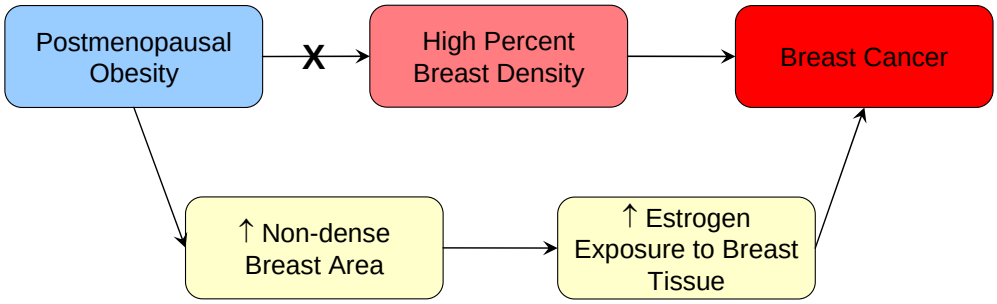


Figure 2.2 Illustration of the validity of mammographic breast density as a surrogate endpoint for breast cancer

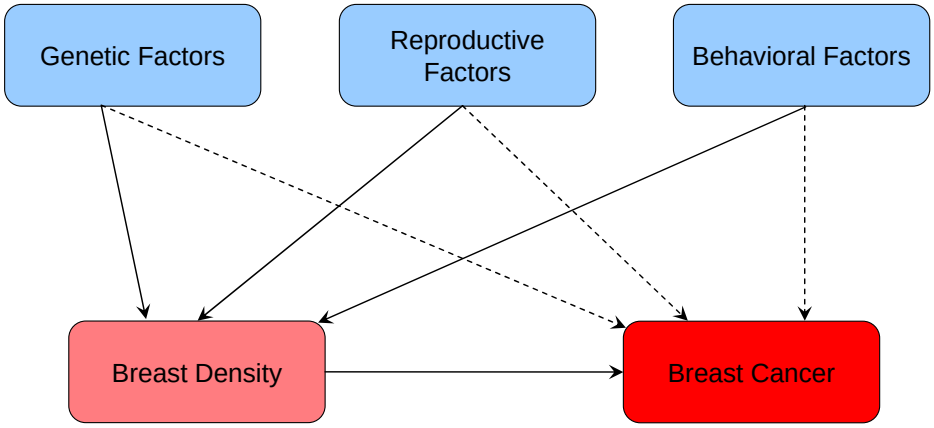


Figure 2.3 Proposed causal pathway for breast cancer

Further, information on the types of hormone therapy used was not available. This could explain the lack of significant findings if the study population was composed of users of combined estrogen formulations as well as users of unopposed estrogen formulations. As previously discussed, the Women's Health Initiative has shown that combined estrogen formulations resulted in increased risk of breast cancer while unopposed formulations did not.[41, 42, 75, 76] Thus, although this appears to be the first study to evaluate breast density according to the criteria for surrogate endpoints, the study did have some important limitations and cannot be considered conclusive without corroboration by future studies.

On the other hand, postmenopausal obesity (Figure 2.2.B) is known to result in decreased percent breast density (see section 2.4.4); thus the direction of the association between this exposure and breast density does not satisfy criterion #1. Therefore, breast density is unlikely to be valid as a surrogate endpoint for breast cancer in studies evaluating postmenopausal obesity (or other anthropometric measures). This lack of validity for the exposure of postmenopausal obesity may be explained by the fact that obese women have substantially increased non-dense breast areas. Percent breast density is thus decreased in obese women because the non-dense area of the breast is so large relative to the size of the dense area. An obese woman might have a higher dense area (and thus more tissue at risk for developing breast cancer) than a non-obese woman, however this fact would be obscured by studying only percent density. Therefore absolute measures of breast density, such as the dense breast area, are likely to be more informative and less confounded by anthropometry than are relative measures, such as percent density. Further, it is possible that for the exposure of obesity, the non-dense area of the breast may be highly relevant to breast cancer development. As shown in Figure 2.2.B, higher non-dense breast area, which represents adipose tissue, may result in increased estrogen exposure to

the nearby dense breast tissue due to the conversion of androstenedione to estrogen that occurs in adipose tissue.[133, 187] This increased estrogen exposure of the ducts and lobules where cancers arise may result in increased risk of breast cancer. Thus it appears that mammographic breast density may be a valid surrogate endpoint for some, but not all, relevant exposures.

2.3.6 Proposed causal model for breast cancer

A causal pathway for breast cancer relating breast density, genetic factors, reproductive factors, and behavioral factors is proposed in Figure 2.3. In this model, genetic factors (e.g. family history of breast cancer, BRCA1/2 mutations), reproductive factors (e.g. age at menarche, parity), and behavioral factors (e.g. hormone therapy use, anthropometry) exert some of their effects on breast cancer risk through a pathway involving breast density. Indeed, it has been suggested that mammographic breast density represents the cumulative exposure to estrogen over the lifetime.[191] Residual effects of these factors (represented by dashed arrows) still can impact breast cancer development, however, through pathways that do not include changes in breast density. Further, breast density has its own effect on breast cancer that is not entirely explained by its associations with genetic, reproductive, and behavioral factors. Thus, this model proposes that mammographic breast density is a step on a causal pathway to breast cancer, of which there are potentially many, yet is also an independent risk factor for breast cancer in its own right. This complex relationship thereby necessitates careful control of genetic, reproductive, and behavioral factors in studies examining breast density and breast cancer.

2.4 ANGIOGENESIS AND BREAST CANCER

Angiogenesis is the process by which the body forms new blood vessels, and in healthy adults occurs only in the female reproductive system. Folkman's work generated much of the initial interest in angiogenesis, demonstrating that the ability of tumors to grow is dependent on angiogenesis.[192, 193] Without vascularization, tumors are unable to grow beyond 1-2mm^3 in size.[194]

2.4.1 Roles of vascular endothelial growth factor (VEGF)

While angiogenesis is a tightly controlled biological process involving multiple factors, vascular endothelial growth factor (VEGF) has been identified as a primary promoter of angiogenesis.[195] VEGF is a potent endothelial cell mitogen, and does not act on other types of cells.[195] VEGF has a wide range of functions, including promoting endothelial cell mitogenesis and survival, increasing stromal degradation by promoting the expression of enzymes involved in these processes, and promoting vascular permeability.[196] These functions are carried out through interactions with three types of receptors for VEGF present on endothelial cells.[195] There are multiple types of VEGF proteins, including VEGF-A, VEGF-B, VEGF-C, and VEGF-D.[196] VEGF-A is the most common form of VEGF; for this reason, VEGF-A is commonly referred to simply as "VEGF," and from this point forward use of the acronym "VEGF" will refer to VEGF-A unless otherwise stated. VEGF-C has been shown to be present in tumors which metastasize to the lymph nodes, and VEGF-D appears to be present only in cases of inflammatory breast cancer.[195, 197]

VEGF-A occurs in six different isoforms due to alternative splicing of the VEGF gene.[198, 199] The isoforms contain 121, 145, 165, 183, 189, or 206 amino acids. The 121 and 165 isoforms are the most abundant and are secreted as a soluble protein, while the other isoforms remain in the extracellular matrix.[198, 199] An investigation of twenty-six existing breast cancer cell lines reported that five of the six isoforms were present in all cell lines, while the 206 isoform was not observed in any of the cell lines.[199] Further, while the expression levels of many of the isoforms were correlated with one another, no correlations were observed between expression of the various isoforms and clinicopathological features of the cell lines, such as stage or grade.[199] The distribution of the VEGF isoforms is preserved even when the cells are stressed by changes in oxygen concentration, pH, or glucose availability, suggesting that these isoforms possess distinct biological functions.[200] At least one study has reported that the 121 isoform promotes tumorigenesis more strongly in a mouse model than the other VEGF isoforms.[201] Further research is necessary to elucidate the differing roles for each of the VEGF isoforms.

2.4.2 Angiogenesis and cardiovascular disease

Angiogenesis is also recognized as playing an important role in other diseases, including psoriasis, diabetic retinopathy, rheumatoid arthritis, and atherosclerotic plaque formation.[202] Whether VEGF levels are increased among diabetics is unclear, with at least one study reporting no difference between diabetics and non-diabetics[203] and another reporting increased VEGF levels among diabetics with hypertension versus subjects with only hypertension.[204] Angiogenesis appears to be related to cardiovascular disease in general. One study reported that plasma VEGF levels among hypertensive individuals were positively correlated with 10-year risk of cardiovascular disease and cerebrovascular accident, as determined by Framingham

cardiovascular risk scores.[205] A separate study of men recruited from a urology clinic reported significant increases in risk of acute myocardial infarction (HR 3.36, 95% CI 1.35-8.41), stroke (HR 3.98, 95% CI 1.61-9.86), and death (HR 1.74, 95% CI 1.01-3.00) for men with serum VEGF levels greater than 500 pg/mL at enrollment versus those with VEGF levels below 300 pg/mL.[206] In one study, plasma VEGF levels were higher among subjects with a history of myocardial infarction compared to participants with a primary acute myocardial infarction and healthy controls.[207]

Despite these strong relationships, whether increased angiogenesis is the cause, or rather the result, of cardiovascular disease is unclear.[208] One animal study reported that VEGF mRNA expression was increased in rats genetically modified to become hypertensive versus wild-type rats, and these investigators suggested that VEGF expression was increased as a means of compensating for the hypertension.[209] Several human studies report that hypertensives have increased levels of VEGF compared to normotensive controls.[205, 210-213] Further, some investigators have reported that treatment to lower blood pressure also has the effect of decreasing VEGF levels.[205, 213, 214] The effect of VEGF on blood pressure is far less clear, however, when one considers the literature from other fields. For example, blood pressure has been demonstrated to increase among cancer patients treated with angiogenesis inhibiting drugs.[214, 215] One randomized trial of the angiogenesis inhibitor sorafenib for treatment of metastatic renal cell carcinoma reported that 43% of participants experienced hypertension.[216] Similarly, investigators have noted that women with preeclampsia, a condition of pregnancy characterized by hypertension and proteinuria, have substantially lower levels of VEGF than do healthy women.[215] The mechanisms by which VEGF may regulate blood pressure are clearly complex, and further research in this area is warranted.

2.4.3 *In vitro* and animal studies of VEGF and breast cancer

There is a vast literature reporting studies of VEGF and angiogenesis using *in vitro* and animal models of carcinogenesis. This literature review will primarily focus on those studies directly related to carcinogenesis of the breast. Angiogenesis can be induced in mice upon transplantation of tissue from biopsy samples of women with confirmed breast cancer.[217] Xenografts of tissue taken from healthy women undergoing breast reduction did not induce angiogenesis, demonstrating that the vascularization was caused by the tumor cells.[217] *In vitro* experiments have demonstrated that VEGF is necessary for the invasion of breast cancer cells, though not for their proliferation.[218, 219] VEGF is expressed in a variety of cell lines, as previously noted, and two studies reported that VEGF expression was highest in MDA-MB-231 cells compared to other cell lines.[220, 221] VEGF is necessary for the initial growth of tumors. Studies in which either VEGF or its receptors VEGFR-1 and VEGFR-2 are blocked show decreased tumor growth, size, and metastases in animal models.[222-224] In a mouse model using inoculation with T47-D breast cancer cells to induce carcinogenesis, Yoshiji et al. reported that VEGF was essential for initial growth but its expression was not important once the tumors were established and of large size.[225] Other angiogenic factors, including basic fibroblast growth factor (bFGF) and transforming growth factor-α (TGF-α), appear to be important for the continued growth of tumors after the initial growth for which VEGF is needed.[225]

Part of VEGF's role in promoting tumor growth may also be carried out through preventing apoptosis. Bachelder et al. reported that blocking expression of VEGF in MDA-MB-231 and MDA-MB-435 breast cancer cells resulted in a three-fold increase in apoptosis.[226] In fact, VEGF has been shown to promote expression of the anti-apoptotic protein Bcl-2. Pidgeon et al. demonstrated that MDA-MB-231 cells induced to express VEGF or supplemented with

exogenous VEGF overexpressed Bcl-2.[227] Further, treatment of the cells with antibody to VEGF decreased expression of Bcl-2 and resulted in increased apoptosis.[227] It is also possible that Bcl-2 may regulate VEGF expression. Biroccio et al. reported that MCF-7 breast cancer cells manipulated to overexpress Bcl-2 showed significantly increased expression of VEGF protein under hypoxic conditions as compared to the xenografts using control MCF-7 cells without altered Bcl-2 expression.[228] Additionally, these investigators reported that mouse xenografts using the Bcl-2 overexpressing MCF-7 cells were more vascularized and showed increased VEGF by immunohistochemistry than did the control MCF-7 cells.[228] Thus there does appear to be a relationship between VEGF, Bcl-2, and apoptosis, though the direction of this relationship remains unclear.

The role of VEGF has been studied in relation to numerous factors believed to be involved in breast carcinogenesis. *In vitro* and animal studies have reported that VEGF expression is correlated with cyclin I[229] and NF-κB[221] and is upregulated by both TGF-β1[220] and tumor necrosis factor-α (TNF-α).[230] VEGF levels have been shown to be increased via multiple pathways, including those involving extracellular matrix metalloproteinase inducer (EMMPRIN),[231] heparanase and Src,[232] and mitogen-activated protein kinase (MAPK) and phospatidylinositol-3-kinase (PI3K).[233] The chemokines CCL5 and CCL2 were reported to have no effect on VEGF expression,[234] though VEGF has been reported to regulate expression of the chemokine receptor CXCR4.[218] The protein ErbB2 (also known as HER-2/neu) is related to increased VEGF protein synthesis *in vitro*. Klos et al. reported that MDA-MB-435 cells modified to constitutively express ErbB2 had higher levels of VEGF protein synthesis *in vitro* and generated tumors with increased microvascular density in mice as compared to wild type MDA-MB-435 cells.[235] High VEGF levels were also observed in the SKBR-3 cell line which

has endogenous overexpression of ErbB2.[236] Further, the addition of 4D5, a monoclonal antibody specific to ErbB2 which is now known as trastuzamab or Herceptin, decreased VEGF protein production in these cells.[236] Indeed, the role of VEGF in angiogenesis and carcinogenesis appears to be complex and likely involves interaction with multiple pathways.

One of the strongest inducers of VEGF expression is the hypoxic environment of tumors. Bachelder et al. demonstrated that hypoxia significantly increased the expression of VEGF by MDA-MB-231 and MDA-MB-435 breast cancer cells and that cell survival under hypoxic conditions was dependent upon the actions of VEGF.[226] Scott et al. studied VEGF expression in seven different breast cancer cell lines (MDA-MB-231, MDA-MB-435, MDA-MB-453, SKBR-3, BT20, MCF-7, and T47-D) in response to perturbations in the cellular environment, such as hypoxia, lowered pH, hypoglycemia, and hormonal concentration.[200] All cell lines except for MDA-MB-435 were observed to significantly increase VEGF expression under hypoxic conditions, with inductions ranging from 1.7 times normoxic induction in MDA-MB-231 cells to a 6.9-fold induction in T47-D cells. Differences in the VEGF expression under hypoxic conditions were unrelated to differences in p53, ER status, or normoxic VEGF production among the cell lines.[200] Hypoxia is believed to increase VEGF expression through a signaling pathway involving HIF-1α.[200, 237] The hypoxia-induced expression of VEGF is further increased when BRCA1 is present.[237] The mechanism by which BRCA1 exerts this effect is believed to involve an interaction with HIF-1α, as BRCA1 had no effect on VEGF levels when the VEGF promoter contained a mutation blocking its interaction with HIF-1α.[237] Further experimentation implied that under hypoxic conditions BRCA1 is able to induce VEGF by blocking the degradation of HIF-1α normally carried out by the proteasome.[237] Additional work in understanding the mechanistic relationship between BRCA1 and VEGF is needed.

Hormonal regulation of VEGF may also be important to angiogenesis in breast cancer. Hyder et al. reported identifying two estrogen response elements (ERE) in the rat VEGF gene, with one ERE in each of the 5' and 3' untranslated regions.[238] Mouse models have shown that VEGF increases estrogen dependent tumor growth[239] and that estrogen increases both tumor and extracellular levels of VEGF.[240] Most studies using the estrogen-sensitive (expressing both ER-α and ER-β) breast cancer cell line MCF-7 have reported that estrogen increases VEGF mRNA and/or protein expression by these cells,[241-246] although one study found no effect[247] and another reported decreased VEGF expression induced by estrogen.[248] The increase in VEGF expression can be blocked by incubation of MCF-7 cells with the pure anti-estrogen ICI 182,780,[245] though this is not a consistent result.[246] Further, estrogen has been reported to decrease expression of soluble VEGFR-1 in MCF-7 cells.[242, 249] Estrogen has been shown to have no effect on MDA-MB-231 cells, which expresses only ER-β,[241, 243-245] or in the estrogen-insensitive cell line T47-D.[200] Estrogen can also increase VEGF expression in estrogen-sensitive ZR-75 cells[250] and in SKBR-3 cells, in which a variant of ER-α was identified.[251]

The effects of tamoxifen, a selective estrogen-receptor modulator, on VEGF are far more controversial than those of estrogen. Some studies report that tamoxifen treatment increases VEGF expression in MCF-7 cells,[245, 248] while others report that VEGF expression in MCF-7 cells is decreased by tamoxifen.[242, 243, 252] Takei et al. reported that while tamoxifen had no effect on VEGF protein expression at concentrations ≤1 μM, a concentration of 1 μM showed a slight increase in VEGF protein and a concentration of 10 μM showed a strong increase.[246] Similar to the *in vitro* results, while one study reported that tamoxifen increased both vascular permeability and VEGF protein in a mouse breast cancer model,[248] another reported that extracellular VEGF was lower in tumors of mice treated with estrogen and tamoxifen as compared to those in mice

treated with estrogen alone.[252] Levels of soluble VEGFR-1 were reported to increase in MCF-7 cells treated with tamoxifen, with an approximately nine-fold increase in the ratio of soluble VEGFR-1:VEGF after tamoxifen treatment.[242] One study using the MDA-MB-231 cell line reported that tamoxifen had no effect on levels of VEGF expression.[243] Tamoxifen had no effect on VEGF expression in GI-101A cells compared to untreated cells.[253] The synthetic estrogen diethylstilbestrol (DES) was shown to increase VEGF expression in these cells, however, and the addition of tamoxifen was able to block this DES-induced VEGF expression.[253]

Most studies exploring the effects of progestins in the progestin-sensitive T47-D breast cancer cell line report that progestin exposure increases VEGF expression,[247, 254, 255] though at least one study has found no effect.[200] Both natural and synthetic progestins are able to increase VEGF expression *in vitro*, and the synthetic progestin medroxyprogesterone acetate (MPA) is reported to have the strongest effect on VEGF expression by T47-D cells.[247, 254] MPA is a synthetic progestin commonly used in postmenopausal hormone therapy, and Hyder et al. suggest that perhaps the strong effect of MPA on VEGF expression at least partially explains the increased breast cancer risk observed in women using estrogen + progestin preparations versus those using estrogen alone.[254] Though more evidence is needed for this hypothesis, it is interesting to note that MPA was the synthetic progestin used in the Women's Health Initiative clinical trial of hormone therapy.[75] Further evidence for an effect of progestins on VEGF expression comes from studies showing that the anti-progestin RU-486 blocks the progestin-induced expression of VEGF[254] and from studies reporting no effect of progestin on VEGF expression in breast cancer cell lines known to be progestin-insensitive.[238] Interestingly, the breast cancer drug faslodex (also known as ICI 182,780), an anti-estrogen drug which is used in women resistant to tamoxifen, has been reported to block VEGF mRNA and protein expression

through anti-progestin actions.[256] Although further research is needed to fully describe the hormonal regulation of VEGF expression, it is quite clear that both estrogens and progestins, as well as their antagonists, play a substantial role in regulating VEGF and, therefore, angiogenesis.

2.4.4 Human studies of VEGF and breast cancer

The following sections review studies of the role of VEGF in breast cancer that have been conducted using human subjects and/or stored tumor tissue samples from breast cancer patients.

2.4.4.1 VEGF and breast cancer prognosis

The association between VEGF and prognosis after breast cancer has been extensively studied. Gasparini reviewed such studies in 2000, and concluded that the vast majority of them (8 of the 9 published at that time) supported the conclusion that increased VEGF expression conferred a poorer breast cancer prognosis.[257] The studies included in this review utilized populations of both node-negative and node-positive breast cancer patients.[257] The results of more recent studies, however, appear to be far less consistent. Although some studies have reported decreased disease-free survival and overall survival with increased tumor[258-261] or serum[262] VEGF levels, numerous studies have reported no association between tumor VEGF levels measured by immunohistochemistry or enzyme-linked immunosorbant assay (ELISA) and disease-free survival[263-268] or overall survival.[264, 265, 268, 269] Interestingly, Nishimura et al. reported no association with plasma VEGF levels and overall survival in their full study population, but a significant inverse association was observed when analyses were restricted to postmenopausal women.[269] One study reported that VEGF levels measured by immunohistochemistry were negatively associated with overall survival in univariate analyses,

but this relationship was no longer significant in multivariate analyses.[270] Meo et al. reported that VEGF was not independently associated with risk of recurrence, though in a bivariate model they did observe a significant interaction between VEGF and plasminogen activator inhibitor-1 (PAI-1) such that risk of recurrence was increased when both VEGF and PAI-1 levels were increased.[271] The null findings of at least one study may be due to the few outcome events occurring in the study population.[265]

The inconsistency in the findings may be due to the inclusion of heterogeneous types of breast cancer patients. In fact, one study reported varying effects of tumor VEGF on survival depending upon the type of treatment the patient had received. Foekens et al[272] studied a population of 845 women diagnosed with invasive breast cancer who had experienced a recurrence of their disease; 618 were treated for the recurrence with tamoxifen, while 227 were treated with chemotherapy. VEGF levels in the primary tumor were measured by ELISA and classified as low (<0.22 ng/mg protein), intermediate (0.22 – <1.73 ng/mg protein), or high (≥1.73 ng/mg protein). Among those treated with tamoxifen, median time to progression of disease was 12.2 months for the high VEGF group compared to 18.4 months for the low VEGF group (OR for response to treatment 0.45, 95% CI 0.26-0.78 for high vs. low VEGF). Among those treated with chemotherapy, median time to progression of disease was 6.6 months for the high VEGF group compared to 7.6 months for the low VEGF group (OR for response to treatment 0.31, 95% CI 0.14-0.68 for high vs. low VEGF).[272] These differences in survival may reflect differences in the underlying pathology of the disease which dictate treatment course rather than differences due to treatment, however. In exploratory analyses VEGF appeared to only confer a poor prognosis among the estrogen receptor negative (ER−) women treated with chemotherapy and the estrogen receptor positive (ER+) women treated with tamoxifen.[272]

Other studies have also reported different findings when stratifying by ER or nodal status. Linderholm et al.[273] reported that while the overall associations with disease-free and overall survival were either null or of borderline significance, VEGF was significantly predictive of both disease-free and overall survival among women diagnosed with ER+ disease. On the other hand, Coradini et al.[274] reported that addition of tumor VEGF level provided further information on prognosis to the Nottingham Prognostic Index only among women with ER− disease. Some studies have reported significant associations with both types of survival in populations of women with node-negative disease,[275-277] while others have not.[278] Studies in populations of women with node-positive disease generally report no association between VEGF and survival.[268, 279-282] A study of node-positive and ER+ breast cancer patients did report an interaction with ER level and VEGF such that increasing ER levels lessened or eliminated the negative impact of increased VEGF levels on disease-free survival.[283] Polymorphisms in the VEGF gene may also be related to breast cancer survival; a recent study reported that individuals with a GG genotype at the common polymorphic site +405G/C had lower survival after breast cancer compared to women with the CC genotype (HR 1.6, 95% CI 1.0-2.5).[284] A separate study reported that the -7C/T polymorphism was associated with overall survival.[285] Overall, it currently appears that the effect of VEGF on breast cancer prognosis is controversial and further, more methodologically rigorous, studies are needed.

2.4.4.2 VEGF and breast cancer tumor characteristics

VEGF expression has been reported to be increased in tumor tissue as compared to adjacent, non-cancerous breast tissue[286, 287] or samples from healthy controls,[288] and expression of VEGF in tumor tissue is correlated with angiogenesis.[289] Interestingly, one study reported a significant difference between cancerous and non-neoplastic tissue in postmenopausal women only.[290]

VEGF is believed to be important early in the development of cancer. Indeed, a study of cases of ductal carcinoma in situ reported that approximately 50% of cases had strong expression of VEGF mRNA.[291] In addition to being expressed by the tumor cells themselves, there is evidence that tumor-associated macrophages also produce VEGF and contribute to angiogenesis.[292]

Numerous studies have investigated the relationships between VEGF in breast tumors and clinicopathologic features of the disease, yet these studies provide conflicting results. For example, many studies report no association between tumor VEGF and tumor size,[259, 264-266, 272, 278-280, 287, 292-302] yet others report that tumor VEGF is positively associated with tumor size.[258, 268, 270, 271, 276, 277, 281, 303-305] In a study of 257 tumor samples, Coradini et al. reported a correlation of 0.28 ($p \leq 0.01$) between VEGF and tumor size.[303] Likewise, Linderholm et al. reported VEGF levels greater than the population median (2.40 pg/μg DNA) in 58.2% of larger size tumors and in 45.9% of smaller size tumors ($p=0.008$).[276] Of note is that while half of the studies reporting no association between VEGF and size used immunohistochemistry and a subjective classification for VEGF expression, all but two[270, 281] of the studies reporting positive associations with size measured VEGF using ELISA on tumor cytosols. There are substantial differences in the sensitivity and reliability of these two methods, which thus may explain some of the conflicting results that have been reported. Further, the studies reporting a significant, positive association tended to have larger sample sizes than those reporting no association.

Similarly, some studies report positive associations between VEGF and grade of disease,[258, 268, 270, 276, 281, 301, 304, 305] while most report no association with grade.[259, 264, 272, 277, 279, 280, 292, 294-296, 299, 306, 307] Additionally, one study reported a negative association between VEGF mRNA expression and grade of disease.[302] Similar to the case with tumor size, the studies reporting positive associations tended to have larger sample sizes than those reporting no

associations. Thus these differences in methodology may explain some of the inconsistency in the findings.

The majority of studies comparing VEGF expression or levels by nodal positivity status have reported no association.[197, 259, 264-266, 269, 272, 279, 280, 287, 292-300, 302, 304, 307] Coradini et al., however, reported a significant correlation between number of metastatic lymph nodes and VEGF measured by ELISA ($r=0.21$, $p\leq0.01$),[303] and Zaman et al. reported a trend toward increased plasma VEGF among women with node-positive disease.[305] Mohammed et al. reported that tumors with positive lymph nodes were also more likely to show high levels of VEGF expression by immunohistochemistry (76% node-positive vs. 29% node-negative, $p<0.001$),[270] and a study by Valkovic et al. reported similar findings.[301] Konecny et al. observed a positive association between VEGF and presence of positive lymph nodes when using an assay that recognizes isoforms $VEGF_{121-206}$ ($p=0.015$) but not when using an assay that recognized only isoforms $VEGF_{165-206}$ ($p=0.444$).[268] Few studies have tested for associations between stage of disease and VEGF, but those that have report no significant association.[279, 281, 293, 308]

Associations between hormonal receptor status and VEGF have been extensively studied. Many studies report an inverse association between VEGF and ER status, such that VEGF levels are higher among women with ER-negative disease or are negatively correlated with ER expression.[259, 265, 268, 270, 272, 279, 294, 300, 304] For example, one of the larger studies (N=845) reported that median tumor VEGF levels were higher in ER-negative disease than in ER-positive disease (0.45 versus 0.18 ng/mg protein, $p<0.0001$).[272] Conversely, one study reported increased VEGF expression in ER-positive disease[264] and another reported a positive correlation between VEGF and ER levels.[295] Further, multiple studies report no association between ER status and tumor VEGF[266, 271, 277, 278, 280, 293, 296-298, 302, 306] or serum VEGF.[220, 299] Those studies reviewed

here which reported no association generally employed a small number of participants (range N=25 – N=574, with 11 of 13 having less than 300) while studies reporting inverse associations tended to have larger sample sizes (range N=177 – N=887, with 5 of 9 having more than 300). Thus it is possible that the lack of significant associations between ER status and VEGF may have been due to lack of statistical power in some studies.

Results are similarly mixed for testing associations between PR status and VEGF. Many studies report no association with tumor[264-266, 271, 278, 279, 293, 294, 297, 298, 302, 306] or serum VEGF,[299] while many others report that levels of VEGF are negatively correlated with PR levels and/or that tumor VEGF levels are higher among women with PR-negative disease.[259, 268, 270, 272, 277, 304] Similar to the case of ER status and VEGF, the studies reporting significant inverse associations between PR status and VEGF tended to have larger study populations than did those reporting no associations; thus lack of statistical power may explain some of the conflicting results relating VEGF to PR status as well.

Fewer studies have examined whether VEGF is associated with HER-2/neu expression. While one study reported a positive association between VEGF and HER-2/neu expression,[268] most studies report no association.[264, 279, 294, 306, 308] One study did report a positive association among postmenopausal women only, however.[294] On the other hand, a separate study reported a negative correlation between tumor VEGF expression and HER-2/neu expression.[309] It is therefore unclear if an association exists between VEGF and HER-2/neu at the present time.

Common to most of the studies exploring relationships between VEGF and clincopathological features of breast cancer is the use of a small sample size, as has been previously mentioned. Further, most studies fail to do an adequate job of providing inclusion and exclusion criteria for their studies or of describing the demographic characteristics of their

population. Thus from an epidemiologic perspective more careful design and reporting of such studies is needed before one can make valid conclusions about VEGF and tumor characteristics.

2.4.4.3 VEGF in breast cancer cases and healthy controls

It is also believed that serum and plasma levels of VEGF are higher among breast cancer patients as compared to healthy controls (Table 2.3).[220, 305, 310-321] For example, Heer et al. reported that in a sample of 196 incident breast cancer cases and 88 healthy controls, the median serum VEGF level was higher in cases (305.9 pg/mL; interquartile range 156.7-451.6 pg/mL) than in controls (167.5 pg/mL; interquartile range 101.5-245.3 pg/mL).[317] These studies, however, are all limited by small sample sizes, incomplete description of the study populations, and/or failure to control for potential confounders in the analyses. Thus, the true associations between VEGF levels and breast cancer remain unclear.

At least four studies have investigated the relationship between polymorphisms in the VEGF gene and breast cancer risk.[285, 322-324] Jin et al. studied seven common polymorphisms in the VEGF gene (-2578C/A, -2549del/ins, -2489C/T, -2447G/del, -1154G/A, -634G/C, and +936C/T) and found no association with any of the polymorphisms and breast cancer risk.[324] Likewise, Balasubramanian et al. found no association between the -460C/T, +405G/C, -7C/T, or +936C/T polymorphisms and breast cancer risk, though there was an indication that the -460T/+405C/-7C/963C haplotype was associated with a decreased risk of breast cancer.[285] In a nested case-control study Jacobs et al. reported that homozygous presence of the -2578C or -1154G alleles, both of which are believed to result in increased VEGF expression, were related to increased odds of invasive breast cancer (OR 1.46, 95% CI 1.00-2.14 and OR 1.64, 95% CI 1.02-2.64, respectively), while another allele believed to increase expression of VEGF, +936C, was not associated with invasive breast cancer but reduced the risk of in situ breast cancer

Table 2.3 Summary of studies examining differences in serum or plasma VEGF levels between breast cancer cases and controls

Author	Study Design	Study Population	Results	Comments
Yamamoto[320] (1996)	Case-control	N=184 healthy controls; 132 male, 52 females); age 21-59 N=286 cancer patients (N=175 breast; 137 primary, 38 recurrent); age not stated Menopausal status not stated	Mean serum VEGF in controls 77.0pg/mL; 19.7% of breast cancer cases have serum VEGF level > 180pg/mL (mean + 2SD of control population)	Unclear if cases were incident or prevalent and if blood samples were taken prior to onset of treatment; no formal significance testing comparing VEGF levels in breast cancer cases and controls reported; analyses not adjusted for confounders
Dirix[313] (1997)	Prospective Cohort	N=146 cancer patients (N=17 treated for metastatic breast cancer, N=22 untreated non-metastatic breast cancer) Age and menopausal status distributions not stated	Serum VEGF levels >95th percentile of controls (500pg/mL) in 38% of untreated non-metastatic breast cancer cases; no statistics provided for metastatic breast cancer cases	Control values taken from those reported by manufacturer of assay kit; no formal significance testing comparing VEGF levels in breast cancer cases and controls reported; small number of breast cancer cases; analyses not adjusted for confounders; also followed for prognosis
Donovan[220] (1997)	Case-control	N=15 healthy controls; age-matched to cases N=26 breast cancer patients; age not stated	Mean serum VEGF higher among cases (407.67±272.07pg/mL) than among controls (230.0±127.18pg/mL), p=0.03	Small sample sizes; no selection criteria provided for cases or controls; analyses not adjusted for confounders
Salven[318] (1997)	Case-control	N=113 controls with diabetes (N=7), rheumatoid arthritis (N=5), or healthy (N=81); 19 male, 94 female; age 20-82 N=97 cancer patients (N=33 breast cancer) (N=45 male, N=58 female); age 23-85 Menopausal status not stated	Serum VEGF levels higher among cancer cases (median 197pg/mL, range 8-1711pg/mL) vs. controls (median 17pg/mL, range 1-177pg/mL) (p<0.0001); median serum VEGF level higher among disseminated breast cancer cases (median 205, range 21-1347pg/mL) versus locoregional breast cancer (median 150pg/mL, range 132-244 pg/mL) (no significance testing performed)	Small numbers of breast cancer cases; no specific testing of breast cancer cases versus controls; analyses not adjusted for confounders

Table 2.3 (continued)

Salven[314] (1999)	Case-control	N=73 patients undergoing breast surgery; histology reveals 18 benign tumors, 7 in situ disease, and 48 invasive cancers; median age 59, range 22-95 N=32 patients with metastatic breast cancer; 27 undergoing active treatment at time of blood draw; age distribution not stated Menopausal status not stated	Serum VEGF levels higher among invasive cancer cases (median 104pg/mL, range 11-593pg/mL) and metastatic cases (median 186pg/mL, range 7-1347pg/mL) versus benign tumors (median 57pg/mL, range 18-328pg/mL) (p=0.13 invasive, p=0.0018 metastatic)	Analyses not adjusted for confounders; controls not "healthy" but actually have benign breast conditions
Adams[315] (2000)	Case-control	N=12 women with benign breast disease; median age 47, range 32-63; 6 pre- and 6 postmenopausal N=62 women with localized breast cancer; median age 56, range 29-85; 4 peri-, 14 pre- and 44 postmenopausal N=42 women in remission from breast cancer; median age 54, range 38-85; 3 peri-, 7 pre- and 32 postmenopausal N=22 women with metastatic breast cancer; median age 52, range 32-82; 7 pre- and 15 postmenopausal N=63 healthy women; median age 37, range 20-72; 49 pre- and 14 postmenopausal	Plasma and serum VEGF levels significantly different among groups (p<0.001 plasma, p=0.048 serum); median plasma levels: 40.9pg/mL metastatic, 44.5pg/mL remission, 31.3pg/mL local disease, 28.3pg/mL benign disease, 27.3pg/mL controls; median serum levels: 252.9pg/mL metastatic, 293.7pg/mL remission, 244.2pg/mL local disease, 264.8pg/mL benign disease, 186.0pg/mL controls	Analyses not adjusted for potential confounders; also noted that 80% of tumors expressed VEGF
Heer[317] (2001)	Prospective Cohort	N=200 breast cancer cases; age range not stated N=88 healthy controls; age range 22-79 Menopausal status not stated	Median serum VEGF higher in cases (305.9pg/mL, interquartile range 156.7-451.6pg/mL) compared to controls (167.5pg/mL, interquartile range 101.5-245.3pg/mL) (p<0.0005)	Analyses not adjusted for potential confounders; also followed for prognosis

Table 2.3 (continued)

Benoy[310] (2002)	Case-control	N=104 breast cancer cases N=26 healthy controls Age and menopausal status distributions not stated	Mean serum (347±276pg/mL) and plasma (45±32pg/mL) VEGF higher in cases than in controls (105±74pg/mL and 14±9 pg/mL, respectively), both p<0.0001	Analyses not adjusted for potential confounders
Wu[319] (2002)	Prospective Cohort	N=125 African American (N=57) and Hispanic (N=68) breast cancer cases; mean age 50.7±1.3 African American, 46.6±1.5 Hispanic; 51% African American and 35% Hispanic postmenopausal N=20 African American and Hispanic healthy women; age and menopausal status distributions not stated	Cases (median 39.3 pg/mL, range 20.9-417.2pg/mL) have higher plasma VEGF than controls (median 24.4 pg/mL, range 18.0-77.7pg/mL) (p<0.0002)	Analyses not adjusted for potential confounders; menopausal status determined by < or ≥ age 50; also followed for prognosis
Coskun[311] (2003)	Case-control	N=38 metastatic breast cancer patients; mean age 50.0±13 N=23 breast cancer patients in remission; mean age 48.3±9.6 N=16 healthy controls; mean age 47.4±9.4 Menopausal status unknown	Serum VEGF higher in patients with metastatic disease (252.7±147.5pg/mL) compared to patients in remission (137.2±64.7pg/mL, p<0.001) and controls (107.3±50.0pg/mL, p<0.001); remission and control groups not statistically significantly different (p>0.05)	Analyses not adjusted for potential confounders

Table 2.3 (continued)

Nishimura[269] (2003)	Case-control	N=15 patients with benign breast disease; mean age 46.5, range 32-79 N=187 patients with primary breast cancer; mean age 54.3, range 29-95 N=32 patients with non-recurrent breast cancer; mean age 54.0, range 35-86 N=56 patients with recurrent breast cancer; mean age 57.9, range 32-85	Significant differences in mean plasma VEGF levels observed among all groups: benign breast disease 16.0±2.1 pg/mL, primary breast cancer 25.7±26.6 pg/mL, non-recurrent breast cancer 18.9±12.7 pg/mL, recurrent breast cancer 44.7±53.8 pg/mL	Analyses not adjusted for potential confounders; no inclusion or exclusion criteria stated; mean plasma VEGF values stated in text do not match those listed in figure displaying these results
Granato[316] (2004)	Case-control	N=69 breast cancer patients; median age 67, range 51-92; menopausal status not stated N=54 healthy controls; median age 60, range 51-77; menopausal status not stated	Serum VEGF higher among cases (median 192.7pg/mL, range 22.7-953.5pg/mL) than controls (median 145.7pg/mL, range 0.0-707.6pg/mL) (p=0.055)	Analyses not adjusted for potential confounders; not clear if cases were incident or prevalent; unclear why some subjects not included in analyses
Hoar[312] (2004)	Case-control	N=51 breast cancer patients; mean age 57.3±11.1 N=51 age and sex matched healthy controls; mean age 54.0±5.4 Menopausal status not stated	Plasma VEGF higher among cases (median 120pg/mL, interquartile range 15-8,000pg/mL) than controls (median 34pg/mL, interquartile range 18-90pg/mL), p=0.03	Controls said to have been age matched to cases, but difference in age is borderline significant (p=0.06), analyses not adjusted for potential confounders

Table 2.3 (continued)

Teh[321] (2004)	Prospective Cohort	N=17 breast cancer patients; mean age 61, range 48-72 N=7 patients with fibroadenomas; mean age 30.2, range 25-35 N=7 healthy controls; mean age 55, range 45-72 Menopausal status not stated	No difference in plasma VEGF levels between cases (pre-surgery 96.58±26.6pg/mL) and controls (88±12.3pg/mL) (p=0.07)	Small numbers therefore likely underpowered; analyses not adjusted for potential confounders; controls age-matched to breast cancer cases but age distributions are not equal
Zaman[305] (2006)	Case-control	Unknown number of healthy controls; age ≥18 N=17 breast cancer patients; age ≥18 Menopausal status not stated	Plasma VEGF significantly higher in cases (52.9±29.9pg/mL) than controls (37.6±25.5pg/mL) (p=0.04)	Small numbers of cases and unknown number of controls; analyses not adjusted for confounders; little information provided regarding subject recruitment

(OR 0.49, 95% CI 0.37-0.93).[322] Krippl et al. reported that individuals with a +936T allele which is believed to result in lower VEGF levels, either homo- or heterozygous, had a reduced risk of breast cancer (OR 0.51, 95% CI 0.38-0.70).[323] Despite these findings and those showing the importance of VEGF to cancer growth and progression, only one study has been conducted to determine if biological levels of VEGF are related to risk of breast cancer. This study reported no association between the -460C/T, +405G/C, -7C/T, or +936C/T polymorphisms and plasma, serum, or tumor VEGF levels.[285] Future studies are needed to verify these results and to relate blood and tumor VEGF levels to other common polymorphisms in the VEGF gene.

2.4.4.4 VEGF and personal characteristics

The variation of VEGF by personal characteristics of breast cancer patients or healthy controls has not been extensively studied. In fact, a review of the literature identified comparisons based only on age, menopausal status, or BRCA1 genotype. Many studies have examined associations between VEGF and age, with most finding no association.[264, 270, 272, 276, 279, 280, 292, 295-299, 304, 306, 307, 325] Three studies have reported a positive association between age and tumor VEGF levels,[269, 277, 293] while two other studies found inverse associations between age and tumor VEGF.[278, 281] Greb et al. reported a negative correlation between VEGF mRNA expression and age in non-neoplastic breast tissue, yet no association with age in cancer tissue from the same women.[290] Further, one study reported no association between serum VEGF and age among women treated with chemotherapy, yet a positive association among women treated with tamoxifen.[300] The majority of studies report no association between VEGF and menopausal status.[259, 264, 266, 272, 278-280, 302, 326] However, one reported that VEGF levels were increased among premenopausal women[294] while others reported higher tumor or plasma VEGF levels in postmenopausal women.[269, 306] Also, Greb et al. reported that VEGF mRNA expression was

higher in the non-neoplastic breast tissue of premenopausal women, though no difference by menopausal status was observed for cancerous tissue.[290] One study reported that serum VEGF was lower in breast cancer patients with a BRCA1 mutation compared to those without such a mutation.[326] Studies do not appear to have investigated whether VEGF levels vary by other relevant factors such as race/ethnicity, family history of breast cancer, or use of HT or oral contraceptives; thus future research of the variation of VEGF by personal characteristics is warranted.

2.4.4.5 Hormonal regulation of VEGF

In healthy adults angiogenesis normally does not occur except in the female reproductive system. Many *in vitro* studies have reported that estradiol (E2) also induces VEGF (see section 2.5.2). Further, Dabrosin reported high correlations between plasma E2 ($r=0.814$, $p<0.0001$) and breast tissue E2 ($r=0.67$, $p=0.004$) with VEGF in breast tissue of 16 pre- and postmenopausal healthy women.[327] Though the evidence is still preliminary, a relationship between estrogen and VEGF could suggest another mechanism by which estrogen acts to promote carcinogenesis in breast tissue and also adds support to the hypothesis that increased levels of VEGF are positively associated with breast cancer risk. Other evidence that sex hormones can regulate VEGF include the observation that administration of follicle stimulating hormone (FSH) increases follicular VEGF levels among women undergoing in vitro fertilization,[328] as well as evidence that administration of testosterone induces VEGF mRNA expression in murine breast cancer cells.[245] Thus sex hormones may play a significant role in regulating angiogenesis.

2.4.4.6 VEGF and the menstrual cycle

There is some controversy regarding whether levels of angiogenic factors are affected by the menstrual cycle. Some studies have not observed a cyclic variation in VEGF during normal menstrual cycles.[315, 329-332] One study has reported that neither serum nor plasma VEGF levels varied during the menstrual cycle in a cohort of 20 healthy premenopausal women.[329] A limitation of this study was that most of the plasma measurements were below the lower limit of detection of the assay used, thus preventing variability in the plasma levels from truly being assessed. Other studies have reported cyclic variation in VEGF levels during the menstrual cycle. For example, two studies reported that serum VEGF levels were lower in the luteal phase as compared to the follicular phase.[333, 334] Both of these studies, however, included a small number of participants (14 in Heer et al. and 6 in Kusumanto et al.) and also failed to properly analyze the results as repeated measures, instead analyzing the results as independent observations. Thus the results of these studies must be viewed with caution. Other investigators have reported that VEGF is higher in the luteal phase.[335, 336] Dabrosin et al. reported that VEGF levels in the extra-cellular fluid of breast tissue was increased in the luteal phase but that plasma levels showed no variability with phase of the menstrual cycle.[336, 337] Further research is needed to understand the effects of estrogen and the menstrual cycle on VEGF levels.

2.5 SUMMARY

As evidenced by the substantial morbidity and mortality it causes, breast cancer is truly a significant public health problem in the United States. Though great strides have been made in improving methods of early detection and treatment, both of which have led to reduced breast

cancer mortality, effective prevention of breast cancer still remains an elusive goal. Much is currently known about risk factors for breast cancer, including reproductive history, genetic mutations, and hormone use. It is also recognized that overweight and obesity are related to risk of postmenopausal breast cancer. Interestingly, body weight is inversely related to percent breast density, which is considered to be a potential surrogate endpoint for breast cancer. Studies have not been designed to look at relationships between anthropometry and breast density longitudinally. Such investigations may help to understand the reasons for this contradiction and how to appropriately control for it in studies utilizing breast density as a surrogate endpoint. It is also increasingly apparent that angiogenesis, and specifically vascular endothelial growth factor, plays an important role in the etiology of breast cancer. Though *in vitro* and animal studies relating VEGF to breast cancer have been extensive, the focus of such research in humans has been primarily on prognosis after breast cancer diagnosis. Those studies that have addressed the role of VEGF in breast cancer etiology using human subjects have been severely limited by selection bias, confounding, and lack of statistical power. Further, very little is known about the correlates of VEGF among healthy women and women with breast cancer. Thus we do not know if factors such as age, menopausal status, and hormone use affect serum levels of VEGF.

We present a longitudinal evaluation of anthropometric measures (BMI, weight) and breast density (percent and absolute), a method for estimating outcome data collected off-schedule from planned study visits, and a case-control study evaluating differences in serum VEGF levels between women with and without breast cancer. This research tests novel hypotheses intended to enhance existing knowledge of breast cancer etiology and thus contribute to public health.

3.0 ARTICLE 1: LONGITUDINAL INFLUENCE OF ANTHROPOMETRY ON MAMMOGRAPHIC BREAST DENSITY: THE STUDY OF WOMEN'S HEALTH ACROSS THE NATION (SWAN)

To be submitted for publication

Katherine W. Reeves, MPH,[1] Roslyn A. Stone, PhD,[2] Francesmary Modugno, PhD, MPH,[1] Roberta B. Ness, MD, MPH,[1] Victor G. Vogel, MD, MHS,[3] Joel L. Weissfeld, MD, MPH,[1,4] Laurel Habel, PhD,[5] Barbara Sternfeld, PhD,[5] Jane A. Cauley, DrPH, MPH[1]

[1]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[2]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[3]Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

[4]University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania

[5]Division of Research, Kaiser Permanente, Oakland, California

This ancillary study was supported by National Cancer Institute grant R01 CA89552. The Study of Women's Health Across the Nation (SWAN) has grant support from the National Institutes of Health, DHHS, through the National Institute on Aging, the National Institute of Nursing Research and the NIH Office of Research on Women's Health (Grants NR004061; AG012505, AG012535, AG012531, AG012539, AG012546, AG012553, AG012554, AG012495).

3.1 ABSTRACT

High percent mammographic breast density is strongly associated with an increased risk of breast cancer. Though body mass index (BMI) is positively associated with risk of postmenopausal breast cancer, BMI is negatively associated with percent breast density in cross-sectional studies. Longitudinal studies evaluating associations between changes in BMI and weight and mammographic breast density have not been conducted. We studied the longitudinal relationships between anthropometry and breast density in a prospective cohort of 834 pre- and perimenopausal women enrolled in an ancillary study to the Study of Women's Health Across the Nation (SWAN). Routine screening mammograms were collected and read for breast density. Random intercept regression models were used to evaluate whether longitudinal changes in BMI and weight were associated with changes in dense breast area and percent density. The study population was racially diverse (7.4% African American, 48.8% Caucasian, 21.8% Chinese, and 21.9% Japanese). Mean follow-up was 4.8 years, and mean annual weight change was +0.22 kg/year. In fully adjusted models, changes in BMI and weight were not associated with changes in dense breast area (β=-0.0105, p=0.34 and β=-0.0055, p=0.20, respectively), but were strongly negatively associated with changes in percent density (β=-1.18, p<0.001 and β=-0.44, p<0.001, respectively). This study provides evidence that longitudinal changes in BMI and weight are not associated with the dense area, yet are negatively associated with percent density. Thus, effects of changes in anthropometry on percent breast density may reflect effects on non-dense tissue, rather than on the dense tissue where cancers arise.

3.2 INTRODUCTION

In 2007 alone an estimated 178,480 U.S. women will be diagnosed with breast cancer.[1] Though substantial efforts have been devoted to studying breast cancer etiology and prevention, the pace of this research is often slow due to the decades needed for breast cancer to develop. Therefore, cancer epidemiology studies are incorporating surrogate endpoints, which allow for studies to be conducted with fewer subjects and over a shorter time period while aiding in the understanding of the mechanisms of cancer development.[2]

The breast is composed of different types of tissues, and the composition varies from woman to woman; fat appears dark on a mammogram because it is radiologically lucent, while epithelial and connective tissues appear bright because they are radiologically dense. The areas of density can be measured and summed to determine the dense breast area. Mammographic breast density is typically expressed as a percentage (dense area/total breast area*100%). High mammographic breast density increases the risk of breast cancer more than fourfold compared to women with low breast density.[3]

Body mass index (BMI) bears an idiosyncratic relationship to breast density and breast cancer. It is related to an increased risk of postmenopausal breast cancer and a decreased risk of or no association with premenopausal breast cancer.[4] However, cross-sectional studies of mammographic breast density in both pre- and postmenopausal women consistently report that increased weight or BMI is associated with lower percent breast density.[5-12] For example, one study reported percent density was 5.2% lower among premenopausal women and 4.7% lower among postmenopausal women in the 3rd quartile versus 1st quartile of BMI.[7]

The amount of fat present in the breast is strongly related to BMI.[13] Recent studies have evaluated the association between anthropometric measures and the dense and non-dense areas

of the breast.[8, 10, 14, 15] Several studies have shown that weight and BMI are positively associated with the size of the non-dense area;[8, 10, 14, 15] yet for the dense breast area, two studies reported negative correlations,[8, 14] one reported a non-significant association,[10] and another reported that the direction of the association varied by race/ethnicity.[15] Only a few studies have used longitudinal data to evaluate associations between anthropometric measures and breast density.[12, 16, 17] These longitudinal studies have been limited by small numbers, self-reported anthropometric measurements, or use of a non-quantitative measure of breast density.

In the current analysis we performed a prospective cohort study of 834 women enrolled in an ancillary study to the Study of Women's Health Across the Nation (SWAN). Our primary hypothesis was that time-related changes in weight and BMI would be positively associated with time-related changes in dense area and negatively associated with percent density.

3.3 METHODS

3.3.1 Study population

SWAN began in 1994 to study the health of women as they transition from premenopause to postmenopause.[18] Briefly, women eligible for SWAN were age 42-52, had ≥1 menstrual period in the previous three months, had a uterus and ≥1 intact ovary, and were not currently taking hormone therapy (HT) or oral contraceptives (OCs). The SWAN Mammographic Density Substudy is an ancillary study to SWAN with the goal of examining factors related to mammographic breast density and how mammographic breast density changes as women progress through the menopausal transition. Separate written informed consent and institutional

review board approvals were obtained for this ancillary study. Participants from three SWAN sites (Los Angeles, CA, Oakland, CA, and Pittsburgh, PA) representing four races/ethnicities (African American, Caucasian, Chinese, and Japanese) were enrolled at their 5th or 6th annual visit. Caucasian participants were enrolled from all three sites, while all African Americans were from the Pittsburgh site, all Chinese from the Oakland site, and all Japanese from the Los Angeles site. Eligible mammograms included routine screening mammograms that were taken two years prior to the baseline SWAN visit through two years after the 6th annual SWAN visit. Mammograms of breasts that had undergone a biopsy or more extensive surgery were ineligible for the SWAN Mammographic Density Study. Of those eligible for the SWAN Mammographic Density Substudy, 86.1% (N=1,055) consented and at least one mammogram was obtained from 95.3% of these women (N=1,007).

Participants were excluded from this analysis if they reported a history of breast cancer at SWAN enrollment (N=6) or had only one available mammogram (N=139). Women who were diagnosed with breast cancer during their SWAN follow-up (N=21) were censored at the visit they reported this diagnosis. Ten of these women had no mammograms prior to their breast cancer diagnosis and were excluded from this analysis. After these exclusions 852 participants and 3,784 mammograms remained in the analysis. Three participants reported being either pregnant or breastfeeding at the time of a SWAN visit, and these specific visits were not included in the analysis.

3.3.2 Mammographic density readings

The mammographic density assessments were performed by a single reader using craniocaudal views of the right breast. Films of the left breast were used if a participant had a previous

surgery to the right breast or if the right breast films were of poor quality. Prior to performing the density assessment, the reader rated the quality of each film as excellent, fair, good, or poor.

Quantitative measures of density were made by tracing the total area of the breast and outlining the areas of density (excluding biopsy scars, Cooper's ligaments, and breast masses) onto clear acetate placed over the mammogram. A compensating polar planimeter (LASICO, Los Angeles, CA) was used to measure the total breast area and the dense breast area in cm^2. A blinded random sample of films (N=449) was used to assess the reproducibility of the density assessments. This re-review showed good association between the initial and repeat readings of percent density (within-person Spearman correlation coefficient=0.96).

3.3.3 Anthropometric measures

Height was measured without shoes using either a metric folding wooden ruler or measuring tape (home and some clinic visits), or a fixed stadiometer (clinic visits). Weight was measured without shoes, and in light indoor clothing, using a portable digital scale (home and some clinic visits) or either a digital or balance beam scale in the clinic. Portable scales were calibrated weekly, and stationary clinic devices were calibrated monthly. BMI was calculated as the weight in kilograms divided by the square of the height in meters.

3.3.4 Additional variables

Additional data were collected at the clinic visits by either interviewer- or self-administered questionnaires. These data included demographic information, and personal and family history of breast cancer. Cancer history was updated at each SWAN visit. Reproductive variables

ascertained at baseline included age at menarche, age at first birth, number of births, and history of breastfeeding. Number of births and breastfeeding history were updated at each SWAN visit. History and number of previous breast biopsies were reported at the time of enrollment into the SWAN Mammographic Density Substudy. These data were used to calculate the Gail score for each participant.[19, 20] Gail scores of 1.66 or higher indicate a high 5-year risk of breast cancer. History of atypical hyperplasia was not collected in this study and was listed as "unknown" for each woman in the calculation of the Gail scores.

At baseline women reported their history of HT and OC use, as current users were initially excluded; current HT and OC use was updated at each SWAN visit. Menopausal status was ascertained using an interviewer-administered questionnaire at each visit. Women were asked about the frequency and regularity of their menstrual bleeding. Women reporting no decrease in their menstrual regularity over the past year were classified as premenopausal. Early perimenopause was classified as decreased menstrual regularity within the previous three months, and late perimenopausal was defined as no menstrual bleeding in the 3-11 months prior to the interview. Women reporting no menstrual bleeding for at least 12 months prior to the interview were classified as postmenopausal. Women reporting a bilateral oophorectomy and/or hysterectomy were defined as having a surgical menopause. Those women reporting use of HT with some bleeding within the past 12 months were classified as unknown menopausal status. Women classified as postmenopausal remained classified as such thereafter, regardless of their HT use.

3.3.5 Statistical analysis

Mammogram data and SWAN visit data for each participant were ordered chronologically. Because these routine screening mammograms were not performed as part of the SWAN study, the mammogram dates did not match the dates of SWAN visits. We developed an algorithm to match retrieved screening mammograms to the nearest SWAN study visit, regardless of whether the mammogram preceded or followed the study visit. To reduce potential error associated with variable time between mammograms and visits, only mammograms taken within 90 days of a SWAN visit were used as matches. Mammographic breast density variables (total breast area, dense breast area, and quality of film) at other study visits were estimated with linear interpolation using the "ipolate" command in Stata version 10.0 (Stata Corporation, College Station, TX). Separately for each participant, the interpolated estimates were calculated by constructing straight line segments between measurements from consecutive mammograms. The value for an unmatched SWAN visit was calculated from the equation of the line segment, assuming that the change in density variables was linear between the two measurement dates that defined the line segment. We did not use extrapolation to estimate breast density; therefore, visits without a mammogram both before and after the visit time were not included in the analysis.

We added a noise term to each interpolated value of total and dense breast area to account for the error introduced by estimating the breast density measurements. These noise terms were randomly generated from a normal distribution with a mean of 0 and standard deviation (SD) corresponding to the person-specific SD of the observed measurements for each participant. Multiple imputation was used to create ten analytic datasets. Percent breast density was calculated by dividing the final (i.e. observed data if mammogram within 90 days, interpolated

data otherwise) dense area measurement by the final total breast area and expressing this value as a percentage. Due to the addition of the random noise terms, some interpolated values (<1%) were considered to be implausible (i.e. total area<0, dense area<0, or dense area>total area); such values were discarded and additional random noise terms generated until acceptable imputed values were obtained. After implementation of the matching and interpolation algorithms, 18 participants were left with less than two mammographic density measurements due to the timing of their mammograms. These participants were excluded from further analyses; the remaining 834 participants had 3,746 eligible mammograms. Women enrolled in the SWAN Mammographic Density Substudy but excluded from this analysis were of similar age, educational level, and menopausal status as those included. Participants included were of slightly lower BMI (25.4 versus 26.4 kg/m^2, p=0.06), less likely to be African American (7.4% versus 16.8%, p<0.001), and less likely to be from the Pittsburgh clinical site (23.7% versus 43.9%, p<0.001; data not shown).

Summary statistics were calculated for demographic, anthropometric, reproductive history, and mammographic breast density variables. The averages of the mammographic density variables across the ten multiply imputed datasets were calculated and used for the descriptive statistics. The length of time between the participants' study visits and their nearest mammograms was calculated. A square root transformation was applied to dense breast area after imputation to improve normality. Analysis of variance (ANOVA) and Kruskal-Wallis tests were used to test for baseline differences in normally and non-normally distributed continuous variables, respectively, and chi-square tests were used to test for differences in categorical variables. Bivariate cross-sectional associations between the anthropometry and breast density variables were assessed using ANOVA.

Random intercept models were fit with the participants' age in days as the time scale using the Stata "xtreg" command. Separate regressions were performed for the two primary mammographic breast density outcomes (dense breast area and percent density) for each of the two primary independent variables of interest (BMI and weight). A total of four regressions were performed. Variables included as possible covariates were: age (continuous), combined race and site (Caucasian/Pittsburgh, African American/Pittsburgh, Caucasian/Oakland, Chinese/Oakland, Caucasian/Los Angeles, Japanese/Los Angeles), family income (<$35,000, $35,000-$49,999, $50,000-$74,999, $75,000-$99,999, ≥$100,000), education (≤high school, >high school, college, post-college), age at menarche (<12, 12, 13, ≥14), age at first birth (<20, 20-24, 25-29, 30-34, ≥35), breastfeeding history (nulliparous, parous/never, 1-4 months, 5-11 months, 12-22 months, ≥22 months), number of births (0, 1, 2, ≥3), menopausal status (pre-/early perimenopausal, late perimenopausal/postmenopausal/hysterectomy with bilateral oophorectomy, unknown due to HT use), ever use of OCs prior to baseline (no, yes), ever use of HT prior to baseline (no, yes), current HT/OC use (no, yes), ever HT/OC use at each visit (no, yes), number of 1st degree relatives with breast cancer (0, ≥1), number of 2nd degree relatives with breast cancer (0, 1, ≥2), history of breast biopsy (no, yes), number of breast biopsies (0, 1, ≥2), and Gail score (continuous). Continuous variables were centered on the population mean. Time-varying variables were included as appropriate (e.g. menopausal status, hormone use).

Model building using the first imputed dataset followed a backward selection of covariates, retaining covariates that were significant at the 0.10 level in the final model. The "mijoin" and "micombine" commands in Stata were used to estimate the regression models and calculate multiply imputed estimates of the regression coefficients and their variances following the method of Rubin.[21, 22] We report the results of three different models: Model 1, BMI or

weight only; Model 2, BMI or weight and age; and Model 3, BMI or weight and additional covariates.

Analyses were repeated separately by race/ethnicity, baseline BMI, and menopausal status. All tests performed were two-sided with a $p \leq 0.05$ considered statistically significant. All analyses were conducted using Stata version 10.0.

3.4 RESULTS

3.4.1 Characteristics of study population

The 834 participants comprising the study population are described in Table 3.1. The average age of the participants at SWAN enrollment was 46.5 years (SD 2.7). This population was racially diverse, with 62 (7.4%) African American, 407 (48.8%) Caucasian, 182 (21.8%) Chinese, and 183 (21.9%) Japanese participants. The majority of participants (57.5%) were categorized as normal weight at SWAN enrollment. On average, participants had a low risk of being diagnosed with breast cancer within the next 5 years, with a mean Gail score of 1.06 (SD 0.5). By design all women were either premenopausal (58.3%) or early perimenopausal (41.7%) at enrollment. The vast majority reported a previous use of OCs (75.4%) but no previous use of other exogenous hormones (86.4%) at enrollment. Use of OCs or HT since the previous visit increased during follow-up, reaching a maximum of 27.4% at visit 6. At visit 7, 26.2% of participants were premenopausal/early perimenopausal, 68.0% were late perimenopausal/postmenopausal, and 5.9% were of unknown menopausal status due to use of hormone therapy (data not shown).

The average follow-up time between the first and last SWAN visits included in this analysis was 4.8 years (SD 1.8). Overall, participants tended to gain weight over follow-up (Table 3.2), with a mean annual weight increase of 0.22 kg (SD 1.84) and a mean annual BMI increase of 0.09 kg/m^2 (SD 0.70).

3.4.2 Breast density characteristics and cross-sectional associations with anthropometry

Approximately half of all mammograms matched to SWAN visits 1-6 were taken within 90 days of the visit. Of mammograms matched to SWAN visits 0 and 7, 28.4% and 36.9% were taken within 90 days of the visit, respectively (data not shown). The mean dense breast area from participants' first available mammogram was 46.2 cm^2 (SD 26.7), and the mean percent breast density was 42.3% (SD 19.6; Table 3.3). Dense breast area and percent density were positively correlated ($r = 0.48$, $p<0.001$). When participants were cross-classified by quartiles of dense breast area and percent density from their first mammogram, 39.8% were ranked in the same quartile of both dense breast area and percent density.

In cross-sectional analyses of the participants' first mammogram and their height and weight at that time, dense breast area was positively associated with BMI category ($p=0.01$) and weight quartile ($p=0.002$). Percent density was inversely associated with both BMI category ($p<0.001$) and weight quartile ($p<0.001$).

3.4.3 Longitudinal associations between anthropometry and breast density

The average annual change over follow-up was -0.58 cm^2 (SD 3.32) for dense area and -1.01% (SD 2.38) for percent density (Table 3.2). No longitudinal associations between changes in BMI

or weight and changes in the dense breast area were apparent (Table 3.4). Age-adjusted regressions (Model 2) resulted in small, non-significant associations between BMI (β=0.00003, p=0.99) or weight (β=-0.0004, p=0.91) and square-root transformed dense breast area. In models adjusted for age, race/site, menopausal status, family history of breast cancer, number of previous breast biopsies, and hormone use since previous visit (Model 3), small, non-significant, negative associations were observed for BMI (β=-0.0015, p=0.34) and weight (β=-0.0035, p=0.20).

In contrast, changes in BMI and weight were significantly negatively associated with percent breast density (Table 3.4). Percent breast density decreased by 1.30% per 1 unit increase in BMI in age-adjusted analyses (p<0.001) (Model 2), and this association was only slightly attenuated with additional adjustment for age, race/site, education level, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of OC use at baseline, and hormone use since previous visit (β=-1.18, p<0.001) (Model 3). Similar relationships were observed for weight, with a decrease of 0.44% in percent breast density per kilogram increase in weight in a fully adjusted model (p<0.001) (Model 3).

3.4.4 Longitudinal associations across initial BMI categories

To assess whether the longitudinal associations between BMI and weight and breast density variables varied by initial BMI, regression coefficients for Caucasian participants were compared across categories of BMI at SWAN enrollment. The Caucasian subgroup was the only one that included enough participants in the normal, overweight, and obese categories to provide reliable estimates. For dense breast area, the regression coefficients for BMI decreased to negative values as the baseline BMI category increased (normal β=0.023; overweight β=0.018; obese β=-

0.006), although they remained small and non-significant. For percent density the regression coefficients remained negative yet became smaller as baseline BMI category increased: normal β=-1.27; overweight β=-0.56; obese β=-0.29 (data not shown). Similar relationships were observed for regressions with weight as the primary independent variable.

3.4.5 Sensitivity analyses

The stability and consistency of these relationships were investigated through a series of sensitivity analyses grouping on potentially confounding factors. Similar associations, as judged by the magnitude and direction of regression coefficients, to those observed in the complete cohort were observed in analyses restricted to racial subgroups, to women who did not use any exogenous hormones throughout follow-up (N=441), to women who remained premenopausal/early perimenopausal throughout follow-up (N=183), to women with ≥80% of mammograms taken within 90 days of a SWAN visit (N=98), to women with no breast cancer diagnosis during follow-up (N=824), and to observations with mammograms of good or excellent quality (N=819; data not shown).

Further, we tested for interactions between our anthropometry variables and menopausal status. We observed no significant interaction between BMI (p=0.29) or weight (p=0.11) and menopausal status in regressions on dense breast area. We observed borderline significant interactions between BMI (p=0.06) and weight (p=0.09) and menopausal status in regressions on percent density.

3.5 DISCUSSION

This analysis of data from a prospective, multi-ethnic cohort of 834 women revealed no association between longitudinal changes in anthropometric measures and the dense breast area, yet significant negative associations between changes in anthropometry and percent breast density. This study is among the first to prospectively evaluate the effects of anthropometry on both relative and absolute measures of breast density. Over time, a one unit increase in BMI was associated with a decrease of 1.18% in percent breast density and a one kilogram increase in weight was associated with a decrease of 0.44% in percent breast density. Figure 3.1 shows estimated breast density measurements as a function of BMI to illustrate the observed associations between change in BMI and change in breast density. Though our observations regarding percent density were in agreement with our stated hypothesis, the lack of a relationship with dense breast area was counter to what we hypothesized. We did observe highly significant positive cross-sectional associations between BMI and weight and the dense breast area, but these associations did not persist in longitudinal analyses.

Our results are largely in agreement with the many previous cross-sectional studies reporting significant inverse relationships between BMI or weight and percent density.[5-7, 9-11, 14, 15, 23, 24] We did observe a significant positive cross-sectional relationship between BMI and weight and dense area, in agreement with some,[10] but not all, previous studies.[14, 24] In one study, however, the association was made positive after adjustment for the non-dense breast area.[24]

Few studies have used longitudinal data to analyze associations between anthropometry and mammographic breast density. McCormack et al. reported that women with larger increases in BMI between ages 43 and 53 had an increased risk of having high-risk Wolfe patterns.[12] Their study, however, is not directly comparable to ours due to their use of a qualitative breast

density measurement and their use of longitudinal change in BMI yet only a single mammographic density assessment. Boyd et al.[16] related weight change over two years to change in breast density measurements over the same time period. As in our analysis, Boyd et al. reported a significant negative association between weight change and percent density, such that percent density was increased in those who lost weight; however, their study also reported a significant positive association between weight change and the size of the dense breast area, such that dense breast area was decreased in those who lost weight.[16] It is unclear why our findings differed from these latter results. The characteristics of the two populations differed markedly in their age, ethnicities, baseline breast cancer risk, and observed weight change. These differences may explain at least some of the discordant results of these two studies. Indeed, a nested case-control study of Native Hawaiian, Japanese, and Caucasian women recruited from the general population reported results similar to those we observed. Maskarinec et al.[17] reported that overweight and obese women had a more gradual decline in percent density over time as compared to women of normal weight. Likewise, we observed that the regression coefficients for both BMI and weight with the outcome of percent density were more strongly negative among women of normal weight at study enrollment than among those who were overweight or obese at that time. The authors did not report on the outcome of dense breast area, however, precluding a direct comparison to our results.

The observation that anthropometry is related to the dense breast area in cross-sectional studies but not in longitudinal studies appears to be paradoxical at first consideration. Overweight and obese women may have a larger dense breast area than underweight or normal weight women simply because the total breast size is generally larger in women of greater weight; this explains the highly significant cross-sectional associations. Indeed, in a previous

cross-sectional analysis of mammographic data from SWAN, significant positive correlations between BMI and the total breast area and also between the total breast area and the dense breast area were observed.[15] After these cross-sectional differences are accounted for, further increases in weight and BMI appear to result in the accumulation of fat in the breast rather than altering the dense breast tissue. Thus the total breast area increases while the dense breast area remains relatively constant. As total breast area is the denominator when calculating percent breast density, increased total breast area results in a decrease in percent density.

Overall, our results provide evidence that the consistently demonstrated relationship between anthropometry and breast cancer risk may not proceed through a direct effect of anthropometry on the size of the dense breast area. The significant inverse association between increases in BMI and weight and percent breast density most likely reflects the effect of anthropometry on the non-dense area. Indeed, this effect on the non-dense area may explain the effect of anthropometry on breast cancer risk. In adipose tissue, such as that comprising the non-dense area of the breast, androstenedione is converted to estrogen.[25, 26] Higher non-dense breast area may therefore result in increased estrogen exposure to the nearby dense breast tissue due to this peripheral production of estrogen.[10, 17] This increased estrogen exposure of the ducts and lobules where cancers arise may result in increased risk of breast cancer. Therefore, observing a longitudinal decrease in percent density may actually reflect an increase in breast cancer risk if the decreased percent density results from an increase in non-dense tissue rather than a decrease in the dense breast area.

Haars et al.[10] noted that percent breast density may not be valid for etiologic inference because this measure incorporates information about both the dense breast area, believed to represent cells at risk for developing cancer, and BMI, an independent risk factor for breast

cancer. Further, Haars et al. reported that only 37% of their participants were ranked in the same quartile of both percent density and dense breast area; in other words, a group of women with equivalent percent density may actually have a wide range in the size of their dense breast areas.[10] Likewise, 39.8% of our participants had concordant classifications for quartiles of dense breast area and percent density. Thus when one studies percent breast density as an intermediate endpoint, the results are also reflective of associations with BMI and do not necessarily reflect unique effects of the exposure being evaluated on the dense breast area.[10] Our results, which show no longitudinal relationship between anthropometry and the dense breast area, yet a strong negative longitudinal relationship between anthropometry and percent density, support the recommendation by Haars et al. that the dense breast area be used as the outcome in studies using mammographic breast density to make inference to breast cancer etiology.[10] We add to their recommendation that the non-dense area should be considered etiologically relevant to breast cancer as well.

Strengths of this study include its large sample size and multi-ethnic, population-based cohort. Also, menopausal status and use of HT were carefully monitored in SWAN. Quantitative measurements of mammographic breast density were used, which are preferable to the qualitative and subjective measurements used in many previous studies. The high reliability of the single reader of the mammograms is also a substantial strength. Finally, we had the opportunity to demonstrate consistent findings when analyzing groups defined by race, exogenous hormone use, and menopausal status.

Limitations to this study include potential residual confounding, despite careful adjustment for confounders. Also, the participants in the SWAN Mammographic Density Substudy are not a representative sample of the areas from which they were recruited, and this

may limit the external validity of these results. The most significant limitation is that the mammograms were not taken at the same time as the SWAN visits. Therefore we used linear interpolation with multiple imputation of random noise terms to estimate the participant's breast density at the time of her SWAN visit if the mammogram was not taken within 90 days of the nearest SWAN visit. We have provided and in depth description and validation of our interpolation and imputation method in a separate manuscript (Reeves et al., in preparation)[*]. Further, we observed similar results to those observed in the entire cohort when we repeated analyses among women with the majority of their mammograms occurring within 90 days of a SWAN visit. Future studies may benefit from incorporating mammograms into their study visits to avoid the estimation that was required in this study.

This study provides evidence that changes in anthropometry are not longitudinally associated with changes in the dense breast area, yet are strongly associated with percent breast density, at least among women transitioning through menopause. Our findings suggest that as a surrogate for breast cancer, the absolute dense breast area is likely to be the most relevant outcome, though the non-dense area may be important to disease etiology as well.

[*] Reeves KW, Stone RA, Modugno F, Ness RB, Vogel VG, Weissfeld JL, Habel L, Vuga M, Cauley JA. Linear Interpolation with Multiple Imputation to Account for Off-schedule Observations in a Longitudinal Study. In preparation.

3.6 ACKNOWLEDGEMENTS

We would like to thank SWAN participants, study staff, and personnel at contributing mammography facilities. We also thank Dr. Joseph Costantino and the NSABP for providing Gail score calculations.

Clinical Centers: University of Michigan, Ann Arbor - MaryFran Sowers, PI; Massachusetts General Hospital, Boston, MA - Robert Neer, PI 1994 - 1999; Joel Finkelstein, PI 1999- present; Rush University, Rush University Medical Center, Chicago, IL - Lynda Powell, PI; University of California, Davis/Kaiser - Ellen Gold, PI; University of California, Los Angeles - Gail Greendale, PI; University of Medicine and Dentistry - New Jersey Medical School, Newark –Gerson Weiss, PI 1994 – 2004; Nanette Santoro, PI 2004 – present; and the University of Pittsburgh, Pittsburgh, PA - Karen Matthews, PI.

NIH Program Office: National Institute on Aging, Bethesda, MD - Marcia Ory 1994 – 2001; Sherry Sherman 1994 – present; National Institute of Nursing Research, Bethesda, MD – Program Officers.

Central Laboratory: University of Michigan, Ann Arbor - Daniel McConnell; (Central Ligand Assay Satellite Services).

Coordinating Center: New England Research Institutes, Watertown, MA - Sonja McKinlay, PI 1995 – 2001; University of Pittsburgh, Pittsburgh, PA – Kim Sutton-Tyrrell, PI 2001 – present.

Steering Committee: Chris Gallagher, Chair; Susan Johnson, Chair

3.7 LITERATURE CITED

1. American Cancer Society. *Cancer Facts & Figures 2007.* Atlanta: American Cancer Society, Inc.; 2007.

2. Cooper R, Kaanders JH. Biological surrogate end-points in cancer trials: potential uses, benefits and pitfalls. *Eur J Cancer.* 2005;41(9):1261-1266.

3. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2006;15(6):1159-1169.

4. van den Brandt PA, Spiegelman D, Yaun SS, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol.* 2000;152(6):514-527.

5. Bartow SA, Pathak DR, Mettler FA, Key CR, Pike MC. Breast mammographic pattern: a concatenation of confounding and breast cancer risk factors. *Am J Epidemiol.* 1995;142(8):813-819.

6. Brisson J, Morrison AS, Kopans DB, et al. Height and weight, mammographic features of breast tissue, and breast cancer risk. *Am J Epidemiol.* 1984;119(3):371-381.

7. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer. *Cancer Causes Control.* 2000;11(7):653-662.

8. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 1998;7(12):1133-1144.

9. Gram IT, Funkhouser E, Tabar L. Anthropometric indices in relation to mammographic patterns among peri-menopausal women. *Int J Cancer.* 1997;73(3):323-326.

10. Haars G, van Noord PA, van Gils CH, Grobbee DE, Peeters PH. Measurements of breast density: no ratio for a ratio. *Cancer Epidemiol Biomarkers Prev.* 2005;14(11 Pt 1):2634-2640.

11. Lam PB, Vacek PM, Geller BM, Muss HB. The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont. *Cancer.* 2000;89(2):369-375.

12. McCormack VA, dos Santos Silva I, De Stavola BL, et al. Life-course body size and perimenopausal mammographic parenchymal patterns in the MRC 1946 British birth cohort. *Br J Cancer.* 2003;89(5):852-859.

13. Cruz-Korchin N, Korchin L, Gonzalez-Keelan C, Climent C, Morales I. Macromastia: how much of it is fat? *Plast Reconstr Surg.* 2002;109(1):64-68.

14. Maskarinec G, Meng L, Ursin G. Ethnic differences in mammographic densities. *Int J Epidemiol.* 2001;30(5):959-965.

15. Habel LA, Capra AM, Oestreicher N, et al. Mammographic density in a multiethnic cohort. *Menopause.* 2007;14(5):891-899.

16. Boyd NF, Greenberg C, Lockwood G, et al. Effects at two years of a low-fat, high-carbohydrate diet on radiologic features of the breast: results from a randomized trial. Canadian Diet and Breast Cancer Prevention Study Group. *J Natl Canc Inst.* 1997;89(7):488-496.

17. Maskarinec G, Pagano I, Lurie G, Kolonel LN. A longitudinal investigation of mammographic density: the multiethnic cohort. *Cancer Epidemiol Biomarkers Prev.* 2006;15(4):732-739.

18. Sowers M, Crawford S, Sternfeld B, et al. SWAN: A Multicenter, Multiethnic, Community-Based Cohort Study of Women and the Menopausal Transition. In: Lobo RA, Kelsey JL, Marcus R, eds. *Menopause : Biology and Pathobiology*. San Diego, CA: Academic Press; 2000:175-188.

19. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst.* 1999;91(18):1541-1548.

20. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81(24):1879-1886.

21. Royston P. Multiple imputation of missing values. *Stata J.* 2004;4(3):227-241.

22. Rubin DB. *Multiple Imputation for Non-response in Surveys*. New York: John Wiley & Sons; 2004.

23. Jeffreys M, Warren R, Gunnell D, McCarron P, Smith GD. Life course breast cancer risk factors and adult breast density. *Cancer Causes Control.* 2004;15(9):947-955.

24. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. *Br J Cancer.* 1998;78(9):1233-1238.

25. Key TJ, Allen NE, Verkasalo PK, Banks E. Energy balance and cancer: the role of sex hormones. *Proc Nutr Soc.* 2001;60(1):81-89.

26. Lorincz AM, Sukumar S. Molecular links between obesity and breast cancer. *Endocr Relat Cancer.* 2006;13(2):279-292.

Table 3.1 Characteristics of the study population at SWAN enrollment, N=834

Baseline characteristic	Total N=834	Pittsburgh N=198	Los Angeles N=321	Oakland N=315	P value*
General characteristics					
Age, years; mean (SD)	46.5 (2.7)	46.1 (2.5)	46.7 (2.7)	46.5 (2.7)	0.09
Race/ethnicity; N (%)					<0.001
African American	62 (7.4)	62 (31.3)	0 (0.0)	0 (0.0)	
Caucasian	407 (48.8)	136 (68.7)	138 (43.0)	133 (42.2)	
Chinese	182 (21.8)	0 (0.0)	0 (0.0)	182 (57.8)	
Japanese	183 (21.9)	0 (0.0)	183 (57.0)	0 (0.0)	
Family income; N (%)					<0.001
<$35,000	126 (15.4)	50 (25.5)	25 (8.0)	51 (16.4)	
$35,000-$49,999	140 (17.1)	39 (19.9)	38 (12.2)	63 (20.2)	
$50,000-$74,999	212 (25.9)	53 (27.0)	74 (23.8)	85 (27.2)	
$75,000-$99,999	141 (17.2)	32 (16.3)	62 (19.9)	47 (15.1)	
≥ $100,000	200 (24.4)	22 (11.2)	112 (36.0)	66 (21.2)	
Education; N (%)					0.001
≤ High school	136 (16.3)	38 (19.2)	38 (11.8)	60 (19.1)	
>High school	250 (30.0)	65 (32.8)	112 (34.9)	73 (23.2)	
College	223 (26.7)	38 (19.2)	95 (29.6)	90 (28.6)	
Post-college	225 (27.0)	57 (28.8)	76 (23.7)	92 (29.2)	
History of any cancer; N (%)	5 (0.6)	0 (0.0)	1 (0.3)	4 (1.3)	0.26†
Anthropometric characteristics					
Body mass index, kg/m^2; mean (SD)	25.4 (5.9)	28.4 (6.0)	23.9 (4.7)	25.2 (6.2)	<0.001
Underweight: <18.5 kg/m^2; N (%)	16 (1.9)	0 (0.0)	9 (2.9)	7 (2.2)	<0.001
Normal: 18.5 - <25.0 kg/m^2; N (%)	474 (57.5)	63 (32.1)	216 (68.4)	195 (62.5)	
Overweight: 25.0 - <30.0 kg/m^2; N (%)	202 (24.5)	71 (36.2)	64 (20.3)	67 (21.5)	
Obese: ≥30.0 kg/m^2; N (%)	132 (16.0)	62 (31.6)	27 (8.5)	43 (13.8)	
Weight, kg; mean (SD)	66.3 (17.1)	75.6 (17.1)	61.5 (14.0)	65.4 (17.9)	<0.001
Reproductive history					
Age at menarche, years; N (%)					0.004
<12	171 (20.6)	53 (26.9)	74 (23.2)	44 (14.0)	
12	234 (28.2)	55 (27.9)	91 (28.5)	88 (28.0)	
13	251 (30.2)	59 (30.0)	92 (28.8)	100 (31.9)	
≥14	174 (21.0)	30 (15.2)	62 (19.4)	82 (26.1)	
Age at first birth, years; N (%)					<0.001
Not applicable	148 (17.8)	28 (14.1)	58 (18.1)	62 (19.8)	
<20	55 (6.6)	29 (14.7)	13 (4.1)	13 (4.1)	
20-24	160 (19.2)	52 (26.3)	55 (17.2)	53 (16.9)	
25-29	235 (28.3)	46 (23.2)	95 (29.7)	94 (29.9)	
30-34	145 (17.4)	28 (14.1)	61 (19.1)	56 (17.8)	
≥ 35	89 (10.7)	15 (7.6)	38 (11.9)	36 (11.5)	

Table 3.1 (continued)

Cumulative breastfeeding, months; N (%)					<0.001
Nulliparous, never	148 (17.8)	28 (14.1)	58 (18.1)	62 (19.8)	
Parous, never	138 (16.6)	53 (26.8)	30 (9.4)	55 (17.5)	
1-4 months	145 (17.4)	32 (16.2)	59 (18.4)	54 (17.2)	
5-11 months	135 (16.2)	37 (18.7)	57 (17.8)	41 (13.1)	
12-22 months	141 (17.0)	24 (12.1)	63 (19.7)	54 (17.2)	
≥ 23 months	125 (15.0)	24 (12.1)	53 (16.6)	48 (15.3)	
Number of births; N (%)					0.22
0	148 (17.8)	28 (14.1)	58 (18.1)	62 (19.7)	
1	138 (16.6)	32 (16.2)	57 (17.8)	49 (15.6)	
2	351 (42.1)	78 (39.4)	138 (43.1)	135 (42.9)	
≥ 3	196 (23.5)	60 (30.3)	67 (20.9)	69 (21.9)	
Menopausal status; N (%)					0.91
Premenopausal	483 (58.3)	112 (57.1)	189 (59.1)	182 (58.2)	
Early Perimenopausal	346 (41.7)	84 (42.9)	131 (40.9)	131 (41.9)	
Ever used birth control pills; N (%)	627 (75.4)	152 (77.2)	85 (26.5)	239 (76.1)	0.60
Ever used hormones other than birth control pills; N (%)	110 (13.2)	23 (13.3)	42 (13.2)	42 (14.4)	0.99
Other characteristics					
Number of 1st degree relatives with breast cancer; N (%)					0.99
0	755 (91.2)	180 (90.9)	291 (91.2)	284 (91.3)	
≥ 1	73 (8.8)	18 (9.1)	28 (8.8)	27 (8.7)	
Number of 2nd degree relatives with breast cancer; N (%)					0.12
0	657 (79.4)	145 (73.2)	255 (79.9)	257 (82.6)	
1	132 (15.9)	39 (19.7)	50 (15.7)	43 (13.8)	
≥ 2	39 (4.7)	14 (7.1)	14 (4.4)	11 (3.5)	
Number of breast biopsies; N (%)					0.02
0	730 (87.5)	171 (86.4)	279 (86.9)	280 (88.9)	
1	76 (9.1)	20 (10.1)	24 (7.5)	32 (10.2)	
≥ 2	28 (3.4)	7 (3.5)	18 (5.6)	3 (1.0)	
Gail score; mean (SD)	1.06 (0.45)	0.95 (0.5)	1.13 (0.5)	1.05 (0.4)	<0.001
<1.66; N (%)	756 (91.8)	182 (92.9)	284 (89.6)	290 (93.3)	0.20
≥1.66; N (%)	68 (8.3)	14 (7.1)	33 (10.4)	21 (6.8)	

*P values from two-sample t tests or Kruskal-Wallis tests for continuous variables and chi-square tests for categorical variables

†P value from Fisher's Exact Test

Table 3.2 Summary statistics for annual change in anthropometric and breast density measures from participants' first to last observations with mammogram data

	N	Mean (SD)	25^{th} – 75^{th} Percentile	P Value*
Annual change in BMI (kg/m^2)†				
Total population	802	0.09 (0.70)	-0.10 – 0.33	
Race/ethnicity				0.05
African American	59	0.05 (1.26)	-0.29 – 0.53	
Caucasian	389	0.10 (0.79)	-0.10 – 0.39	0.18
Pittsburgh	132	0.11 (0.73)	-0.15 – 0.45	
Oakland	133	0.08 (0.72)	-0.13 – 0.33	
Los Angeles	124	0.12 (0.90)	-0.05 – 0.45	
Chinese	176	0.06 (0.37)	-0.10 – 0.22	
Japanese	178	0.12 (0.44)	-0.07 – 0.31	
BMI category at SWAN enrollment				0.51
Underweight: <18.5 kg/m^2	15	0.16 (0.22)	0.04 – 0.28	
Normal: 18.5 - <25.0 kg/m^2	460	0.14 (0.45)	-0.04 – 0.32	
Overweight: 25.0 - <30.0 kg/m^2	196	0.11 (0.51)	-0.16 – 0.38	
Obese: ≥30.0 kg/m^2	126	-0.03 (1.32)	-0.42 – 0.42	
Annual change in weight (kg) †				
Total population	807	0.22 (1.84)	-0.25 – 0.89	
Race/ethnicity				0.03
African American	60	0.12 (3.13)	-0.50 – 1.47	
Caucasian	390	0.23 (2.16)	-0.33 – 1.06	0.16
Pittsburgh	133	0.22 (1.88)	-0.33 – 1.19	
Oakland	133	0.21 (1.83)	-0.38 – 0.91	
Los Angeles	124	0.26 (2.66)	-0.10 – 1.18	
Chinese	178	0.16 (0.86)	-0.24 – 0.52	
Japanese	179	0.29 (1.10)	-0.16 – 0.80	
BMI category at SWAN enrollment				0.32
Underweight: <18.5 kg/m^2	15	0.44 (0.67)	0.12 – 0.81	
Normal: 18.5 - <25.0 kg/m^2	462	0.35 (1.12)	-0.08 – 0.83	
Overweight: 25.0 - <30.0 kg/m^2	196	0.24 (1.38)	-0.46 – 0.97	
Obese: ≥30.0 kg/m^2	126	-0.14 (3.51)	-0.97 – 1.13	
Annual change in dense area (cm^2)				
Total population	834	-0.58 (3.32)	-1.70 – 0.59	
Race/ethnicity				0.03
African American	62	-0.46 (7.22)	-3.21 – 3.27	
Caucasian	407	-0.84 (3.36)	-2.18 – 0.45	0.54
Pittsburgh	133	-1.06 (4.18)	-2.30 – 0.35	
Oakland	138	-0.76 (3.04)	-2.05 – 0.54	
Los Angeles	136	-0.69 (2.71)	-2.01 – 0.45	
Chinese	182	-0.39 (1.51)	-1.46 – 0.35	
Japanese	183	-0.22 (2.27)	-1.23 – 0.69	
BMI category at SWAN enrollment				0.28
Underweight: <18.5 kg/m^2	16	-0.0003 (1.25)	-0.78 – 1.01	
Normal: 18.5 - <25.0 kg/m^2	474	-0.61 (2.53)	-1.66 – 0.36	
Overweight: 25.0 - <30.0 kg/m^2	202	-0.61 (3.59)	-1.80 – 0.91	
Obese: ≥30.0 kg/m^2	132	-0.50 (5.16)	-2.17 – 1.28	

Table 3.2 (continued)

Annual change in percent density (%)				
Total population	834	-1.01 (2.38)	-2.16 – 0.13	
Race/ethnicity				0.07
African American	62	-0.52 (3.75)	-2.07 – 1.00	
Caucasian	407	-1.22 (2.17)	-2.43 – -0.04	0.99
Pittsburgh	133	-1.30 (2.79)	-2.39 – 0.08	
Oakland	138	-1.20 (1.86)	-2.06 – -0.08	
Los Angeles	136	-1.16 (1.70)	-2.45 – -0.09	
Chinese	182	-0.91 (1.96)	-2.07 – 0.08	
Japanese	183	-0.81 (2.60)	-2.06 – 0.27	
BMI category at SWAN enrollment				<0.001
Underweight: <18.5 kg/m^2	16	-1.14 (2.42)	-2.86 – 0.38	
Normal: 18.5 - <25.0 kg/m^2	474	-1.16 (2.35)	-2.56 – -0.11	
Overweight: 25.0 - <30.0 kg/m^2	202	-1.03 (2.10)	-1.87 – 0.10	
Obese: ≥30.0 kg/m^2	132	-0.50 (2.78)	-1.08 – 0.43	

* P values from Kruskal-Wallis test due to non-normality and heteroskedasticity

†Number of observations for annual change in BMI and weight are <834 because some participants were missing height and/or weight data at their first or last study visits with mammogram data

Table 3.3 Summary statistics of participants' initial mammographic breast density measurements, N=834[*]

		Dense Breast Area (cm^2)			Percent Density (%)		
	N	**Mean (SD)**	**25^{th} – 75^{th} Percentile**	**P value**	**Mean (SD)**	**25^{th} – 75^{th} Percentile**	**P value**
Total population	834	46.2 (26.7)	28.9 – 59.2		42.3 (19.6)	29.3 – 57.4	
Body mass index category				0.01			<0.001
Underweight: <18.5 kg/m^2	15	32.5 (12.9)	21.3 – 42.2		61.9 (19.0)	45.4 – 77.6	
Normal: 18.5 – <25.0 kg/m^2	453	43.8 (23.4)	28.8 – 54.2		50.7 (16.8)	39.1 – 63.3	
Overweight: 25.0 – <30.0 kg/m^2	206	50.0 (25.3)	32.9 – 62.9		39.1 (16.1)	27.8 – 50.1	
Obese : ≥30.0 kg/m^2	142	49.8 (36.7)	22.7 – 73.1		23.7 (16.5)	10.2 – 35.0	
Weight, kg				0.002			<0.001
1^{st} Quartile: 39 – <55.0	207	39.0 (19.9)	25.4 – 48.7		54.6 (16.5)	42.8 – 68.0	
2^{nd} Quartile: 55.0 – <63.2	208	47.5 (24.6)	32.8 – 58.9		49.0 (16.3)	38.3 – 61.3	
3^{rd} Quartile: 63.2 – <73.8	201	48.3 (24.4)	31.2 – 60.9		41.9 (17.1)	28.8 – 53.9	
4^{th} Quartile: 73.8 – 153.9	205	50.1 (34.4)	25.1 – 67.4		27.5 (17.4)	12.8 – 41.2	

[*]Mammographic and personal characteristics are from the first timepoint at which the participant has mammographic density values, averaged across all imputations; in some cases the first timepoint did not correspond to the enrollment visit, thus the distributions of BMI and weight presented here differ from those in Table 3.1

[†]P values from ANOVA across groups using the following transformations: square root (dense breast area), untransformed (percent density); means, SD, and range are all reported in the natural scale

Table 3.4 Random effects regression estimates for the outcomes of dense area and percent density using multiple imputation[*]

	N	β[†]	Standard Error	95% CI	P Value
Dense breast area					
Body mass index, kg/m^2					
Model 1: BMI	830	-0.0152	0.0097	-0.0348 – 0.0043	0.12
Model 2: BMI + age	830	0.00003	0.0101	-0.0205 – 0.0205	0.99
Model 3: Fully adjusted[**]	824	-0.0105	0.0108	-0.0327 – 0.0117	0.34
Weight, kg					
Model 1: Weight	830	-0.0055	0.0036	-0.0129 – 0.0019	0.14
Model 2: Weight + age	830	-0.0004	0.0038	-0.0082 – 0.0073	0.91
Model 3: Fully adjusted[**]	824	-0.0055	0.0042	-0.0142 – 0.0031	0.20
Percent breast density					
Body mass index, kg/m^2					
Model 1: BMI	830	-1.4601	0.0811	-1.6206 – -1.2996	<0.001
Model 2: BMI + age	830	-1.2999	0.0858	-1.4710 – -1.1289	<0.001
Model 3: Fully adjusted[‡]	823	-1.1845	0.0934	-1.3714 – -0.9976	<0.001
Weight, kg					
Model 1: Weight	830	-0.5170	0.0288	-0.5740 – -0.4600	<0.001
Model 2: Weight + age	830	-0.4675	0.0301	-0.5273 – -0.4076	<0.001
Model 3: Fully adjusted[‡]	823	-0.4374	0.0341	-0.5056 – -0.3692	<0.001

[*]Dense area was square root transformed due to non-normality; percent density was modeled in the natural scale

[†]Regression coefficients have the following units: $\sqrt{cm^2}/(kg/m^2)$ for regression of body mass index on dense breast area; $\sqrt{cm^2}/kg$ for regression of weight on dense breast area; $\%/(kg/m^2)$ for regression of body mass index on percent density; %/kg for regression of weight on percent density

[**]Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

[‡]Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

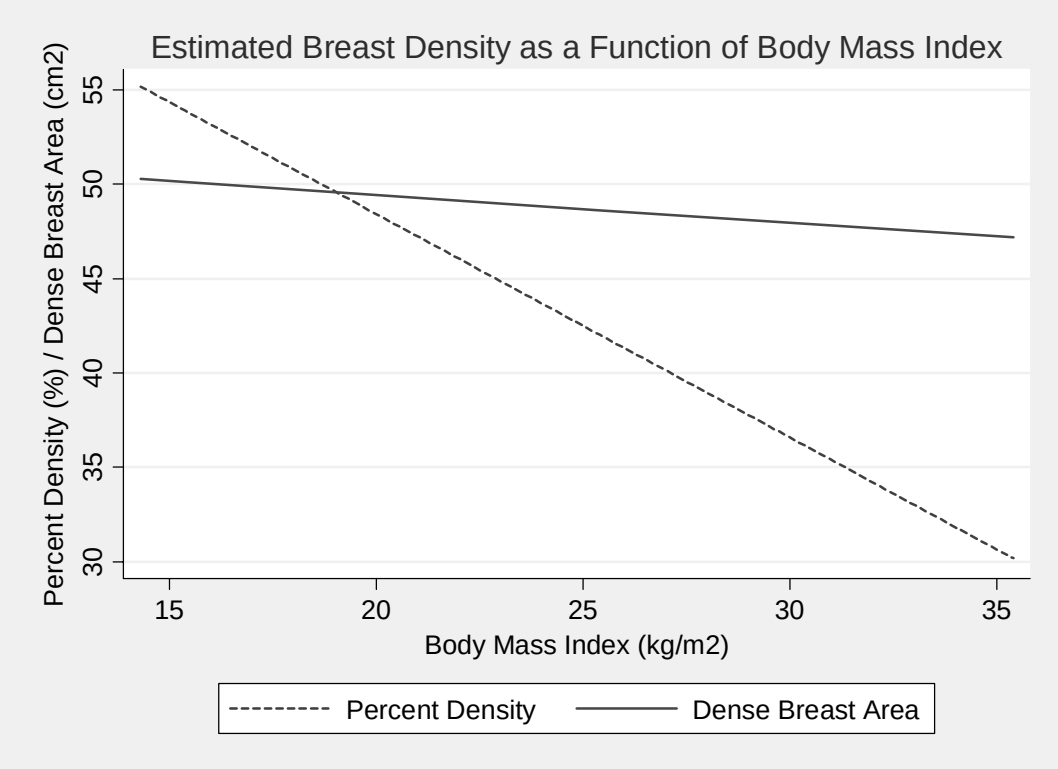


Figure 3.1 Dense breast area and percent density as a function of body mass index estimated from multivariable random intercept regression models[*]

[*]Values were estimated by varying the values of body mass index while keeping all other variables in the model fixed at their mean or referent values. For dense breast area, the values are for premenopausal, Caucasian women from the Pittsburgh site age 46.5 years with no 1st degree family history of breast cancer, no hormone use since previous visit, and no previous breast biopsies. For percent density, the values are for premenopausal, Caucasian women from the Pittsburgh site age 46.5 years with no previous use of oral contraceptives at enrollment, a high school or lower education, no hormone use since previous visit, no previous breast biopsies, and who were nulliparous and age <12 at menarche.

4.0 ARTICLE 2: LINEAR INTERPOLATION WITH MULTIPLE IMPUTATION TO ACCOUNT FOR OFF-SCHEDULE OBSERVATIONS IN A LONGITUDINAL STUDY

Katherine W. Reeves, MPH,[1] Roslyn A. Stone, PhD,[2] Francesmary Modugno, PhD, MPH,[1] Roberta B. Ness, MD, MPH,[1] Victor G. Vogel, MD, MHS,[3] Joel L. Weissfeld, MD, MPH,[1,4] Laurel Habel, PhD,[5] Marike Vuga, PhD,[1] Jane A. Cauley, DrPH, MPH[1]

[1]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[2]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[3]Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

[4]University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania

[5]Division of Research, Kaiser Permanente, Oakland, California

This ancillary study was supported by National Cancer Institute grant R01 CA89552. The Study of Women's Health Across the Nation (SWAN) has grant support from the National Institutes of Health, DHHS, through the National Institute on Aging, the National Institute of Nursing Research and the NIH Office of Research on Women's Health (Grants NR004061; AG012505, AG012535, AG012531, AG012539, AG012546, AG012553, AG012554, AG012495).

4.1 ABSTRACT

Data in epidemiological studies are sometimes collected off-schedule from planned study visits. In an ancillary study, longitudinal outcome data were collected retrospectively from mammograms that were not acquired at the study visits. This created a missing data problem because the outcome of interest, breast density, was unknown at the time of the study visits when covariate data were collected. We developed a method to estimate the off-schedule mammographic breast density measurements at study visits using a novel approach of linear interpolation combined with multiple imputation. We evaluated the validity of this approach by using it to estimate known values of breast density and comparing the estimated values to the observed values. We compared results of random intercept models assessing the association between body mass index (BMI) and dense breast area when breast density was estimated with our approach to results obtained by simply matching each mammogram to the nearest study visit. Our method had a small bias on average (0.11 cm^2). The association between BMI and dense breast area was statistically significant when estimation was based on simple matching (β=-0.0155 p=0.04), yet was non-significant when based on interpolation and multiple imputation (β=-0.0098, p=0.38). Simple matching may produce inaccurate estimates because it does not incorporate the time difference and change in breast density over time. Our method of linear interpolation with multiple imputation may be applicable to other longitudinal datasets where important data were collected outside the scheduled study visits and the variable of interest changes linearly over time.

4.2 INTRODUCTION

There are many situations in epidemiologic research where data are collected at times other than those planned in the initial study design. Some studies may plan regular visits to collect data, yet additional data are collected at other times and may be considered off-schedule. This can occur when longitudinal studies use retrospectively collected data, which are unlikely to exactly match the timing of the study visits, in addition to data from the regularly scheduled visits. For example, medical records or other routinely collected medical data could be used in ancillary studies conceived after data collection for the main study has begun. To address this issue in our own work, we propose a novel method for estimating the value of off-schedule measurements at the time of the study visits.

We encountered the need to estimate breast density in the context of an ancillary study to the Study of Women's Health Across the Nation (SWAN), the SWAN Mammographic Density Substudy. In this ancillary study, routine screening mammograms from a subset of SWAN participants were collected retrospectively for measures of breast density. Though these mammograms were taken during the period of time in which women were actively participating in SWAN, the timing of the mammograms rarely coincided with the timing of the SWAN visits, when other data of interest were collected. This created a missing data problem, because the values of the breast density measurements at the time of the SWAN visits were mostly unobserved. Thus, a method of estimating the mammogram data at the SWAN visits was necessary. Simply matching each mammogram to the nearest SWAN visit was problematic, due to the high degree of both between- and within-subject variability in the timing of the mammograms and study visits. A simple matching algorithm would not account for either the time between study visits and mammograms or changes in breast density occurring over those

periods. Addressing this problem was crucial to our planned analysis using the longitudinal mammographic density data as the outcome in a study where the primary independent variable was body mass index (BMI). Height and weight were measured at the annual SWAN visits and were used to calculate BMI as the weight in kilograms divided by the squared height in meters.

To our knowledge, methods for estimating retrospectively collected, supplementary data at study visits have not been reported. In this paper we describe a method to estimate off-schedule outcome data at the time of the study visits using techniques commonly employed for handling missing data. We used linear interpolation with multiple imputation to estimate breast density measurements at the time of SWAN visits based on the observed mammogram data.

4.3 METHODS

4.3.1 Study population

SWAN is a prospective cohort study focused on the health of women as they transition from premenopause to postmenopause.[1] Eligibility criteria for SWAN were age 42-52, ≥1 menstrual period in the previous three months, intact uterus and ≥1 intact ovary, and not currently on hormone therapy (HT) or taking oral contraceptives (OCs). The SWAN Mammographic Density Substudy is an ancillary study to SWAN with the goal of examining how mammographic breast density changes as women transition through menopause.[2, 3] Participants from three SWAN sites (Los Angeles, CA, Oakland, CA, and Pittsburgh, PA) were enrolled at their 5th or 6th annual visit. At the time of this analysis, data were available from the baseline visit and up to seven follow-up visits for each participant. Of those eligible for the SWAN Mammographic Density Substudy,

86.1% (N=1,055) consented and at least one mammogram was retrieved for 95.3% of these women (N=1,007 participants, with 3,980 mammograms). Women who enrolled in the SWAN Mammographic Density Substudy were of similar age, leaner, and less likely to be African American or Caucasian than women who were eligible but chose not to participate. Participants provided written informed consent, and institutional review board approvals were obtained for both SWAN and the SWAN Mammographic Density Substudy.

Participants were excluded from this analysis if they reported a history of breast cancer at SWAN enrollment (N=6) or had <2 available mammograms (N=157). Women diagnosed with breast cancer during their SWAN follow-up (N=21) were censored at the time of diagnosis; ten of these women had no mammograms prior to their breast cancer diagnosis and were excluded. Mammograms of breasts that had undergone a biopsy or more extensive surgery were not read for density in SWAN. After these exclusions, 834 participants and 3,746 mammograms remained in the analysis. Women enrolled in the SWAN Mammographic Density Substudy but excluded from this analysis were of similar age, educational level, and menopausal status as those included.

4.3.2 Description of data

The collection of mammograms in the SWAN Mammographic Density Substudy has been described previously.[2, 3] Briefly, starting after the time of the 6th annual SWAN visit, investigators obtained routine screening mammograms taken two years prior to the baseline SWAN visit through two years after the 6th annual SWAN visit. Because these mammograms were not taken as part of the SWAN study, but rather depended upon each woman's compliance with current breast cancer screening guidelines, the number of and interval between

mammograms obtained on each participant was highly variable. The number of mammograms collected on the 834 participants in the present analysis ranged from 2 to 10 with a median of 4.

Breast density measurements were performed by a single reader using manual planimetry.[2, 3] The reader provided measurements of the total breast area and the dense breast area, as well as a rating of the quality of the mammogram film. A blinded 10% random sample of films showed good association between the initial and repeat readings of dense area (within-person Spearman correlation coefficient=0.96). For the purpose of illustrating the statistical methods to estimate the mammogram measures at the time of SWAN visits, we focused on the dense breast area.

The timing of each participant's mammograms rarely coincided with the timing of her SWAN visits; mammograms and visits occurred on the same day for only 10 of 3,746 mammograms (0.27%). To illustrate the method, these coincident mammograms were estimated as well. The length of time between each SWAN visit and the closest mammogram was quite variable, both between and within participants. To illustrate the complexity of the available data for this analysis, Figure 4.1 shows a timeline of the SWAN visit and mammogram data on a representative participant. This participant, referred to as Participant X, is used throughout this paper to illustrate the implementation of estimation algorithms. Noteworthy features of this participant's data include the lack of mammograms near the time of some SWAN visits (Baseline through Visit 2) and the variable timing of the mammograms in relation to the nearest SWAN visit (Mammogram 1, 37 days before Visit 3; Mammogram 2, 92 days after Visit 4; Mammogram 3, 135 days after Visit 5; Mammogram 4, 134 days after Visit 6).

4.3.3 Approaches for estimating dense breast area

We considered five approaches for estimating breast density at the visit times: 1) a simple matching algorithm to match each mammogram to its nearest visit, 2) linear interpolation, 3) linear interpolation with addition of a singly imputed noise term, 4) linear interpolation with multiply imputed noise terms, and 5) linear interpolation with multiply imputed noise terms for visits >90 days from nearest mammogram and matching otherwise.

For Approach 1, the matching occurred regardless of whether the mammogram preceded or followed the study visit and regardless of the time between the two events. Because we believe breast density was unlikely to undergo substantial changes within a few months time, matches were maintained for mammograms taken within 90 days of a SWAN visit for Approach 5.

To implement the linear interpolation, mammograms and SWAN visits were ordered chronologically. Data from the mammograms before and after a target SWAN visit were used to estimate breast density at that SWAN visit according to the following equation:

$$D_{i,t_j} = D_{i,t_-} + s(D_{i,t_+} - D_{i,t_-})$$

with

$$s = \frac{A_{i,t_j} - A_{i,t_-}}{A_{i,t_+} - A_{i,t_-}}$$

where D is the dense breast area, i is a unique indicator for each participant, t denotes time with t_j representing the time of the missing data (i.e., time of target SWAN visit) and t_- and t_+ representing the times of the mammograms before and after t_j, respectively. A scaling factor (s) for the time between the visit and the nearest mammograms was computed using age (A) at the

time of these events, measured in years to the first decimal place (using a similar notation as for *D*). This method was based on the assumption that the change in breast density was linear between the two mammograms. We did not use extrapolation to estimate breast density; therefore, visits without a mammogram both before and after the target SWAN visit time were not included in the analysis. The linear interpolation was implemented using the "ipolate" command in Stata version 10.0 (Stata Corporation, College Station, TX).

To illustrate the use of the above equations, we calculated the dense breast area at Visit 4 for Participant X. Her age at Visit 4 was 46.2 ($A_{i,tj}$=46.2), at Mammogram 1 she was age 45.1 ($A_{i,t-}$=45.1), and at Mammogram 2 she was age 46.4 ($A_{i,t+}$=46.4). Her measured dense breast area was 93.6 cm^2 at Mammogram 1 ($D_{i,t-}$=93.6) and 90.2 cm^2 at Mammogram 2 ($D_{i,t+}$=90.2). Substituting these values into the above equations, we find:

$$s = \frac{46.2 - 45.1}{46.4 - 45.1} = 0.85$$

and

$$D_{i,t_j} = 93.6 + (0.8462) * (90.2 - 93.6) = 90.7$$

Therefore the interpolated dense breast area at Visit 4 for Participant X was 90.72 cm^2.

Additionally, for Approaches 3-5 we added a noise term to each interpolated value to account for the error introduced by estimating breast density rather than measuring it at the time point corresponding to the SWAN visit. For each participant we calculated the standard deviation (SD) of all observed dense breast area measurements (i.e. the SD of the observed mammogram data). Person-specific normal distributions of noise terms were generated, which had a mean of 0 and SD equal to the participant's calculated SD of dense breast areas. Noise terms were randomly selected from this distribution and added to each interpolated value. For

Approaches 4 and 5, multiple imputation was used to create ten analytic datasets, each with different random noise terms added to the same linearly interpolated value of dense breast area. The literature on multiple imputation recommends 3-5 imputations for most situations but notes that up to 10 may be needed in situations with a higher level of missing data.[4, 5] Because we have used a novel application of multiple imputation, we explored the effect of the choice of number of imputations on the estimated variance of the regression coefficients calculated using up to 15 imputations. We chose to use ten imputations because in our data the variances changed little beyond 10 imputations (data not shown). Due to the addition of the random noise terms, some negative values for dense breast area were obtained (<1%) and were considered to be implausible. These implausible values were discarded and additional random noise terms were generated and added to the interpolated values until acceptable imputed values were obtained. Figure 4.2 displays the observed and estimated mammogram data, using Approach 5, for Participant X.

4.3.4 Design of validation study

Using the observed mammographic density data among participants with at least three mammograms, data for one mammogram between the first and last mammogram was randomly set to missing. Approaches 1-4 were each used to estimate the density measurements for the "missing" mammogram data. For descriptive purposes, in Approach 4 the estimated value was calculated as the average of the ten imputed values. Approach 5 was not considered because none of the participants had mammograms separated by less than 90 days. Bias was calculated by subtracting the observed value from the estimated value for each approach. We calculated summary statistics for bias for each approach. Large values of bias were used as indicators that

the approach provided poor estimates for a participant. We also used analysis of variance (ANOVA) to compare the magnitudes of the bias by the reported quality of the mammograms (poor, fair, good, or excellent). We assessed the correlation between estimates from each approach with the observed data and with estimates from the other approaches using Pearson's correlation coefficients.

4.3.5 Illustration of effects of choice of estimation approach

To demonstrate the impact of the choice of estimation approach on regression results, we developed an example exploring the longitudinal relationships between BMI and the dense breast area. A parsimonious random intercept regression model was built using the first analytic dataset. The regression was repeated using breast density estimated under each of the five approaches. For Approaches 4 and 5, this resulted in a set of ten regression coefficients and their variances. These estimates were combined into an overall multiple imputation estimate of the regression coefficient and its variance using Rubin's rules.[4, 6] The multiple imputation estimate of β, β^*, is calculated as the arithmetic mean of the β from the set of *m* imputations:

$$\beta^* = (1/m)\sum_{j=1}^{m}\beta_j$$

The variance of β^*, *T*, is calculated as a function of the within-imputation variance, *W*, and the between-imputation variance, *B*, as follows:

$$T = W + (1 + 1/m)B; \text{ where:}$$

$$W = (1/m)\sum_{j=1}^{m} W_j \qquad B = 1/(m-1)\sum_{j=1}^{m}(\beta_j - \beta^*)^2$$

Confidence intervals and p values can then be determined using a t distribution with degrees of freedom, *v*, calculated as:

$$v=(m-1)\left\{1+\frac{W}{(1+1/m)B}\right\}^2$$

The participant's age in days was used as the time scale in the regressions. Random intercept models were fit using the "xtreg" command in Stata version 10.0. Multiple imputation estimates of the regression coefficients and their variances were calculated using the "mijoin" and "micombine" commands in Stata. The dependent variable, estimated dense breast area, was square-root transformed to improve normality. Because race was confounded with site, we created a combined race/site variable with the following categories for inclusion in the model: Caucasian/Pittsburgh, African American/Pittsburgh, Caucasian/Oakland, Chinese/Oakland, Caucasian/Los Angeles, Japanese/Los Angeles. Each model was adjusted for age, race/site, menopausal status, first degree relative with history of breast cancer, number of previous breast biopsies, and hormone use since previous visit (Reeves et al., in preparation).[*] We compared the magnitudes of the regression coefficients and their variances across the five approaches considered. All statistical tests performed were two-sided with a $p \leq 0.05$ considered statistically significant.

[*] Reeves KW, Stone RA, Modugno F, Ness RB, Vogel VG, Weissfeld JL, Habel L, Sternfeld B, Cauley JA. Longitudinal influence of anthropometry on mammographic breast density: the Study of Women's Health Across the Nation (SWAN). In preparation.

4.4 RESULTS

4.4.1 Validation study

Table 4.1 describes the study population. At baseline participants were on average 46.5 years old, 58.3% were premenopausal, 57.5% had a BMI in the normal range (18.5 – <25 kg/m^2), and 48.8% were Caucasian.

The validation study included 710 participants with ≥3 mammograms. The bias for estimating dense breast area using Approaches 1–4 compared to the observed values is summarized in Figure 4.3. Mean bias was similar across approaches, yet the greatest variability in bias was observed with Approaches 1 and 3. For Approach 4, mean bias was 0.11 cm^2 and approximately 50% of the estimates were within 4 cm^2 of the observed measurement.

Figure 4.4 shows plots of the observed dense breast area versus the values estimated by each approach. Approaches 2 (r=0.96) and 4 (r=0.96) produced the estimates most highly correlated with the observed values, though correlations with the observed values were also high for Approaches 1 (r=0.94) and 3 (r=0.93). The estimates produced by the four approaches were highly correlated with one another (all r≥0.95), though some outliers were noted.

Figure 4.5 shows examples where Approach 4 did (Participant Y) and did not (Participant Z) work well. For Participant Y, the change in density across mammograms 4, 5, and 6 was approximately linear, and the multiple imputation estimate of dense breast area (32.2 cm^2) was very close to the observed value (32.8 cm^2). Our method worked poorly when density underwent a non-linear change across the estimation interval. We defined the method to work poorly when the bias was greater than 10 cm^2 and the estimated value differed from the observed by more than 10%. This applied to 14.8% (N=105) of the participants. This is illustrated by

Participant Z, for whom the estimated dense breast area for mammogram 5 was 67.4 cm^2 while the observed dense breast area was 122.8 cm^2 (bias=-55.4 cm^2). Interestingly, the quality of film for this mammogram was rated as poor, while the quality was rated higher for her other mammograms.

The majority of the estimated mammogram films were of excellent (N=492, 69.3%) or good (N=178, 25.1%) quality, though some were judged to be of fair (N=31, 4.4%) or poor (N=9, 1.3%) quality. The quality of the estimated film was more likely to be fair or poor in instances where the method worked poorly (10.5%), compared to instances where the method worked well (4.8% fair or poor, p=0.02). The bias from Approach 4 differed by film quality with the mean bias for poor (-3.60 cm^2) or fair (3.14 cm^2) quality films being larger than for those of good (0.22 cm^2) or excellent (-0.06 cm^2) quality (p=0.10 based on the ANOVA model).

4.4.2 Regression analyses

Table 4.2 displays the results of the random intercept regression models for each of the five estimation approaches. The association between BMI and square root transformed dense breast area was statistically significant only when dense breast area was estimated with the simple matching algorithm (Approach 1: β=-0.0155, p=0.04). The magnitude of the estimated regression coefficient was greatest when using the simple matching algorithm. The coefficient was smallest for the linear interpolation without noise terms (Approach 2: β=-0.0070). The two multiple imputation approaches resulted in similar estimated regression coefficients (Approach 4: β=-0.0098 and Approach 5: β=-0.0105).

The estimated variance was lowest for linear interpolation without the addition of noise terms (Approach 2: variance=2.73×10^{-5}) and was approximately doubled using either the simple

matching algorithm (Approach 1: variance=5.43×10^{-5}) or single imputation (Approach 3: variance=5.89×10^{-5}). Variance was highest using multiple imputation of the noise terms (Approach 4: variance=11.86×10^{-5}), and decreased slightly when observed data were used for mammograms within 90 days of a visit (Approach 5: variance=11.73×10^{-5}).

4.5 DISCUSSION

We have presented a procedure for estimating off-schedule outcome data at the times of scheduled study visits using linear interpolation with multiply imputed noise terms. Because mammographic density is generally a stable measure, linear interpolation was expected to work well. In a validation study using the observed SWAN mammogram data, our estimate measurements were, on average, very close to the observed data. Problems arose when the mammogram data exhibited a non-linear trend over the estimation interval. We also have demonstrated how variations on our proposed method affect the results of regression analyses.

Multiple imputation is recognized as a statistically valid method for obtaining unbiased estimates and appropriate variances in datasets with missing data.[5, 7-11] This technique has been used to estimate missing data in a wide variety of variables, including depression scores,[10] medical costs,[9] and serum cholesterol levels.[12] Multiple imputation performs better than other methods of handling missing data, such as analyzing only cases with complete data, carrying the last observation forward, or using only a single imputation.[7-14] For our method we used a novel application of multiple imputation as a way to account for uncertainty in the linear interpolation.

Though this paper is not intended to draw etiologic inference, it is important to note how one's conclusions might change based on the approach used to estimate dense breast area. The

estimated regression coefficient was largest and statistically significant only when using a simple matching algorithm. Though mammographic breast density is a relatively stable measure, substantial changes in breast density may occur over the course of a year due to changes in menopausal status,[15] use of hormone therapy,[16-18] and the aging process.[19-22] Thus the simple matching algorithm is unlikely to provide the best estimates of breast density measurements at the time of the study visits, and appeared to overestimate the effects of BMI on the dense breast area. Importantly, based on the simple matching algorithm we would infer that change in BMI was significantly associated with the dense breast area, while we would judge this association to be attenuated and non-significant when using our other approaches.

Using raw interpolated values or only a single imputation of the noise terms resulted in substantially different estimates of the regression coefficient than did multiple imputation. Though the variance was considerably larger when multiple imputation was used, relying on only a single imputation could lead to bias. As shown in Figure 4.3, estimating dense breast area using the single imputation approach resulted in a wider range of bias than did the multiple imputation approach. The multiply imputed noise terms provided a more appropriate reflection of the error introduced by estimating the data. These noise terms were randomly generated from distributions reflecting the variability of each participant's observed mammographic density data. However, the standard deviations used as the basis for the noise terms assume no trend in the mammogram data and are poorly estimated for women with fewer measurements. The assumed distribution of noise terms could greatly influence the results. Although other methods of choosing noise terms are possible, the relatively small number of mammograms for many participants limited our ability to consider more complicated functional forms.

In Approach 5, we preserved matches where the mammogram was taken within 90 days of the study visit. We evaluated the effect of this assumption by comparing results using this approach to those using linear interpolation with multiple imputation for all observations (Approach 4). The regression coefficients were very similar using either method, though the variances were slightly larger when no matches were maintained. Using the observed data in cases where the measurements were taken close to the time of the study visit appears to be appropriate in our example, although not necessarily elsewhere.

In cases such as ours that have missing data for a variable that is likely to change linearly over the study interval yet has short-term stability, we recommend using linear interpolation with multiple imputation while preserving closely aligned data (Approach 5) to account for off-schedule observations. Our proposed method has several advantages. It provides a way for investigators to estimate data collected outside of planned study visits with data collected at those visits when necessary for specific analyses. This increases the possibilities for ancillary studies and other analyses not planned at the onset of the study to utilize off-schedule data. Use of linear interpolation produces estimates based on the observed data. A previous simulation study demonstrated that linear interpolation was generally the best method for predicting missing values in time-oriented data.[23] Also, the multiply imputed noise terms increase the appropriateness of the variance estimates so that statistical inferences can be made that account for uncertainty. Further, this method is relatively easy to implement with the use of Stata routines to perform the interpolation, imputations, and combining of results according to Rubin's rules.[6]

The proposed method is not without its limitations, however. Some study visits were excluded as participants lacked mammograms both before and after the visit and we did not use

extrapolation to estimate the density measurements. This loss of observations may result in decreased power and raises the possibility of a selection bias. Error was introduced when measurements were estimated rather than observed, and was further, albeit purposefully, introduced by the addition of the noise terms. Thus our tests of statistical significance may be overly conservative. This method does not account for measurement error from the original density readings or error in measurements introduced by differences in the degree of breast compression from mammogram to mammogram. In at least one case where the method performed poorly (Participant Z), the quality of the film was rated as poor and possibly indicates measurement error.

Finally, our method does not perform well when there is a substantially non-linear trend in the measurements over the estimation interval. Our validation study, however, used observations to estimate the missing data that were typically a year before or after the timepoint with the missing data and were separated from one another by approximately two years. In reality, both mammograms and study visits occurred at approximately one-year intervals, with most participants having less than six months between a mammogram and its nearest study visit. Thus, the time period over which the estimates for the regression analyses were made was generally much shorter than those in the validation study. For this reason we believe that the level of error in the application of the method to real data is likely to be less than that observed in the validation study.

Issues with off-schedule data are likely to be present in other studies and longitudinal datasets. While we have developed and demonstrated our method in a study of mammographic breast density, the method is applicable to a wide variety of situations. For example, one might use the proposed method to estimate body weight at the time of a phone interview based on

measurements recorded at preceding and subsequent clinic visits. Further development and testing of our method may be necessary before implementing elsewhere. Our hope is that this method will allow investigators to conduct statistically valid analyses in datasets with similar missing data problems due to off-schedule data. As we have demonstrated, failure to account for the fact that observations are off-schedule observations can lead to incorrect conclusions about the magnitude and significance of associations.

4.6 ACKNOWLEDGEMENTS

We would like to thank SWAN participants, study staff, and personnel at contributing mammography facilities.

Clinical Centers: University of Michigan, Ann Arbor - MaryFran Sowers, PI; Massachusetts General Hospital, Boston, MA - Robert Neer, PI 1994 - 1999; Joel Finkelstein, PI 1999- present; Rush University, Rush University Medical Center, Chicago, IL - Lynda Powell, PI; University of California, Davis/Kaiser - Ellen Gold, PI; University of California, Los Angeles - Gail Greendale, PI; University of Medicine and Dentistry - New Jersey Medical School, Newark –Gerson Weiss, PI 1994 – 2004; Nanette Santoro, PI 2004 – present; and the University of Pittsburgh, Pittsburgh, PA - Karen Matthews, PI.

NIH Program Office: National Institute on Aging, Bethesda, MD - Marcia Ory 1994 – 2001; Sherry Sherman 1994 – present; National Institute of Nursing Research, Bethesda, MD – Program Officers.

Central Laboratory: University of Michigan, Ann Arbor - Daniel McConnell; (Central Ligand Assay Satellite Services).

Coordinating Center: New England Research Institutes, Watertown, MA - Sonja McKinlay, PI 1995 – 2001; University of Pittsburgh, Pittsburgh, PA – Kim Sutton-Tyrrell, PI 2001 – present.

Steering Committee: Chris Gallagher, Chair; Susan Johnson, Chair

4.7 LITERATURE CITED

1. Sowers M, Crawford S, Sternfeld B, et al. SWAN: A Multicenter, Multiethnic, Community-Based Cohort Study of Women and the Menopausal Transition. In: Lobo RA, Kelsey JL, Marcus R, eds. *Menopause : Biology and Pathobiology*. San Diego, CA: Academic Press; 2000:175-188.

2. Crandall CJ, Zheng Y, Karlamangla A, et al. The association between mammographic breast density and bone mineral density in the Study of Women's Health Across the Nation. *Ann Epidemiol.* 2007.

3. Habel LA, Capra AM, Oestreicher N, et al. Mammographic density in a multiethnic cohort. *Menopause.* 2007;14(5):891-899.

4. Rubin DB. *Multiple Imputation for Non-response in Surveys*. New York: John Wiley & Sons; 2004.

5. Schafer JL. Multiple imputation: a primer. *Statistical Methods in Medical Research.* 1999;8(1):3-15.

6. Royston P. Multiple imputation of missing values. *Stata J.* 2004;4(3):227-241.

7. Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol.* 2006;59(10):1087-1091.

8. Moore L, Lavoie A, LeSage N, et al. Multiple imputation of the Glasgow Coma Score. *J Trauma.* 2005;59(3):698-704.

9. Oostenbrink JB, Al MJ, Rutten-van Molken MPMH. Methods to analyse cost data of patients who withdraw in a clinical trial setting. *Pharmacoeconomics.* 2003;21(15):1103-1112.

10. Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC Med Res Methodol.* 2006;6:57.

11. van der Heijden GJMG, Donders ART, Stijnen T, Moons KGM. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. *J Clin Epidemiol.* 2006;59(10):1102-1109.

12. Barzi F, Woodward M. Imputations of missing values in practice: results from imputations of serum cholesterol in 28 cohort studies. *Am J Epidemiol.* 2004;160(1):34-45.

13. Arnold AM, Kronmal RA. Multiple imputation of baseline data in the cardiovascular health study. *Am J Epidemiol.* 2003;157(1):74-84.

14. Catellier DJ, Hannan PJ, Murray DM, et al. Imputation of missing data when measuring physical activity by accelerometry. *Med Sci Sports Exerc.* 2005;37(11 Suppl):S555-562.

15. Boyd N, Martin L, Stone J, Little L, Minkin S, Yaffe M. A longitudinal study of the effects of menopause on mammographic features. *Cancer Epidemiol Biomarkers Prev.* 2002;11(10 Pt 1):1048-1053.

16. Greendale GA, Reboussin BA, Sie A, et al. Effects of estrogen and estrogen-progestin on mammographic parenchymal density. Postmenopausal Estrogen/Progestin Interventions (PEPI) Investigators. *Ann Intern Med.* 1999;130(4 Pt 1):262-269.

17. Greendale GA, Reboussin BA, Slone S, Wasilauskas C, Pike MC, Ursin G. Postmenopausal hormone therapy and change in mammographic density. *J Natl Cancer Inst.* 2003;95(1):30-37.

18. McTiernan A, Martin CF, Peck JD, et al. Estrogen-plus-progestin use and mammographic density in postmenopausal women: women's health initiative randomized trial. *J Natl Cancer Inst.* 2005;97(18):1366-1376.

19. Brisson J, Sadowsky NL, Twaddle JA, Morrison AS, Cole P, Merletti F. The relation of mammographic features of the breast to breast cancer risk factors. *Am J Epidemiol.* 1982;115(3):438-443.

20. Grove JS, Goodman MJ, Gilbert FI, Jr., Mi MP. Factors associated with mammographic pattern. *Br J Radiol.* 1985;58(685):21-25.

21. Bergkvist L, Tabar L, Bergstrom R, Adami HO. Epidemiologic determinants of the mammographic parenchymal pattern. A population-based study within a mammographic screening program. *Am J Epidemiol.* 1987;126(6):1075-1081.

22. Hart BL, Steinbock RT, Mettler FA, Jr., Pathak DR, Bartow SA. Age and race related changes in mammographic parenchymal patterns. *Cancer.* 1989;63(12):2537-2539.

23. Albridge KM, Standish J, Fries JF. Hierarchical time-oriented approaches to missing data inference. *Comput Biomed Res.* 1988;21(4):349-366.

Table 4.1 Characteristics of study population at SWAN enrollment

Characteristic	Total study population N=834
Age, years; mean (SD)	46.5 (2.7)
Race/site; N (%)	
African American/Pittsburgh	62 (7.4)
Caucasian	407 (48.8)
Los Angeles	138 (16.5)
Oakland	133 (15.9)
Pittsburgh	136 (16.3)
Chinese/Oakland	182 (21.8)
Japanese/Los Angeles	183 (21.9)
Menopausal status; N (%)	
Premenopausal	483 (58.3)
Early perimenopausal	346 (41.7)
Body mass index; N (%)	
Underweight: <18.5 kg/m^2	16 (1.9)
Normal: 18.5 – <25 kg/m^2	474 (57.5)
Overweight: 25 – <30 kg/m^2	202 (24.5)
Obese: ≥30 kg/m^2	132 (16.0)

Table 4.2 Parameter estimates for the regression of body mass index on the square root transformed dense area by estimation approach, N=834[*]

Approach to Estimating Dense Breast Area	N	β	Variance	P value
1) Simple matching of mammograms to visits	824	-0.0155	5.43×10^{-5}	0.04
2) Linear interpolation (LI)	823[†]	-0.0070	2.73×10^{-5}	0.18
3) LI with single imputation of noise terms	823[†]	-0.0091	5.89×10^{-5}	0.24
4) LI with multiple imputation of noise terms	823[†]	-0.0098	11.86×10^{-5}	0.38
5) LI with multiple imputation for observations with >90 days between mammogram and visit and matching otherwise	824	-0.0105	11.73×10^{-5}	0.34

[*]Density was modeled using a square root transformation due to non-normality; models were adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit; observations missing in these variables were excluded

[†]One observation could not be interpolated due to lack of mammogram data both before and after the study visit, but was included in Approach 5 as the mammogram was within 90 days of the nearest visit

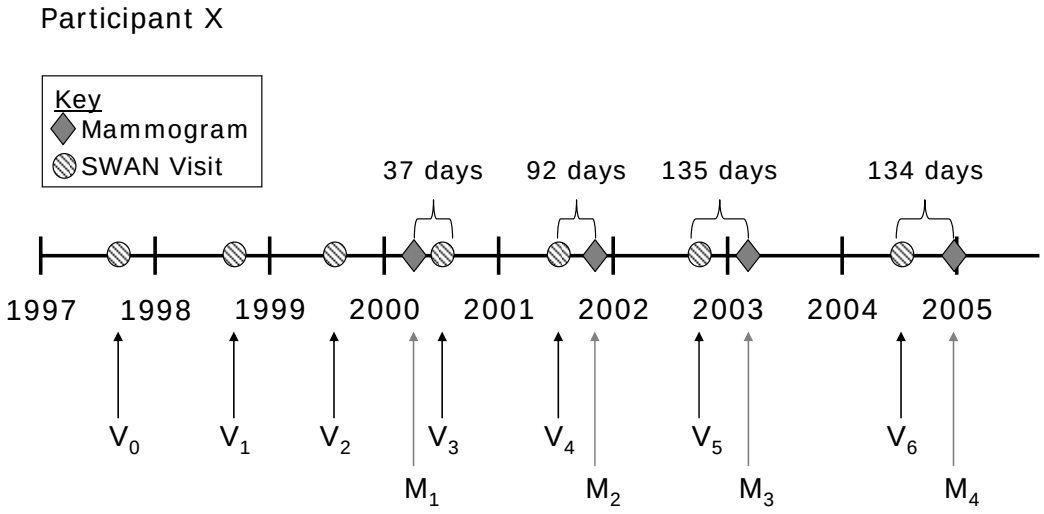


Figure 4.1 Observed data from SWAN visits and retrieved mammograms on a representative participant*

*Note: $V_0 - V_6$: Baseline to 6th annual SWAN visit; $M_1 - M_4$: 1st to 4th mammogram; Exact dates of visits and mammograms are not shown in order to preserve confidentiality

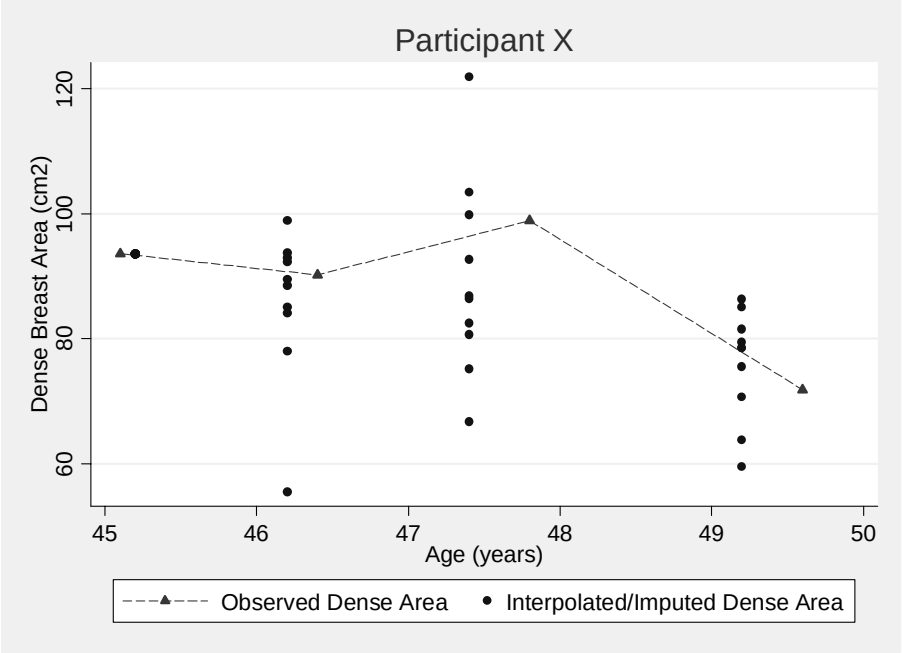


Figure 4.2 Observed and estimated dense area measurements using Approach 5 on a representative participant

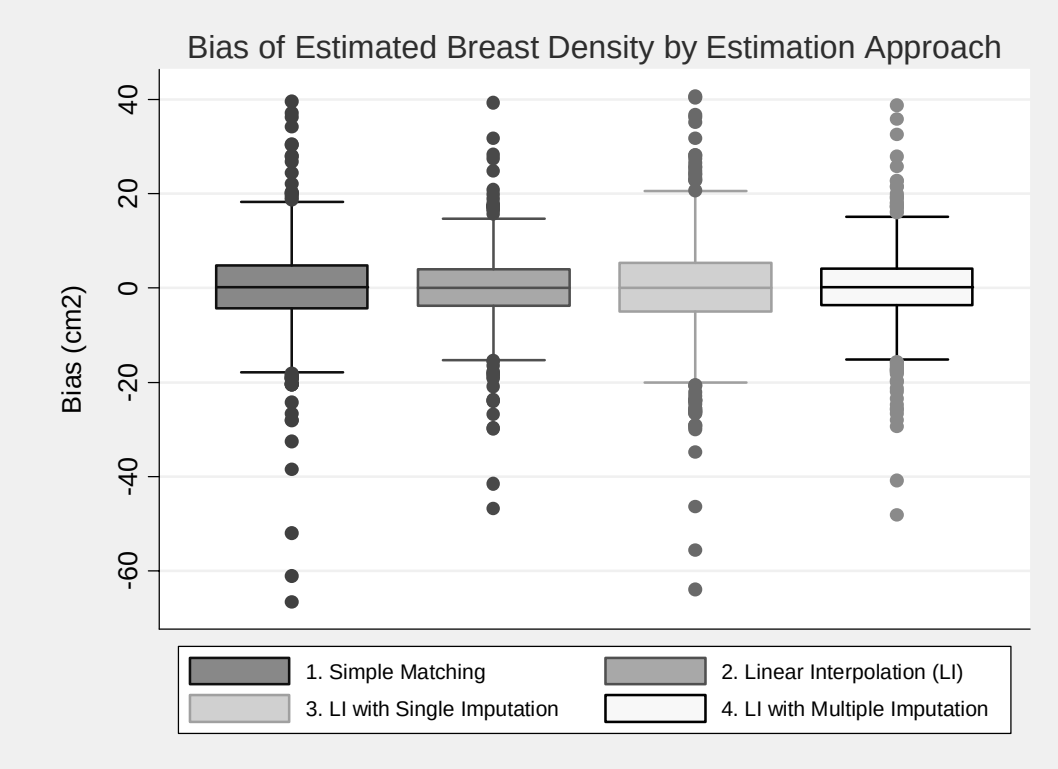


Figure 4.3 Boxplots of bias of estimated breast density in validation study by each estimation approach, N=710

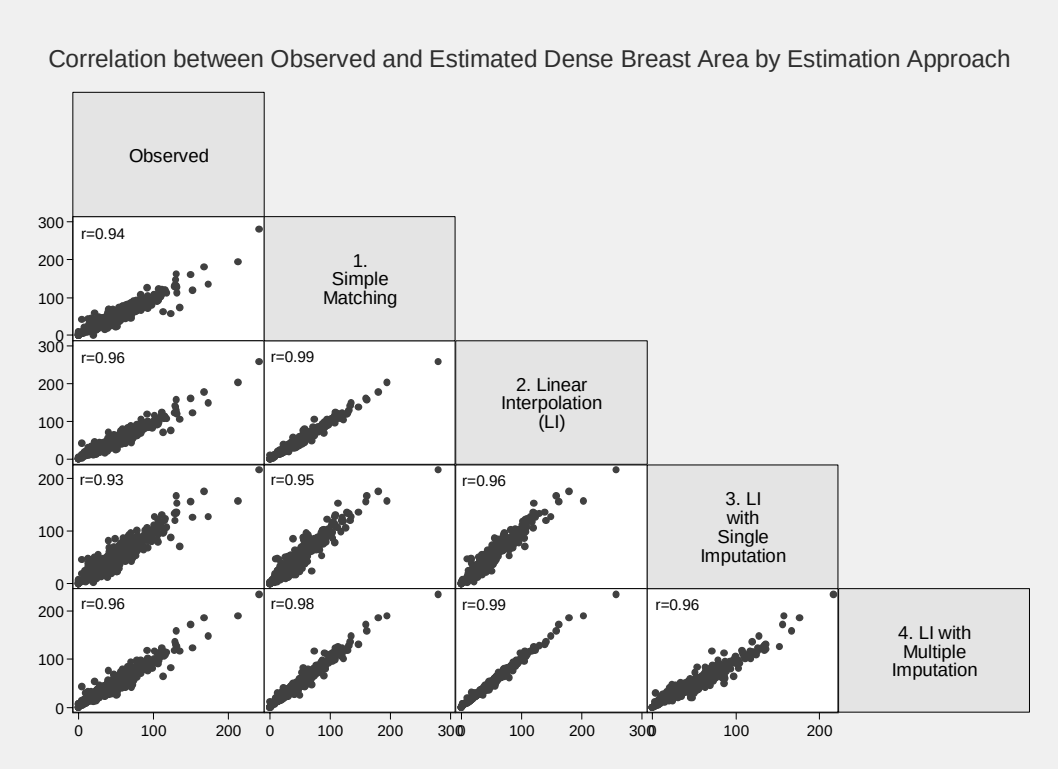


Figure 4.4 Correlation between dense breast area measurements estimated under each approach with observed data and other approaches used in the validation study, N=710[*]

[*]For imputed data, the average dense breast area across 10 imputations is plotted

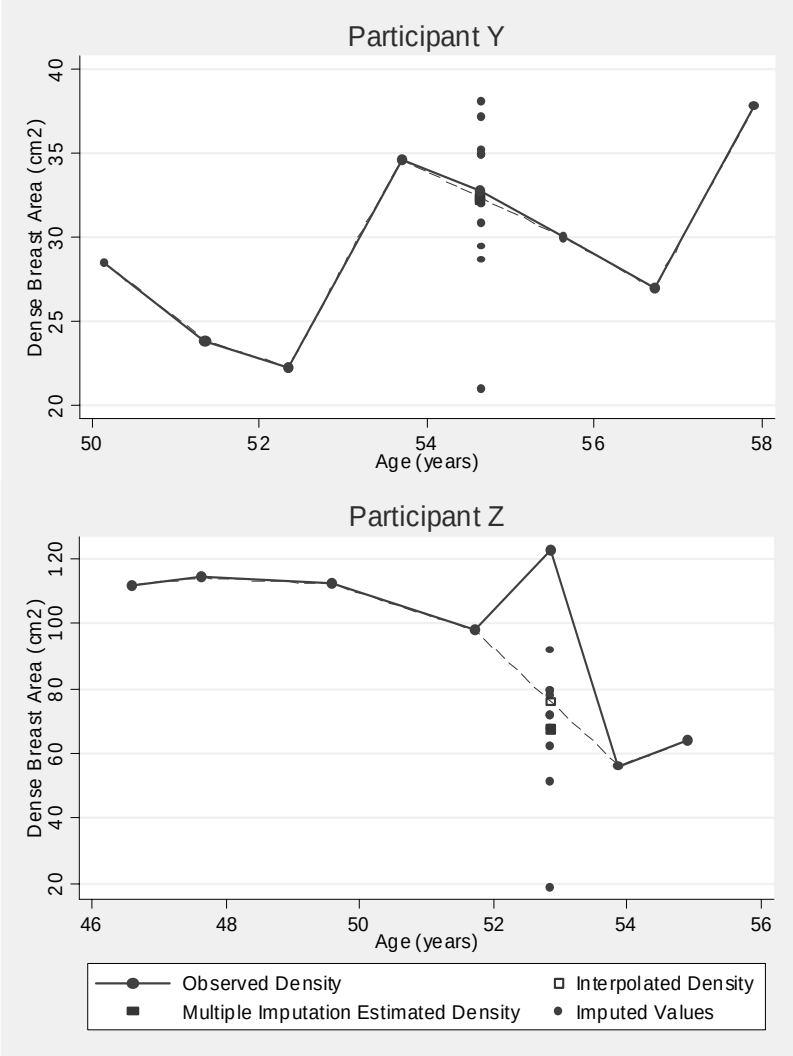


Figure 4.5 Examples of cases where the linear interpolation with multiple imputation (Approach 4) performed well (Participant Y, top) and where the method resulted in a large bias (Participant Z, bottom) in validation study

5.0 ARTICLE 3: VASCULAR ENDOTHELIAL GROWTH FACTOR AND BREAST CANCER RISK

To be submitted for publication

Katherine W. Reeves,[1,*] Roberta B. Ness,[1] Roslyn A. Stone,[2] Joel L. Weissfeld,[1,3]

Victor G. Vogel,[4] Robert Powers,[5,6] Francesmary Modugno,[1] Jane A. Cauley,[1]

[1]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[2]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[3]University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania

[4]Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

[5]Department of Obstetrics & Gynecology and Reproductive Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

[6]Magee-Womens Research Institute, Pittsburgh, Pennsylvania

This work was supported in part by National Institutes of Health grants P20 CA103730, R25-CA57703, K07-CA80668 and R21-CA95113; by Department of Defense grant DAMD17-02-1-0553 and by PA Department of Health grant P2777693. Additional support was provided by funds received from the NIH/NCRR/GCRC Grant MO1-RR000056 and by a small grant from the Department of Epidemiology, University of Pittsburgh.

5.1 ABSTRACT

Vascular endothelial growth factor (VEGF) is a key factor in angiogenesis and thereby plays an important role in carcinogenesis. Previous studies relating circulating levels of VEGF to breast cancer have been limited by small numbers of participants and lack of adjustment for important confounders. We studied the association between serum VEGF and breast cancer in an unmatched case-control study of 407 pre- and postmenopausal women (N=203 cases, N=204 controls). Breast cancer was confirmed through surgical and pathology reports. Controls were selected from women with negative findings on screening mammograms. Logistic regression models of natural log transformed VEGF and breast cancer were adjusted for age, Gail score, education, physical activity, history of breastfeeding, serum testosterone, and hormone therapy use. The majority of the population was postmenopausal (67.6%) and the average age was 56 years; age and menopausal status were similar among cases and controls. Geometric mean VEGF levels were higher in cases (321.4 pg/mL) than controls (291.4 pg/mL), albeit not significantly (p=0.21). In a multivariable model the odds of breast cancer were 37% higher for women with VEGF levels ≥314.2 pg/mL compared to those with levels below 314.2 pg/mL, although this association was not statistically significant (p=0.16). Results were similar in separate regressions of pre- and postmenopausal women. In this case-control study VEGF was non-significantly associated with increased breast cancer risk in pre- and postmenopausal women.

5.2 INTRODUCTION

An estimated 178,480 women in the United States will be diagnosed with breast cancer in 2007.[1] While there are many recognized breast cancer risk factors, including age, family history, and nulliparity, it remains difficult to predict which women will develop the disease. The Gail model is a statistical model often used to predict a woman's 5-year and lifetime risk of breast cancer.[2-4] Application of the Gail model to the Nurse's Health Study cohort, however, demonstrated that the ability of this model to discriminate between women who did and did not develop breast cancer was fairly low (concordance statistic 0.58, 95% CI 0.56-0.60).[5] Little is known about biological factors that may increase breast cancer risk. Biological risk factors that could be measured in blood or urine and used to refine current risk prediction models may enhance our ability to identify women likely to develop breast cancer.[5]

Angiogenesis is the process by which the body forms new blood vessels. Tumor growth is dependent on angiogenesis.[6, 7] Without vascularization tumors are unable to grow beyond 1-2 mm^3 in size.[8] While angiogenesis is a tightly controlled biological process involving multiple factors, vascular endothelial growth factor (VEGF) has been identified as a primary promoter of angiogenesis.[9] VEGF is a potent endothelial cell mitogen and does not act on other types of cells.[9] VEGF has a wide range of functions, including promoting endothelial cell mitogenesis and survival, increasing stromal degradation by promoting the expression of enzymes involved in this process, and promoting vascular permeability.[10] VEGF occurs in six different isoforms due to alternative splicing of the VEGF gene,[11, 12] with the 121 and 165 isoforms secreted as soluble proteins being the most abundant.[11, 12]

Despite strong evidence of an association between VEGF and breast cancer from *in vitro* and animal studies, studies in humans have not shown definitively that VEGF levels are related

to breast cancer risk. Numerous studies have reported that plasma or serum VEGF levels are increased among women with breast cancer compared to those without.[13-25] These studies, however, are limited by small sample sizes, incomplete description of the study populations, and/or failure to control for potential confounders in the analyses. Thus, the true association between VEGF levels and breast cancer remains unclear. Further, estrogen is known to be important to breast cancer development, and there is *in vitro* evidence that estrogen may upregulate VEGF mRNA and protein expression.[26-31] Few studies have investigated associations between VEGF and estrogen in humans, however.

We conducted an unmatched case-control study of serum VEGF levels in relation to breast cancer in an ancillary study to the Mammograms and Masses Study (MAMS). Our primary hypothesis was that serum VEGF levels would be positively associated with breast cancer in this population of pre- and postmenopausal women. We also investigated associations between VEGF and direct and indirect measures of estrogen exposure.

5.3 METHODS

5.3.1 Study population

MAMS is an unmatched case-control study of hormonal determinants of mammographic breast density.[32] Briefly, women were eligible for MAMS if they were age ≥18 and were receiving a) a breast biopsy, b) an initial surgical consultation after breast cancer diagnosis, or c) a routine screening mammogram. Exclusion criteria were prior cancer history other than non-melanoma skin cancer, alcohol intake >5 alcoholic beverages per day, or weight <110 pounds or >300

pounds. Women were enrolled from 2001-2005 through mammography and surgical clinics operated by Magee-Womens Hospital, Pittsburgh, PA. Pathology reports were obtained for women recruited from biopsy and surgical clinics to determine their disease status: benign breast disease, in situ disease, or invasive cancer. Controls were recruited from women with negative findings on screening mammograms. Of the eligible respondents, 55% of cases and 55% of controls enrolled in MAMS. The MAMS study population consists of 1,133 women, including 264 cases with in situ or invasive breast cancer, 313 women with benign breast disease, and 556 controls.

A subset of MAMS participants was selected for this investigation of VEGF and breast cancer. We included only breast cancer cases and healthy controls; thus participants with benign breast disease were excluded (N=313). Cases and controls were excluded from this analysis if they had no available mammogram (38 cases, 36 controls), were missing questionnaires (13 cases, 7 controls), had no available serum sample (5 cases, 5 controls), or if their blood draw was >14 days from enrollment (3 cases, 0 controls). These exclusions resulted in 205 cases (66 premenopausal and 139 postmenopausal) and 508 controls (105 premenopausal and 403 postmenopausal) eligible for the VEGF analysis. We included all 205 eligible cases. A simple random sample of 66 premenopausal and 139 postmenopausal controls was selected from the 508 eligible controls for this analysis. After completion of the VEGF analyses, three participants (1 control and 2 cases) were discovered to have a previous history of cancer and were excluded. The final population for analysis was 407, including 203 cases and 204 controls.

The Institutional Review Board at the University of Pittsburgh approved this study, and all participants provided written informed consent.

5.3.2 Data collection

Participants completed a self-administered take-home questionnaire upon enrollment into MAMS. Data collected included demographic characteristics as well as current and lifetime data on medical conditions and procedures, medications including hormone therapy (HT) and oral contraceptives (OC), reproductive events, family cancer history, weight, physical activity, smoking, and alcohol use. At enrollment, a research nurse measured participants' height and weight using a stadiometer and a standard balance beam scale while participants wore light clothing and no shoes.

Participants gave a non-fasting, 40 mL sample of peripheral blood at enrollment. The blood draw was not timed with the menstrual cycle for premenopausal women. Premenopausal women reported the date of their last menstrual period and the expected date of their next menstrual period, and menstrual cycle phase was inferred from this information. All blood samples were processed immediately at the Magee-Womens Hospital Satellite Clinical Research Center and stored at <-70°C.

5.3.3 Laboratory assays

Samples were relabeled with dummy identifiers and randomly distributed throughout the boxes transferred to the laboratories. A random sample of 40 masked duplicates (including 10 each from premenopausal cases and controls and postmenopausal cases and controls) were randomly distributed throughout the boxes. Samples were transferred packed in dry ice. All laboratory staff were masked to the identity, disease status, and demographic and risk factor characteristics of the samples.

VEGF was measured in serum by enzyme-linked immunosorbant assay (Quantikine® Human VEGF Immunoassay, R&D Systems, Minneapolis, MN).[33] This assay is specific for the 165 isoform of VEGF-A and has a minimum detectable concentration of <5 pg/mL. The coefficient of variation (CV) for the VEGF assay was 14.2%.

Estradiol (E2), follicle stimulating hormone (FSH), sex hormone binding globulin (SHBG), and testosterone (T) were measured in serum at the Clinical Ligand Assay Service Satellite (CLASS) Laboratory at the University of Michigan, School of Public Health. E2 was measured with a modified, off-line ACS:180 (E2-6) immunoassay (Bayer Diagnostics Corp, Tarrytown, NY).[34] This assay has a detectable range of 1 – 250 pg/mL. FSH was measured with a two-site chemiluminescence (sandwich) immunoassay.[35, 36] This assay measures FSH concentrations from 0.3 – 200 mIU/mL. SHBG was measured using a competitive immunoassay run on Bayer Diagnostic's ACS:180 automated analyzer using chemiluminescent technology.[35] The detectable range for SHBG is 1.95 to 250 nM. Total T was measured using a modification of the ACS:180 total T assay to measure with greater precision samples in the low ranges found in women in the peri- and postmenopause.[37] The limit of detection of this assay is <5.15 ng/dL. CVs were 42.3% for E2, 5.5% for FSH, 14.6% for SHBG, and 13.6% for T. The high CV for E2 was related to the low concentrations of E2 observed in postmenopausal women; categorization into quartiles based on the distribution in controls provided better reliability and therefore was used in analyses.

5.3.4 Statistical analysis

Descriptive statistics were calculated for demographic and behavioral characteristics and the laboratory measures. Biological measures below the detection limit for the assay were reset to

the stated detection limit. A natural log transformation was applied to all biological measures to improve normality. The distributions of the demographic, behavioral, and biological variables were compared by disease status using two-sample t tests for continuous variables and chi-square tests for categorical variables.

Variables associated with breast cancer or VEGF in previous studies were evaluated for their association with VEGF: age (continuous), race (white, other), educational level (high school, >high school), BMI (normal: <25 kg/m^2, overweight: 25–<30 kg/m^2, obese: ≥30 kg/m^2), alcohol intake in year prior to enrollment (none, <12 g/d, ≥12 g/d), current alcohol use (no, yes), smoking status (never, former, current), cumulative physical activity in metabolic equivalent (MET)-h/wk (0, 0.1 – 10, ≥10.1), age at menarche (<13, ≥13), menstrual cycle regularity (no, yes, sometimes), menopausal status (premenopausal, postmenopausal without hysterectomy, postmenopausal with hysterectomy), age at menopause (premenopausal, <50, ≥50), menstrual cycle phase (luteal, follicular, unknown, postmenopausal), ever pregnant (no, yes), number of live births (none, 1, 2, ≥3), age at first pregnancy lasting >6 months (no live births/pregnancy lasted <6 months, <20, 20–24, 25–29, ≥30), breastfeeding history (not applicable, no, yes), breast cancer family history (no, yes), cancer family history (no, yes), Gail score (<1.66%, ≥1.66%), previous breast biopsy (no, yes), diabetes (no, yes), myocardial infarction history (no, yes), heart disease history (no, yes), HT use (never, former, current), OC use (never, former, current), E2 (quartiles), FSH (quartiles), SHBG (quartiles), and T (quartiles). Categorizations of these variables were based on common cutpoints (e.g. BMI) or on the original response categories with collapsing of categories to prevent small cell counts (e.g. age at menarche). Quartiles for E2, FSH, SHBG, and T were based on the distribution of these hormones among controls. VEGF was dichotomized based on the median level among the controls (<314.2 pg/mL, ≥314.2 pg/mL).

Bivariate associations between VEGF and these variables among control participants were assessed using chi-square tests. Fisher's exact test was used in instances where the expected cell count was <5.

Logistic regression was used to evaluate the association between VEGF and breast cancer adjusting for relevant covariates. VEGF was modeled in separate regressions as a continuous variable and as a dichotomous variable. A natural log transformation was applied to VEGF to improve normality when VEGF was modeled as a continuous variable. The aforementioned variables were evaluated for inclusion as potential confounders using backward selection based on Wald tests. Dummy variables were created for categorical variables as appropriate. All covariates with a p value <0.10 from a likelihood ratio test were retained in the model. Fractional polynomials were used to assess the assumption that continuous variables were linear in the logit. The Hosmer-Lemeshow test was used to assess model goodness of fit. Potentially influential observations were identified as those having significant influence on model deviance (as assessed by Hosmer-Lemeshow's delta deviance test) or parameter estimates (as assessed by Pregibon's delta beta test). Likelihood ratio tests were used to evaluate the significance of hypothesized interactions by comparing the model including the interaction term to the main effects model.

All analyses were repeated among subgroups defined by menopausal status. Stata version 10.0 was used for all analyses (Stata Corportation, College Station, TX). Two-sided p values ≤ 0.05 were considered statistically significant, with no adjustment for multiple comparisons.

5.4 RESULTS

The 407 participants comprising the study population are described in Table 5.1. The mean age of both cases and controls was 56 (p=0.99). The vast majority of the participants were non-Hispanic whites, though participants of other ethnicities were somewhat more common among controls than cases (7.4% versus 2.5%, p=0.02). BMI was similar between cases and controls, with a mean BMI of 27.8 kg/m^2 among cases and 27.9 kg/m^2 among controls (p=0.84). Compared to controls, cases exercised less (p<0.001), were less likely ever to have breastfed a child (p=0.01), had a higher mean Gail score (p=0.02), and were more likely to be current users of HT (p<0.001). Of the 203 cases, 52 (25.6%) had in situ disease and 151 (74.4%) had invasive cancer.

The distributions of VEGF and the measured reproductive hormones are summarized in Table 5.2. The geometric mean of serum VEGF among cases, 321.4 pg/mL, was higher than that among controls (291.4 pg/mL, p=0.21), albeit not significantly so. Similar results were obtained within the subgroups of pre- and postmenopausal women (data not shown). No significant differences were observed between cases and controls for the geometric means of E2 (15.5 versus 12.6 pg/mL, p=0.18), FSH (62.9 mIU/mL versus 55.5 mIU/mL, p=0.31), and SHBG (51.6 nM versus 49.3 nM, p=0.45). Geometric mean T levels were higher among cases (32.8 ng/dL) compared to controls (28.1 ng/dL; p=0.01). When restricted to premenopausal women only, geometric mean FSH levels were significantly higher among cases (20.7 mIU/mL) than controls (13.4 mIU/mL; p=0.05). Among postmenopausal women only, E2 levels were significantly higher among cases than controls (geometric mean 9.5 pg/mL versus 6.8 pg/mL, respectively, p=0.02; data not shown).

None of the evaluated personal characteristics were significantly associated with serum VEGF levels among controls (Table 5.3). Characteristics related to breast cancer risk, such as Gail score, physical activity, age at menopause, parity, and HT use, were not different among women with VEGF levels at or above the median level of controls (314.2 pg/mL) compared to those below. We observed a positive association between FSH and VEGF (p=0.04). Similar results were observed in analyses restricted to premenopausal or postmenopausal women (data not shown).

In unadjusted and age-adjusted logistic regression analyses, associations between serum VEGF and breast cancer were not significant (Table 5.4). In a model adjusted for age, Gail score, education, physical activity, history of breastfeeding, serum T, and HT use, VEGF (modeled as a continuous variable) was non-significantly positively associated with breast cancer (OR 1.21 per 1 unit increase in ln(VEGF), 95% CI 0.91 – 1.59). Figure 5.1 illustrates this association by displaying odds ratios for specific levels of VEGF calculated from the continuous logistic regression model. For example, the odds of breast cancer were increased by 16% for a woman with a serum VEGF level of 691.4 pg/mL compared to a similar woman with a VEGF level of 291.4 pg/mL. In a similar multivariable regression model with VEGF as a dichotomous variable, women with VEGF levels ≥314.2 pg/mL had a non-significant 37% increase in the odds of having breast cancer (OR 1.37, 95% CI 0.88 – 2.12), compared to women with VEGF <314.2 pg/mL (Figure 5.1).

The magnitude of the association between serum VEGF, dichotomized at the control group median, and breast cancer was similar among premenopausal (OR 1.40, 95% CI 0.64 – 3.07) and postmenopausal women (OR 1.28, 95% CI 0.74 – 2.22). There was no indication of an interaction between VEGF and menopausal status (p=0.52). Additionally, no significant

interactions were observed between VEGF dichotomized at the control median and Gail score (p=0.41), E2 (p=0.62), or T (p=0.88).

We also performed exploratory analyses to investigate whether HT use affected the association between VEGF and breast cancer (Table 5.4). The logistic regressions were repeated within subgroups defined by HT use (never, past, and current users). In multivariable models with VEGF modeled as a dichotomous variable, no association between VEGF and breast cancer was observed among never HT users (OR 1.23, 95% CI 0.68 – 2.22) or current HT users (OR 1.09, 95% CI 0.31 – 3.87). VEGF was positively, though non-significantly, associated with breast cancer among past HT users (OR 2.28, 95% CI 0.86 – 5.68). A likelihood ratio test of an interaction term between VEGF and HT use status was not statistically significant (p=0.45) in a multivariable model with VEGF as a dichotomous variable.

Results were similar in a regression including cases with invasive cancer only (OR 1.41, 95% CI 0.88 – 2.27; p=0.16). Using the methods described earlier, we identified four potentially influential observations (3 controls and 1 case) in the multivariable regression with VEGF as a dichotomous variable. In a sensitivity analysis excluding these four observations, the increase in breast cancer risk for women with VEGF greater than the median was similar to that observed in the full population (OR 1.45, 95% CI 0.92 – 2.27, p=0.11; data not shown).

5.5 DISCUSSION

In this case-control study of pre- and postmenopausal women, we found that serum VEGF levels were positively associated with breast cancer, although the association was not statistically significant. A woman with a serum VEGF level greater than the control median of 314.2 pg/mL,

for example, had 37% higher odds of breast cancer compared to an otherwise similar woman with a serum VEGF level below 314.2 pg/mL. Results were similar among subgroups of pre- and postmenopausal women.

The role of VEGF in angiogenesis and carcinogenesis is complex and likely involves interaction with multiple pathways. Of particular interest to breast cancer development is the possibility that hormones may regulate VEGF. Numerous *in vitro* studies using the estrogen-sensitive MCF-7 breast cancer cell line have reported that estrogen increases VEGF mRNA and/or protein expression by these cells,[26-31] although one study found no effect[38] and another reported decreased VEGF expression induced by estrogen.[39] VEGF expression can be decreased by the selective estrogen receptor modulator tamoxifen,[27, 28, 40] though some studies report that tamoxifen increases VEGF expression in MCF-7 cells.[30, 39] Most studies exploring the effects of progestins in the progestin-sensitive T47-D breast cancer cell line report that progestin exposure increases VEGF expression,[38, 41, 42] though at least one study found no effect.[43] Both natural and synthetic progestins are able to increase VEGF expression *in vitro*, and the synthetic progestin medroxyprogesterone acetate (MPA) is reported to have the strongest effect on VEGF expression by T47-D cells.[38, 41] MPA is commonly used in postmenopausal HT, and it has been suggested that the strong effect of MPA on VEGF expression may at least partially explain the increased breast cancer risk observed in women using combination estrogen and progestin preparations versus those using estrogen alone.[41]

Few studies in humans have investigated associations between hormones and VEGF in relation to breast cancer risk. A study of 16 healthy, pre- and postmenopausal women reported high correlations between VEGF in breast tissue and E2 in both plasma ($r=0.81$, $p<0.0001$) and breast tissue ($r=0.67$, $p=0.004$).[44] We observed no significant association between serum VEGF

and serum E2 among healthy controls, though women with VEGF levels ≥314.2 pg/mL were more likely to be in the lowest quartile of E2 level. Differences in the study populations and in the medium in which VEGF and E2 were measured preclude a direct comparison between these studies and may explain the divergent results. We also observed a stronger, positive association between VEGF and breast cancer among past HT users. This association was not apparent among current HT users, though the number of current HT users was small. Though such differences in the associations between VEGF and breast cancer by HT use are intriguing, a formal test of this interaction was not significant. These exploratory findings provide limited evidence that steroid hormones have effects on VEGF.

Our results are consistent with those of previous studies that have reported positive associations between serum or plasma VEGF and breast cancer,[13-25, 45] though our results did not achieve statistical significance. These previous studies, however, only performed comparisons of mean or median VEGF levels between cases and controls, and did not examine associations between VEGF and breast cancer in a multivariable context. In our study, VEGF levels above the median were more common in cases compared to controls after adjusting for relevant covariates, although this association was not statistically significant (p=0.16).

Previous studies all included fewer than 100 healthy female controls, and most failed to describe how such controls were selected.[13, 14, 18-20, 24] Control participants in MAMS were recruited from healthy women seeking a screening mammogram. Our study population consisted of pre- and postmenopausal women ranging in age from 35 to 84. Many previous studies have failed to adequately describe the age[13, 16, 20, 22-25] or menopausal status[13-17, 19-21, 23-25, 45, 46] of their study populations, making a direct comparison of results difficult. The distribution of VEGF among the controls in our population differed substantially from those reported by

other studies. For example, the arithmetic mean serum VEGF among controls was 387.5 pg/mL in our population, but ranged from 77.0 pg/mL to 230.0 pg/mL in other studies.[13, 14, 23, 24] Similarly, the median serum VEGF among our controls was 314.2 pg/mL compared to medians ranging from 17.0 pg/mL to 186.0 pg/mL reported in other studies.[17-21] This may reflect the higher sensitivity of the assay used here, the influence of a selection bias within our study, or may indicate that serum VEGF levels among healthy women are higher and more variable than previously thought.

A significant limitation to our study is the low statistical power. Although we determined *a priori* that we would have greater than 80% power to detect a difference of 60 pg/mL in mean serum VEGF between cases and controls as statistically significant, the variability of VEGF in our study population far exceeded that observed in the study used as a basis for power calculations.[13] In fact, *a posteriori* power calculations show that our study had only 24% power to detect a difference between the mean values of serum VEGF we observed among cases and controls. Thus our ability to detect true differences in VEGF between cases and controls was limited.

The case-control design prevents us from making a temporal inference. However, case-control studies are an important step towards recommending prospective studies. VEGF levels were assessed at a single point in time. Thus the levels may not be representative of a participant's usual levels but may reflect recent changes in general health or medication use. The extensive data on such factors collected in MAMS allowed for statistical control of these variables and minimized the impact of confounding on the observed results. Only a small percentage of MAMS participants were non-White, thus these results may not apply to women of other races or ethnicities. Further, MAMS participants were better educated and less likely to

smoke as compared to women from Allegheny County overall,[32, 47] indicating a possible volunteer bias, a concern reflected in the modest enrollment levels.

Our results add to the existing body of literature reporting that circulating levels of VEGF are positively associated with breast cancer. Future studies should investigate whether VEGF levels measured prospectively are indicative of later risk of breast cancer. Also, further work is needed to evaluate the relationships between endogenous and exogenous hormones and VEGF and how such relationships might impact breast cancer risk.

5.6 LITERATURE CITED

1. American Cancer Society. *Cancer Facts & Figures 2007.* Atlanta: American Cancer Society, Inc.; 2007.

2. Gail MH, Costantino JP, Bryant J, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst.* 1999;91(21):1829-1846.

3. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst.* 1999;91(18):1541-1548.

4. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81(24):1879-1886.

5. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst.* 2001;93(5):358-366.

6. Folkman J, Klagsbrun M. Angiogenic factors. *Science.* 1987;235(4787):442-447.

7. Folkman J. Tumor angiogenesis: therapeutic implications. *N Engl J Med.* 1971;285(21):1182-1186.

8. Folkman J. What is the evidence that tumors are angiogenesis dependent? *J Natl Cancer Inst.* 1990;82(1):4-6.

9. Kushlinskii NE, Gershtein ES. Role of vascular endothelial growth factor during breast cancer. *Bull Exp Biol Med.* 2002;133(6):521-528.

10. Sledge GW, Jr. Vascular endothelial growth factor in breast cancer: biologic and therapeutic aspects. *Semin Oncol.* 2002;29(3 Suppl 11):104-110.

11. Ferrara N. The role of vascular endothelial growth factor in pathological angiogenesis. *Breast Cancer Res Treat.* 1995;36(2):127-137.

12. Stimpfl M, Tong D, Fasching B, et al. Vascular endothelial growth factor splice variants and their prognostic value in breast and ovarian cancer. *Clin Cancer Res.* 2002;8(7):2253-2259.

13. Benoy I, Salgado R, Colpaert C, Weytjens R, Vermeulen PB, Dirix LY. Serum interleukin 6, plasma VEGF, serum VEGF, and VEGF platelet load in breast cancer patients. *Clin Breast Cancer.* 2002;2(4):311-315.

14. Coskun U, Gunel N, Sancak B, et al. Significance of serum vascular endothelial growth factor, insulin-like growth factor-I levels and nitric oxide activity in breast cancer patients. *Breast.* 2003;12(2):104-110.

15. Hoar FJ, Lip GYH, Belgore F, Stonelake PS. Circulating levels of VEGF-A, VEGF-D and soluble VEGF-A receptor (sFIt-1) in human breast cancer. *Int J Biol Markers.* 2004;19(3):229-235.

16. Dirix LY, Vermeulen PB, Pawinski A, et al. Elevated levels of the angiogenic cytokines basic fibroblast growth factor and vascular endothelial growth factor in sera of cancer patients. *Br J Canc.* 1997;76(2):238-243.

17. Salven P, Perhoniemi V, Tykka H, Maenpaa H, Joensuu H. Serum VEGF levels in women with a benign breast tumor or breast cancer. *Breast Cancer Res Treat.* 1999;53(2):161-166.

18. Adams J, Carder PJ, Downey S, et al. Vascular endothelial growth factor (VEGF) in breast cancer: comparison of plasma, serum, and tissue VEGF and microvessel density and effects of tamoxifen. *Canc Res.* 2000;60(11):2898-2905.

19. Granato AM, Nanni O, Falcini F, et al. Basic fibroblast growth factor and vascular endothelial growth factor serum levels in breast cancer patients and healthy women: useful as diagnostic tools? *Breast Cancer Res.* 2004;6(1):R38-45.

20. Heer K, Kumar H, Read JR, Fox JN, Monson JR, Kerin MJ. Serum vascular endothelial growth factor in breast cancer: its relation with cancer type and estrogen receptor status. *Clin Cancer Res.* 2001;7(11):3491-3494.

21. Salven P, Manpaa H, Orpana A, Alitalo K, Joensuu H. Serum vascular endothelial growth factor is often elevated in disseminated cancer. *Clin Cancer Res.* 1997;3(5):647-651.

22. Wu Y, Saldana L, Chillar R, Vadgama JV. Plasma vascular endothelial growth factor is useful in assessing progression of breast cancer post surgery and during adjuvant treatment. *Int J Oncol.* 2002;20(3):509-516.

23. Yamamoto Y, Toi M, Kondo S, et al. Concentrations of vascular endothelial growth factor in the sera of normal controls and cancer patients. *Clin Cancer Res.* 1996;2(5):821-826.

24. Donovan D, Harmey JH, Toomey D, Osborne DH, Redmond HP, Bouchier-Hayes DJ. TGF beta-1 regulation of VEGF production by breast cancer cells. *Ann Surg Oncol.* 1997;4(8):621-627.

25. Zaman K, Driscoll R, Hahn D, et al. Monitoring multiple angiogenesis-related molecules in the blood of cancer patients shows a correlation between VEGF-A and MMP-9 levels before treatment and divergent changes after surgical vs. conservative therapy. *Int J Cancer.* 2006;118(3):755-764.

26. Buteau-Lozano H, Ancelin M, Lardeux B, Milanini J, Perrot-Applanat M. Transcriptional regulation of vascular endothelial growth factor by estradiol and tamoxifen in breast cancer cells: a complex interplay between estrogen receptors alpha and beta. *Canc Res.* 2002;62(17):4977-4984.

27. Garvin S, Nilsson UW, Dabrosin C. Effects of oestradiol and tamoxifen on VEGF, soluble VEGFR-1, and VEGFR-2 in breast cancer and endothelial cells. *Br J Canc.* 2005;93(9):1005-1010.

28. Lee JE, Chung K-W, Han W, et al. Effect of estrogen, tamoxifen and epidermal growth factor on the transcriptional regulation of vascular endothelial growth factor in breast cancer cells. *Anticancer Res.* 2004;24(6):3961-3964.

29. Maity A, Sall W, Koch CJ, Oprysko PR, Evans SM. Low pO2 and beta-estradiol induce VEGF in MCF-7 and MCF-7-5C cells: relationship to in vivo hypoxia. *Breast Cancer Res Treat.* 2001;67(1):51-60.

30. Ruohola JK, Valve EM, Karkkainen MJ, Joukov V, Alitalo K, Harkonen PL. Vascular endothelial growth factors are differentially regulated by steroid hormones and antiestrogens in breast cancer cells. *Mol Cell Endocrinol.* 1999;149(1-2):29-40.

31. Takei H, Lee E-S, Jordan VC. In vitro regulation of vascular endothelial growth factor by estrogens and antiestrogens in estrogen-receptor positive breast cancer. *Breast Cancer.* 2002;9(1):39-42.

32. Reeves KW, Gierach GL, Modugno F. Recreational physical activity and mammographic breast density characteristics. *Cancer Epidemiol Biomarkers Prev.* 2007;16(5):934-942.

33. Quantikine Human VEGF Immunoassay. Catalog number DVE00. http://www.rndsystems.com/pdf/dve00.pdf. Accessed December 27, 2006.

34. England BG, Parsons GH, Possley RM, McConnell DS, Midgley AR. Ultrasensitive semiautomated chemiluminescent immunoassay for estradiol. *Clin Chem.* 2002;48(9):1584-1586.

35. Randolph JF, Jr., Sowers M, Gold EB, et al. Reproductive hormones in the early menopausal transition: relationship to ethnicity, body size, and menopausal status. *J Clin Endocrinol Metab.* 2003;88(4):1516-1522.

36. Santoro N, Crawford SL, Allsworth JE, et al. Assessing menstrual cycles with urinary hormone assays. *Am J Physiol Endocrinol Metab.* 2003;284(3):E521-530.

37. Sowers MF, Beebe JL, McConnell D, Randolph J, Jannausch M. Testosterone concentrations in women aged 25-50 years: associations with lifestyle, body composition, and ovarian status. *Am J Epidemiol.* 2001;153(3):256-264.

38. Mirkin S, Wong BC, Archer DF. Effect of 17 beta-estradiol, progesterone, synthetic progestins, tibolone, and tibolone metabolites on vascular endothelial growth factor mRNA in breast cancer cells. *Fertil Steril.* 2005;84(2):485-491.

39. Bogin L, Degani H. Hormonal regulation of VEGF in orthotopic MCF7 human breast cancer. *Canc Res.* 2002;62(7):1948-1951.

40. Garvin S, Dabrosin C. Tamoxifen inhibits secretion of vascular endothelial growth factor in breast cancer in vivo. *Canc Res.* 2003;63(24):8742-8748.

41. Hyder SM, Chiappetta C, Stancel GM. Pharmacological and endogenous progestins induce vascular endothelial growth factor expression in human breast cancer cells. *Int J Cancer.* 2001;92(4):469-473.

42. Liang Y, Hyder SM. Proliferation of endothelial and tumor epithelial cells by progestin-induced vascular endothelial growth factor from human breast cancer cells: paracrine and autocrine effects. *Endocrinology.* 2005;146(8):3632-3641.

43. Scott PA, Gleadle JM, Bicknell R, Harris AL. Role of the hypoxia sensing system, acidity and reproductive hormones in the variability of vascular endothelial growth factor induction in human breast carcinoma cell lines. *Int J Cancer.* 1998;75(5):706-712.

44. Dabrosin C. Positive correlation between estradiol and vascular endothelial growth factor but not fibroblast growth factor-2 in normal human breast tissue in vivo. *Clin Cancer Res.* 2005;11(22):8036-8041.

45. Nishimura R, Nagao K, Miyayama H, et al. Higher plasma vascular endothelial growth factor levels correlate with menopause, overexpression of p53, and recurrence of breast cancer. *Breast Cancer.* 2003;10(2):120-128.

46. Teh SH, Hill ADK, Lee AWS, et al. Raised plasma endostatin levels correlate inversely with breast cancer angiogenesis. *J Surg Res.* 2004;116(1):165-171.

47. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Data: Selected Metropolitan/Micropolitan Area Risk Trends. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2005

Table 5.1 Descriptive characteristics of the study population by breast cancer status, N=407

Characteristic	Controls N (%)	Cases N (%)	P value*
Age, years; mean (SD)	56.5 (10.1)	56.5 (10.3)	0.99
<50	56 (27.5)	56 (27.6)	0.56
50-59	81 (39.7)	73 (36.0)	
60-69	40 (19.6)	51 (25.1)	
≥70	27 (13.2)	23 (11.3)	
Ethnicity			0.02
White	189 (92.7)	198 (97.5)	
Other	15 (7.4)	5 (2.5)	
Education level			0.001
High school	37 (18.1)	66 (32.5)	
Greater than high school	167 (81.7)	137 (67.5)	
Body mass index, kg/m^2; mean (SD)	27.9 (6.2)	27.8 (5.8)	0.84
Normal, <25 kg/m^2	76 (37.3)	76 (37.8)	0.99
Overweight, 25-<30 kg/m^2	66 (32.4)	64 (31.8)	
Obese, ≥30 kg/m^2	62 (30.4)	61 (30.4)	
Physical activity, MET h/wk			<0.001
0	20 (9.8)	54 (26.6)	
0.1 – <10	72 (35.3)	63 (31.0)	
≥10	112 (54.9)	86 (42.4)	
Age at menarche, years			0.99
<13	106 (52.2)	105 (52.2)	
≥13	97 (47.8)	96 (47.8)	
Menopausal status			0.09
Premenopausal	66 (32.4)	66 (32.5)	
Postmenopausal without hysterectomy	100 (49.0)	82 (40.4)	
Postmenopausal with hysterectomy	38 (18.6)	55 (27.1)	
Age at menopause, years			0.12
Premenopausal	66 (32.8)	66 (33.5)	
<50	55 (27.4)	70 (35.5)	
≥50	80 (39.8)	61 (31.0)	
Number of live births			0.36
None	48 (23.5)	41 (20.2)	
1	28 (13.7)	19 (9.4)	
2	68 (33.3)	73 (36.0)	
≥3	60 (29.4)	70 (34.5)	
History of breastfeeding			0.05
Not applicable	48 (23.5)	41 (20.3)	
No	65 (31.9)	88 (43.6)	
Yes	91 (44.6)	73 (36.1)	
Previous breast biopsy	25 (12.3)	51 (25.5)	0.001
First degree relative with breast cancer	28 (13.8)	39 (19.2)	0.14
Gail score; mean (SD)	1.49 (0.67)	1.71 (1.08)	0.02†
< 1.66%	140 (69.0)	126 (62.4)	0.16
≥ 1.66%	63 (31.0)	76 (37.6)	

Table 5.1 (continued)

Hormone therapy use status			<0.001
Never	115 (56.4)	108 (53.7)	
Former	69 (33.8)	45 (22.4)	
Current (within previous 3 months)	20 (9.8)	48 (23.9)	
Oral contraceptive use status			0.44
Never	66 (34.6)	64 (34.0)	
Former	118 (61.8)	121 (64.4)	
Current	7 (3.7)	3 (1.6)	

*P values from t tests for continuous variables and chi square tests for categorical variables
†From t test with unequal variances
Abbreviations used: SD, standard deviation; MET, metabolic equivalent

Table 5.2 Summary statistics of serum levels of VEGF, E2, FSH, SHBG, and T in the study population by breast cancer status, N=407[*]

	N	Mean	SD	Geometric Mean	Median	25th – 75th Percentiles
Controls						
VEGF, pg/mL	204	387.5	293.3	291.4	314.2	180.5 – 510.9
E2, pg/mL	204	44.5	82.8	12.6	11.2	4.1 – 40.9
FSH, mIU/mL	204	92.4	66.2	55.5	88.8	31.4 – 138.9
SHBG, nM	204	58.2	33.8	49.3	51.6	35.3 – 73.1
T, ng/dL	204	32.9	18.9	28.1	29.2	18.4 – 42.2
Cases						
VEGF, pg/mL	202[†]	415.7	287.8	321.4	341.7	190.8 – 579.4
E2, pg/mL	203	36.8	56.2	15.5	13.7	6.9 – 38.7
FSH, mIU/mL	203	97.3	66.7	62.9	101.0	30.7 – 138.0
SHBG, nM	203	62.9	45.0	51.6	49.2	34.4 – 73.4
T, ng/dL	203	38.0	22.0	32.8	34.1	24.6 – 44.9

[*]P values from t tests comparing cases to controls on natural log transformed values: VEGF, p=0.21; E2, p=0.18; FSH, p=0.31; SHBG, p=0.45; T, p=0.01
[†]VEGF could not be measured in one case due to insufficient sample volume
Abbreviations used: VEGF, vascular endothelial growth factor; E2, estradiol; FSH, follicle stimulating hormone; SHBG, sex hormone binding globulin; T, testosterone

Table 5.3 Bivariate associations between serum VEGF level and personal characteristics among controls, by VEGF level, N=204

Characteristic	VEGF <314.2 pg/mL N (%)	VEGF ≥314.2 pg/mL N (%)	P value[†]
Age, years			0.96
<50	29 (28.4)	27 (26.5)	
50-59	39 (38.2)	42 (41.2)	
60-69	21 (20.6)	19 (18.6)	
≥70	13 (12.8)	14 (13.8)	
Ethnicity			0.79
White	95 (93.1)	94 (92.2)	
Other	7 (6.9)	8 (7.8)	
Education level			0.20
High school	22 (21.6)	15 (14.7)	
Greater than high school	80 (78.4)	87 (85.3)	
Body mass index, kg/m^2			0.20
Normal, <25 kg/m^2	35 (34.3)	41 (40.2)	
Overweight, 25-<30 kg/m^2	39 (38.2)	27 (26.5)	
Obese, ≥30 kg/m^2	28 (27.5)	34 (33.3)	
Physical activity, MET h/wk			0.49
0	9 (8.8)	11 (10.8)	
0.1 – <10	40 (39.2)	32 (31.4)	
≥10	53 (52.0)	59 (57.8)	
Age at menarche, years			0.63
<13	55 (53.9)	51 (50.5)	
≥13	47 (46.1)	50 (49.5)	
Menopausal status			0.67
Premenopausal	36 (35.3)	30 (29.4)	
Postmenopausal without hysterectomy	48 (47.1)	52 (51.0)	
Postmenopausal with hysterectomy	18 (17.7)	20 (19.6)	
Age at menopause, years			0.49
Premenopausal	36 (36.0)	30 (29.7)	
<50	24 (24.0)	31 (30.7)	
≥50	40 (40.0)	40 (39.6)	
Number of live births			0.65
None	27 (26.5)	21 (20.6)	
1	12 (11.8)	16 (15.7)	
2	35 (34.3)	33 (32.4)	
≥3	28 (27.5)	32 (31.4)	
History of breastfeeding			0.19
Not applicable	27 (26.5)	21 (20.6)	
No	36 (35.3)	29 (28.4)	
Yes	39 (38.2)	52 (51.0)	
Previous breast biopsy			0.83
No	90 (88.2)	89 (87.3)	
Yes	12 (11.8)	13 (12.8)	
First degree relative with breast cancer			0.43
No	86 (84.3)	89 (88.1)	
Yes	16 (15.7)	12 (11.9)	

Table 5.3 (continued)

Gail score			0.42
<1.66%	73 (71.6)	67 (66.3)	
≥1.66%	29 (28.4)	34 (33.7)	
Hormone therapy use status			0.64
Never	59 (57.8)	56 (54.9)	
Former	35 (34.3)	34 (33.3)	
Current (within previous 3 months)	8 (7.8)	12 (11.8)	
Oral contraceptive use status			0.14[†]
Never	36 (36.7)	30 (32.3)	
Former	61 (62.2)	57 (61.3)	
Current	1 (1.0)	6 (6.5)	
Serum E2 level, pg/mL			0.33
0.0 – <4.1	21 (20.6)	31 (30.4)	
4.1 – <11.2	29 (28.4)	21 (20.6)	
11.2 – <40.9	25 (24.5)	26 (25.5)	
≥40.9	27 (26.5)	24 (23.5)	
Serum FSH level, mIU/mL			0.04
0.3 – 31.4	20 (19.6)	31 (30.4)	
31.4 – <88.8	33 (32.4)	19 (18.6)	
88.8 – <138.9	28 (27.5)	22 (21.6)	
≥138.9	21 (20.6)	30 (29.4)	
Serum SHBG level, nM			0.94
1.95 – 35.3	26 (25.5)	25 (24.5)	
35.3 – 51.6	25 (24.5)	26 (25.5)	
51.6 – 73.1	27 (26.5)	24 (23.5)	
≥73.1	24 (23.5)	27 (26.5)	
Serum T level, ng/dL			0.17
5.2 – <18.4	21 (20.6)	30 (29.4)	
18.4 – <29.2	24 (23.5)	27 (26.5)	
29.2 – <42.2	32 (31.4)	19 (18.6)	
≥42.2	25 (24.5)	26 (25.5)	

[*]P values from chi-square test
[†]P value from Fisher's exact test
Abbreviations used: VEGF, vascular endothelial growth factor; MET, metabolic equivalent, E2, estradiol; FSH, follicle stimulating hormone; SHBG, sex hormone binding globulin; T, testosterone

Table 5.4 Results of multivariable logistic regressions for association between serum VEGF and breast cancer[*]

	Unadjusted			Age adjusted			Multivariable adjusted[†]		
	N	**OR (95% CI)**	**P Value**	**N**	**OR (95% CI)**	**P Value**	**N**	**OR (95% CI)**	**P Value**
Total sample									
Continuous VEGF[**]	406	1.17 (0.91 – 1.50)	0.22	406	1.17 (0.91 – 1.50)	0.21	402	1.21 (0.91 – 1.59)	0.19
Categorical VEGF, split at median	406	1.22 (0.83 – 1.80)	0.32	406	1.22 (0.83 – 1.80)	0.32	402	1.37 (0.88 – 2.12)	0.16
Menopausal status									
Premenopausal	132	1.44 (0.73 – 2.86)	0.30	132	1.46 (0.73 – 2.91)	0.29	131	1.40 (0.64 – 3.07)	0.40
Postmenopausal	274	1.13 (0.70 – 1.81)	0.62	274	1.13 (0.70 – 1.82)	0.62	271	1.28 (0.74 – 2.22)	0.38
Hormone therapy use									
Never user	223	1.13 (0.67 – 1.92)	0.64	223	1.12 (0.66 – 1.90)	0.66	223	1.23 (0.68 – 2.22)	0.50
Past user	114	1.70 (0.79 – 3.65)	0.18	114	1.74 (0.81 – 3.77)	0.16	112	2.28 (0.86 – 5.68)	0.08
Current user (within previous 3 months)	67	0.83 (0.28 – 2.39)	0.72	67	0.84 (0.29 – 2.43)	0.74	67	1.09 (0.31 – 3.87)	0.90

[*]Odds ratios comparing individuals with VEGF≥314.2 pg/mL to those with VEGF<314.2 pg/mL unless otherwise specified
[†]Adjusted for age, Gail score, education, physical activity, history of breastfeeding, serum testosterone, HT use
[**]Odds ratios in this row are for a 1 unit increase in ln(VEGF) beyond the control population mean of 5.67 (219.4 pg/mL in the observed scale)
Abbreviations used: OR, odds ratio; CI, confidence interval

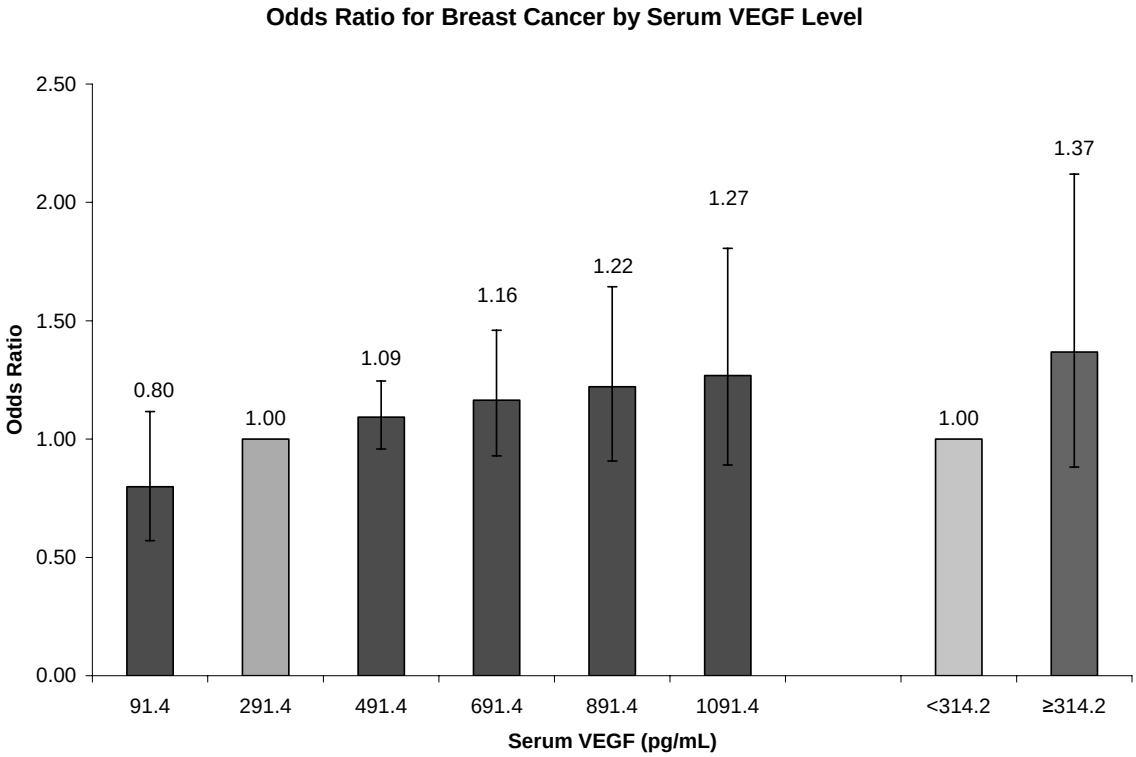


Figure 5.1 Odds of breast cancer by serum VEGF level estimated from a multivariable logistic regression model, N=402[*]

[*]Odds ratios are adjusted for age, Gail score, education, physical activity, history of breastfeeding, serum testosterone, and hormone therapy use

6.0 DISCUSSION

6.1 SUMMARY OF FINDINGS

Although breast cancer survival has improved over recent decades as a result of early detection and improved treatment,[1] breast cancer still impacts the lives of thousands of women each year. During 2007 alone, nearly 180,000 women in the United States were diagnosed with breast cancer.[1] The mechanisms which lead to the development of breast cancer, however, are not completely understood. As a result, opportunities for disease prevention are limited. Therefore, research which focuses on elucidating the mechanisms of breast carcinogenesis is extremely important. Knowledge of the mechanisms which are responsible for breast carcinogenesis may reveal opportunities for disease prevention through either behavioral modification or chemoprevention.

We undertook investigations of two factors of potential etiological importance to breast cancer: anthropometry and angiogenesis. First, we sought to determine how two anthropometric measures, body mass index (BMI) and weight, related to longitudinal changes in mammographic breast density, a well-established, modifiable risk factor for breast cancer. During the course of this investigation we also developed and validated an estimation approach for off-schedule outcomes data. Additionally, we studied whether circulating levels of vascular endothelial growth factor (VEGF), a strong angiogenic factor, increase the risk of breast cancer.

6.1.1 Article 1: Longitudinal influence of anthropometry on mammographic breast density

We used random intercept regression models to study the associations between anthropometry and mammographic breast density in 834 women who participated in an ancillary study to the Study of Women's Health Across the Nation (SWAN). Anthropometric data were measured annually at clinic visits, and breast density was estimated from retrospectively collected mammograms. We found that both BMI and weight were positively and significantly associated with dense breast area in cross-sectional analysis, but neither measure was longitudinally associated with dense breast area. Conversely, BMI and weight were significantly negatively associated with percent density in both cross-sectional and longitudinal analyses.

The results of our cross-sectional analyses are in agreement with previous studies which demonstrate a negative association between anthropometry and percent density.[127, 133, 175-180, 186, 338] Inconsistent results have been reported from cross-sectional analyses of anthropometry and the dense breast area,[133, 179, 186] though our finding of a positive association is in agreement with one previous study.[133] Few studies have investigated these associations using longitudinal studies,[136, 181, 187] and differences in methodology preclude direct comparison between two of these studies and our own.[181, 187] Boyd et al. reported a negative association between weight change and change in percent density, in agreement with our results, yet a positive association between weight change and change in dense breast area.[136] The reasons for the difference between our results and those of Boyd et al. in regard to the dense breast area are unclear, but they may relate to differences between the study populations in characteristics such as age, race/ethnicity, and level of breast cancer risk.

The results of our study indicate that changes in anthropometry do not impact breast cancer development through direct effects on the dense breast area. Overweight and obese women tend to have larger breasts than underweight or normal weight women,[338] explaining the positive cross-sectional associations between BMI and weight and the dense breast area. Further increases in weight and BMI appear to preferentially affect the non-dense breast area. Such an increase in the non-dense area also increases the total breast area. Thus when there is not an equivalent increase in the size of the dense breast area, the overall effect of increased non-dense breast area is to decrease percent breast density. Adipose tissue, such as that in the non-dense area of the breast, is the primary source of estrogen production for postmenopausal women.[115, 116] Local exposure of the neighboring dense breast tissue, where cancers arise, to estrogen produced by aromatization in non-dense breast tissue may be an important factor in breast cancer etiology.[133, 187] Our results suggest that the non-dense breast area be investigated further for a potential role in breast cancer development, though previous research has focused primarily on the dense breast area in terms of disease etiology. Additionally, we have demonstrated that investigators must consider effects of exposures on both the dense and non-dense tissues, rather than focusing solely on percent density.

6.1.2 Article 2: Linear interpolation with multiple imputation to account for off-schedule observations in a longitudinal study

The analysis of the research question posed in the first article was complicated by the design of the SWAN Mammographic Density Study. Mammograms were collected retrospectively from participants' mammogram facilities. Though these mammograms were taken while the participants were enrolled in SWAN, they were not part of the SWAN protocol and therefore

rarely coincided with the dates of the participants' SWAN visits. As a result, we needed a way to estimate breast density measurements at the time of the SWAN visits from the off-schedule mammograms. Using an approach that simply matched each mammogram to the participant's nearest study visit was problematic because the length of time between each mammogram and its nearest visit varied both within and between participants. We developed an estimation approach that used linear interpolation to estimate breast density measurements at the time of each SWAN visit from mammograms taken before and after the SWAN visit. We further added a multiply imputed noise term to the interpolated estimate to account for the error introduced by estimating these measurements.

In a validation study, we demonstrated that the bias of our method for estimating dense breast area was quite low on average (0.11 cm^2). Our linear interpolation with multiple imputation approach was compared to three other approaches, including simple matching, linear interpolation, and linear interpolation with single imputation of noise terms. The bias from our approach was smaller on average and less variable than the bias from the other estimation approaches. We also investigated how each estimation approach would affect the results of random intercept regression models with estimated dense breast area as the outcome variable and BMI as the independent variable. The regression coefficient was largest in magnitude when breast density was estimated using the simple matching algorithm. Further, the association between BMI and dense breast area was statistically significant only when the simple matching algorithm was used. The variance of the regression coefficients was lowest when linear interpolation was used, and was substantially increased when singly or multiply imputed noise terms were added to the linear interpolation estimates. The variance of the regression coefficient was slightly reduced when mammograms within 90 days of the nearest SWAN visit were used as

matches and linear interpolation with multiple imputation was used for visits without a mammogram within 90 days. The magnitude of the regression coefficient was similar between these two approaches. The approach where observed data were used for mammograms within 90 days of the visit and estimated with linear interpolation and multiple imputation otherwise was employed in the analyses for Article 1.

We have demonstrated that our method for estimating off-schedule outcomes data is valid. Further, we have shown that the approach used to estimate such data can have important effects on the inferences made from analyses using the estimated data. For instance, we would have judged the relationship between BMI and dense breast area to be strong and statistically significant when using the simple matching algorithm, a naïve approach, whereas all other estimation approaches showed a weaker, non-significant relationship. Though further development and testing of our approach is warranted, our linear interpolation with multiple imputation approach may be applicable to other longitudinal studies with important data collected outside of the regularly scheduled study visits.

6.1.3 Article 3: Vascular endothelial growth factor and breast cancer risk

The final research aim focused on the relationship between angiogenesis and breast cancer. Angiogenesis is a key step in tumor growth and metastasis, as tumors cannot grow without an adequate blood supply.[194] VEGF is the strongest known angiogenic factor,[195] and is therefore of potential importance to breast carcinogenesis. We performed an unmatched case-control study with 203 cases and 204 controls to investigate whether higher serum VEGF levels were associated with increased breast cancer risk.

No statistically significant differences were detected between serum VEGF levels in cases and controls, though the unadjusted geometric mean VEGF among cases (321.4 pg/mL) was higher than that of controls (291.4 pg/mL). We observed a positive, but non-significant, association between VEGF levels ≥314.2 pg/mL and breast cancer (OR 1.37, 95% CI 0.88-2.12). Our results agree with previous studies in terms of the direction of the association between serum VEGF and breast cancer,[220, 310, 311, 313-318, 320] though the positive association in our study did not achieve statistical significance. The lack of statistical significance may be attributable to the low power of our study. Though *a priori* calculations indicated sufficient power for this analysis, the variability of serum VEGF levels in our control population was much greater than that on which the power calculations were based.[310] Although previous studies did not describe their study population or recruitment procedures in sufficient detail, it appears that ours may be the first breast cancer case-control study to recruit control subjects from the general population. The higher than expected VEGF levels among our controls may reflect unique characteristics of our study population, or they may indicate that levels of VEGF in healthy women are higher than previously recognized. Regardless, it does appear that VEGF levels are elevated among women with breast cancer compared to those without evidence of disease, though perhaps this difference is not statistically significant. Prospective studies will be valuable in determining whether or not measurement of VEGF levels may be useful in predicting breast cancer risk.

6.2 FUTURE RESEARCH

As indicated within each article, there are many opportunities for future research related to the investigations reported here. For example, it would be useful to replicate our findings regarding

anthropometry and breast density in a longitudinal study where mammograms were taken as part of the study protocol. Such a study design would eliminate the need to estimate breast density measurements as in our study, and would help evaluate the extent to which our results were influenced by estimation of our outcome data. Further, identifying the effects of weight loss on both the dense and non-dense breast tissue would aid in understanding the mechanisms by which weight loss affects breast cancer risk. It is also important to demonstrate how changes in breast density relate to changes in breast cancer risk. This latter point is integral to demonstrating that breast density is useful as a surrogate endpoint for breast cancer.

Regarding our estimation approach for off-schedule outcome data, future research is needed to refine the approach. We could also consider other methods for choosing the noise terms. Additionally, our estimation approach should be applied and validated in another longitudinal dataset to evaluate its performance for estimating measures other than breast density.

More research is needed to fully understand the association between VEGF and breast cancer risk. Larger case-control studies with more power are an important next step, as are prospective studies which could evaluate the utility of circulating VEGF levels for predicting which women will develop breast cancer. VEGF levels could also be measured in the participants with benign breast disease in the MAMS population to determine what the VEGF levels are in this population of high-risk women and how they compare to cases and controls. Also, there are many opportunities for future research focusing on the associations between sex steroid hormones and VEGF.

6.3 PUBLIC HEALTH SIGNIFICANCE

Breast cancer is the most commonly diagnosed non-skin cancer among American women, and ranks second in terms of cancer mortality for this population as well.[1] Though such statistics clearly indicate the importance of breast cancer as a public health problem, the symbolic pink ribbons which now appear on everything from cars to cookbooks convey the true importance of this disease to the public. Greater understanding of the factors which influence the development of breast cancer are important for identifying opportunities for prevention of this disease. We, therefore, focused on how anthropometry and angiogenesis relate to breast cancer etiology.

This research makes a significant contribution to public health. First, our demonstration of the longitudinal associations between anthropometry and breast density has increased our understanding of how a well known risk factor for postmenopausal breast cancer, elevated BMI, may influence breast density, another breast cancer risk factor. Further, we have developed an estimation approach for off-schedule data that may be used by other longitudinal studies. Specifically, the application of this method may expand the research questions that can be answered in studies where data of interest were collected outside of planned study visits. Finally, we have demonstrated that VEGF levels are positively associated with breast cancer, though whether this association is statistically or clinically significant is not clear. It is not yet known whether VEGF levels will be useful in distinguishing between women with and without breast cancer. In summary, these investigations have increased our knowledge of breast cancer etiology. Our results, along with those of future studies expanding upon our findings, may lead to improved opportunities for prevention or early detection of breast cancer.

APPENDIX A

LONGITUDINAL INFLUENCE OF ANTHROPOMETRY ON MAMMOGRAPHIC BREAST DENSITY: THE STUDY OF WOMEN'S HEALTH ACROSS THE NATION (SWAN)

A.1 COMPARISON OF PARTICIPANTS TO NON-PARTICIPANTS FROM SWAN

Table A.1 Comparison of baseline characteristics among SWAN participants enrolled in the SWAN Mammographic Density Study to SWAN participants at a participating SWAN site who did not enroll in the SWAN Mammographic Density Study

Characteristic	SWAN Mammographic Density Study participants N=1,007	SWAN participants not enrolled in Mammographic Density Study* N=411	P value†
Age, years; mean (SD)	45.9 (2.7)	45.6 (2.7)	0.06
Body mass index, kg/m^2; mean (SD)	25.6 (5.9)	26.9 (6.8)	<0.001
Weight, kg; mean (SD)	66.8 (17.4)	71.0 (19.3)	<0.001
Race/ethnicity; N (%)			<0.001
African American	91 (9.0)	71 (17.3)	
Caucasian	492 (48.9)	233 (56.7)	
Chinese	198 (19.7)	52 (12.7)	
Japanese	226 (22.4)	55 (13.4)	
Menopausal status; N (%)			0.01
Premenopausal	584 (58.4)	208 (50.9)	
Early perimenopausal	416 (41.6)	201 (49.1)	
Education; N (%)			<0.001
Less than high school	30 (3.0)	10 (2.4)	
High school	142 (14.1)	81 (21.2)	
Some college	313 (31.1)	150 (36.5)	
College	261 (25.9)	78 (19.0)	
Post-college	261 (26.0)	86 (20.9)	
SWAN clinical site; N (%)			<0.001
Oakland, CA	341 (33.9)	118 (28.7)	
Los Angeles, CA	392 (38.9)	104 (25.3)	
Pittsburgh, PA	274 (27.2)	189 (46.0)	

*Includes only SWAN participants from one of three sites offering the Mammographic Density Study (Los Angeles, CA, Oakland, CA, Pittsburgh, PA) but who were either ineligible or chose not to enroll in this ancillary study
†P value from ANOVA for continuous variables and chi square test for categorical variables

Table A.2 Comparison of SWAN Mammographic Density Study participants included and excluded from the present analysis

Characteristic	Included in Analysis N=834	Excluded from Analysis N=173	P value*
Age, years; mean (SD)	46.0 (2.7)	45.7 (2.6)	0.26
Body mass index, kg/m^2; mean (SD)	25.4 (5.9)	26.4 (6.2)	0.06
Weight, kg; mean (SD)	66.3 (17.1)	69.1 (18.3)	0.06
Race/ethnicity; N (%)			<0.001
African American	62 (7.4)	29 (16.8)	
Caucasian	407 (48.8)	85 (49.1)	
Chinese	182 (21.8)	16 (9.3)	
Japanese	183 (21.9)	43 (24.9)	
Menopausal status; N (%)			0.85
Premenopausal	483 (58.3)	101 (59.1)	
Early perimenopausal	346 (41.7)	70 (40.9)	
Education; N (%)			0.08
Less than high school	26 (3.1)	4 (2.3)	
High school	110 (13.2)	32 (18.5)	
Some college	250 (30.0)	63 (36.4)	
College	223 (26.7)	38 (22.0)	
Post-college	225 (27.0)	36 (20.8)	
SWAN clinical site; N (%)			<0.001
Oakland, CA	315 (37.8)	26 (15.0)	
Los Angeles, CA	321 (38.5)	71 (41.0)	
Pittsburgh, PA	198 (23.7)	76 (43.9)	

*P value from ANOVA for continuous variables and chi square test for categorical variables

A.2 HORMONE USE AND MENOPAUSAL STATUS OVER FOLLOW-UP

Table A.3 Hormone use by study population and menopausal status throughout follow-up, N=834

	Visit 1*		Visit 2		Visit 3		Visit 4		Visit 5		Visit 6		Visit 7	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Ever used hormone therapy/oral contraceptives†	67	8.1	138	17.0	201	24.3	253	30.9	300	36.4	328	40.1	348	42.4
Hormone therapy/oral contraceptive use since last visit	67	8.1	124	15.3	158	19.1	200	24.4	222	27.0	224	27.4	143	17.7
Menopausal status														
Pre-/Early perimenopausal	716	86.9	603	74.4	529	64.1	426	52.1	335	40.7	269	32.8	215	26.2
Late perimenopausal/Postmenopausal**	44	5.3	102	12.6	191	23.2	276	33.7	379	46.1	455	55.6	558	68.0
Unknown due to hormone therapy use	64	7.8	106	13.1	105	12.7	116	14.2	109	13.2	95	11.6	48	5.9

*Visit 1 is the first follow-up SWAN visit and is not the baseline SWAN visit; not all 834 participants have complete data at each visit
†This is a time-dependent variable evaluated at each SWAN visit
**Includes women who are postmenopausal due to hysterectomy with bilateral oophorectomy

A.3 CHARACTERISTICS OF STUDY POPULATION BY SUB-GROUPS

Table A.4 Selected baseline characteristics of the study population by race/ethnicity, N=834

Baseline characteristic	African American		Caucasian		Chinese		Japanese		P value*
General characteristics									
Age, years; mean (SD)	46.0	2.3	46.3	2.8	46.5	2.5	46.9	2.6	0.03
Family income; N (%)									<0.001
<$35,000	31	50.0	43	10.7	38	21.1	14	8.1	
$35,000-$49,999	9	14.5	75	18.6	37	20.6	19	10.9	
$50,000-$74,999	12	19.4	105	26.1	46	25.6	49	28.1	
$75,000-$99,999	7	11.3	74	18.4	22	12.2	38	21.8	
≥ $100,000	3	4.8	106	26.3	37	20.6	53	31.0	
Education; N (%)									<0.001
≤ High school	14	22.6	45	11.1	51	28.0	26	14.2	
>High school	22	35.5	121	29.7	41	22.5	66	36.1	
College	11	17.7	100	24.6	52	28.6	60	32.8	
Post-college	15	24.2	141	34.6	38	20.9	31	16.9	
Anthropometric characteristics									
Body mass index, kg/m^2; mean (SD)	31.5	6.4	26.5	6.5	23.4	3.8	23.1	3.5	<0.001
BMI category; N (%)									<0.001
Underweight	0	0.0	7	1.7	4	2.2	5	2.8	
Normal	7	11.7	199	49.4	133	73.9	135	74.6	
Overweight	20	33.3	113	28.0	35	19.4	34	18.8	
Obese	33	55.0	84	20.8	8	4.4	7	3.9	
Reproductive history									
Menopausal status; N (%)									0.64
Premenopausal	34	54.8	231	57.0	112	62.2	106	58.2	
Early Perimenopausal	28	45.2	174	43.0	68	37.8	76	41.8	
Other									
Gail Score; mean (SD)	0.62	0.36	1.13	0.52	1.03	0.32	1.09	0.34	<0.001

*P values from Kruskal-Wallis tests for continuous variables and chi square tests for categorical variables

Table A.5 Selected baseline characteristics of Caucasian participants by site, N=407

Baseline characteristic	Oakland, CA N=131		Los Angeles, CA N=140		Pittsburgh, PA N=142		P value*
General characteristics							
Age, years; mean (SD)	46.5	3.0	46.4	2.8	46.2	2.6	0.88
Family income; N (%)							<0.001
<$35,000	13	9.9	11	8.0	19	14.2	
$35,000-$49,999	26	19.8	19	13.9	30	22.4	
$50,000-$74,999	39	29.6	25	18.3	41	30.6	
$75,000-$99,999	25	18.9	24	17.5	25	18.7	
≥ $100,000	29	22.0	58	42.3	19	14.2	
Education; N (%)							0.02
≤ High school	9	6.8	12	8.7	24	17.7	
>High school	32	24.1	46	33.3	43	31.6	
College	38	28.6	35	25.4	27	19.9	
Post-college	54	40.6	45	32.6	42	30.9	
Anthropometric characteristics							
Body mass index, kg/m^2; mean (SD)	27.6	7.8	24.9	5.9	27.0	5.2	<0.001
BMI category; N (%)							0.003
Underweight	3	2.3	4	3.0	0	0.0	
Normal	62	47.0	81	60.0	56	41.2	
Overweight	32	24.2	30	22.2	51	37.8	
Obese	35	26.5	20	14.8	29	21.3	
Reproductive history							
Menopausal status; N (%)							0.43
Premenopausal	70	52.6	83	60.1	78	58.2	
Early Perimenopausal	63	47.4	55	39.9	56	41.8	
Other							
Gail Score; mean (SD)	1.09	0.4	1.20	0.6	1.10	0.5	0.51
Number of available mammograms; N (%)							<0.001
2	10	7.5	20	14.5	19	14.0	
3	18	13.5	18	13.0	33	24.3	
4	38	28.6	30	21.7	22	16.2	
5	27	20.3	21	15.2	13	9.6	
6	27	20.3	30	21.7	16	11.8	
≥ 7	13	9.8	19	13.8	33	24.3	

*P values from Kruskal-Wallis tests for continuous variables and chi square tests for categorical variables

Table A.6 Selected baseline characteristics of the study population by baseline body mass index category, N=834

	BMI Category at Baseline								
Baseline characteristic	**Underweight**		**Normal**		**Overweight**		**Obese**		**P value***
General characteristics									
Age, years; mean (SD)	45.8	3.5	46.5	2.7	46.5	2.7	46.3	2.7	0.44
Family income; N (%)									0.05
<$35,000	2	12.5	62	13.4	31	15.7	29	22.0	
$35,000-$49,999	2	12.5	71	15.3	40	20.3	26	19.7	
$50,000-$74,999	2	12.5	118	25.4	52	26.4	39	29.6	
$75,000-$99,999	3	18.8	86	18.5	32	16.2	18	13.6	
≥ $100,000	7	43.8	127	27.4	42	21.3	20	15.2	
Education; N (%)									0.15
≤ High school	1	6.3	86	18.1	25	12.4	24	18.2	
>High school	5	31.3	132	27.9	61	30.2	48	36.4	
College	4	25.0	137	28.9	50	24.8	29	22.0	
Post-college	6	37.5	119	25.1	66	32.7	31	23.5	
Reproductive history									
Menopausal status; N (%)									0.64
Premenopausal	11	73.3	279	58.9	117	58.5	73	56.2	
Early Perimenopausal	4	26.7	195	41.4	83	41.5	57	43.9	
Other									
Gail Score; mean (SD)	1.15	0.51	1.10	0.46	1.04	0.44	0.93	0.41	<0.001

*P values from Kruskal-Wallis tests for continuous variables and chi square tests for categorical variables

Table A.7 Selected baseline characteristics of the study population by number of available mammograms, N=834

Baseline characteristic	Number of Mammograms 2		3		4		5		6		≥ 7		P value*
General characteristics													
Age, years; mean (SD)	45.9	2.4	46.4	2.8	46.3	2.8	46.3	2.5	46.8	2.6	47.2	2.8	0.01
Race/ethnicity													0.001
African American	18	14.6	14	8.9	8	4.8	5	3.5	7	5.6	10	8.4	
Caucasian	49	39.8	69	43.7	90	54.2	61	43.0	73	57.9	65	54.6	
Chinese	19	15.5	37	23.4	42	25.3	39	27.5	22	17.5	23	19.3	
Japanese	37	30.1	38	24.1	26	15.7	37	26.1	24	19.1	21	17.7	
Family income; N (%)													0.005
<$35,000	32	26.5	29	18.6	21	13.0	19	13.6	12	9.7	13	11.1	
$35,000-$49,999	24	19.8	24	15.4	40	24.8	17	12.1	20	16.1	15	12.8	
$50,000-$74,999	26	21.5	39	25.0	44	27.3	40	28.6	36	29.0	27	23.1	
$75,000-$99,999	12	9.9	27	17.3	23	14.3	31	22.1	25	20.2	23	19.7	
≥ $100,000	27	22.3	37	23.7	33	20.5	33	23.6	31	25.0	39	33.3	
Education; N (%)													0.11
≤ High school	29	23.6	24	15.2	28	16.9	21	14.8	16	12.7	18	15.2	
>High school	41	33.3	54	34.2	45	27.1	44	31.0	38	30.2	28	23.5	
College	31	25.2	49	31.0	41	24.7	37	26.1	32	25.4	33	27.7	
Post-college	22	17.8	31	19.6	52	31.3	40	28.2	40	31.8	40	33.6	
Anthropometric characteristics													
Body mass index, kg/m^2; mean (SD)	26.3	7.1	25.4	6.0	25.6	6.3	24.7	4.7	25.3	5.8	25.3	5.0	0.80
Body mass index category; N (%)													0.36
Underweight	2	1.7	5	3.2	3	1.8	1	0.7	1	0.8	4	3.4	
Normal	65	54.6	88	55.7	94	57.3	87	61.3	78	62.9	62	53.0	
Overweight	24	20.2	38	24.1	45	27.4	38	26.8	25	20.2	32	27.4	
Obese	28	23.5	27	17.1	22	13.4	16	11.3	20	16.1	19	16.2	
Reproductive history													
Menopausal status; N (%)													0.04
Premenopausal	68	55.3	109	69.0	98	59.8	81	57.0	67	53.2	60	51.7	
Early Perimenopausal	55	44.7	49	31.0	66	40.2	61	43.0	59	46.8	56	48.3	

Table A.7 continued

Other													
Gail Score; mean (SD)	0.94	0.35	0.99	0.37	1.02	0.41	1.06	0.34	1.07	0.42	1.36	0.69	<0.001
Site; N (%)													<0.001
Oakland, CA	29	23.6	55	34.8	80	48.2	66	46.5	49	38.9	36	30.3	
Los Angeles, CA	57	46.3	56	35.4	56	33.7	58	40.9	54	42.9	40	33.6	
Pittsburgh, PA	37	30.1	47	29.8	30	18.1	18	12.7	23	18.3	43	36.1	

*P values from Kruskal-Wallis tests for continuous variables and chi square tests for categorical variables

A.4 MAMMOGRAM CHARACTERISTICS

Table A.8 Summary statistics of time between SWAN visit and matched mammogram, by study visit

	Time between visit and matched mammogram* (days)							
Visit	**N**	**Mean**	**SD**	**Minimum**	**25%**	**Median**	**75%**	**Maximum**
0	518	-166.1	269.9	-1515	-328	-114.5	52	199
1	361	9.4	115.8	-251	-93	12	107	480
2	387	3.5	108.6	-360	-91	8	89	307
3	431	-4.6	116.4	-334	-98	-7	88	564
4	459	-3.8	109.2	-317	-95	4	86	257
5	519	4.6	112.7	-324	-78	11	92	463
6	631	8.6	108.9	-290	-84	22	88	227
7	255	-62.8	130.2	-326	-161	-87	25	350

*Negative value indicates mammogram taken before visit

Table A.9 Distribution of mammograms taken within 90, 120, and 180 days of the matched SWAN visit

		Mammogram within 90 days		Mammogram within 120 days		Mammogram within 180 days	
Visit	**Total N with Mammogram at Visit**	**N**	**%**	**N**	**%**	**N**	**%**
0	518	147	28.4	190	36.7	297	57.3
1	361	163	45.2	226	62.6	339	93.9
2	387	195	50.4	254	65.6	370	95.6
3	431	207	48.0	279	64.7	405	94.0
4	459	230	50.1	300	65.4	438	95.4
5	519	264	50.9	347	66.9	483	93.1
6	631	331	52.5	419	66.4	583	92.4
7	255	94	36.9	125	49.0	203	79.6

Table A.10 Distribution of number of mammograms within 90 days of matched visit by participant for total population and stratitifed by baseline BMI category and by clinic site, N=834[*]

	Total Population		Underweight		Normal		Overweight		Obese		Oakland, CA		Los Angeles, CA		Pittsburgh, PA	
Number of mammograms within 90 days	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
0	116	13.9	1	6.3	60	12.7	32	15.8	23	17.4	35	11.1	50	15.6	31	15.7
1	240	28.8	4	25.0	137	28.9	52	25.7	44	33.3	87	27.6	88	27.4	65	32.8
2	240	28.8	5	31.3	141	29.8	60	29.7	31	32.5	100	31.8	87	27.1	53	26.8
3	127	15.2	4	25.0	70	14.8	30	14.9	20	15.2	50	15.9	51	15.9	26	13.2
4	56	6.7	1	6.3	38	8.0	10	5.0	7	5.3	23	7.3	28	8.7	5	2.5
5	33	4.0	0	0.0	15	3.2	12	6.0	6	4.6	15	4.8	9	2.8	9	4.6
6	14	1.7	1	6.3	8	1.7	3	1.5	1	0.8	4	1.3	4	1.3	6	3.0
7	7	0.8	0	0.0	4	0.8	3	1.5	0	0.0	0	0.0	3	1.0	3	1.5
8	1	0.1	0	0.0	1	0.2	0	0.0	0	0.0	1	0.3	0	0.0	0	0.0

[*]P value from chi square tests of difference by BMI category p = 0.76; p value from chi square test of difference by site p = 0.08

A.5 RESULTS OF RANDOM INTERCEPT REGRESSIONS BY SUB-GROUPS

Table A.11 Random intercept regression estimates for the outcomes of dense area and percent density using multiple imputation, by race/ethnicity*

	African American			Caucasian			Chinese			Japanese		
	N	**β**	**P Value**	**N**	**β**	**P Value**	**N**	**β**	**P Value**	**N**	**β**	**P Value**
Dense breast area												
Body mass index, kg/m^2												
Model 1: BMI	62	-0.058	0.13	403	-0.031	0.03	182	-0.013	0.50	183	0.019	0.28
Model 2: BMI + age	62	-0.053	0.18	403	-0.013	0.39	182	0.005	0.81	183	0.035	0.05
Model 3: Fully adjusted†	62	-0.053	0.18	401	-0.012	0.42	180	0.005	0.82	181	0.029	0.13
Weight, kg												
Model 1: Weight	62	-0.027	0.06	403	-0.013	0.02	182	-0.002	0.82	183	0.007	0.31
Model 2: Weight + age	62	-0.026	0.07	403	-0.007	0.23	182	0.005	0.52	183	0.013	0.07
Model 3: Fully adjusted†	62	-0.026	0.07	401	-0.006	0.27	180	0.005	0.53	181	0.011	0.18
Percent breast density												
Body mass index, kg/m^2												
Model 1: BMI	62	-1.135	<0.001	403	-1.422	<0.001	182	-1.729	<0.001	183	-1.401	<0.001
Model 2: BMI + age	62	-1.104	<0.001	403	-1.230	<0.001	182	-1.455	<0.001	183	-1.200	<0.001
Model 3: Fully adjusted**	61	-1.061	<0.001	401	-1.133	<0.001	180	-1.331	<0.001	181	-1.050	<0.001
Weight, kg												
Model 1: Weight	62	-0.437	<0.001	403	-0.502	<0.001	182	-0.630	<0.001	183	-0.582	<0.001
Model 2: Weight + age	62	-0.431	<0.001	403	-0.438	<0.001	182	-0.531	<0.001	183	-0.514	<0.001
Model 3: Fully adjusted**	61	-0.447	<0.001	401	-0.409	<0.001	180	-0.486	<0.001	181	-0.442	<0.001

*Dense area was square root transformed due to non-normality; percent density was modeled in the natural scale

†Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

**Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

Table A.12 Random intercept regression estimates for the outcome of dense area using multiple imputation, by race/ethnicity and BMI category at SWAN enrollment*

	African American			Caucasian			Chinese			Japanese		
	N	**β**	**P Value**	**N**	**β**	**P Value**	**N**	**β**	**P Value**	**N**	**β**	**P Value**
Normal BMI at SWAN enrollment												
Body mass index, kg/m^2												
Model 2: BMI + age	7	0.041	0.70	197	0.027	0.21	133	0.024	0.34	135	0.018	0.51
Model 3: Fully adjusted†	7	0.040	0.70	197	0.023	0.30	131	0.025	0.33	134	0.018	0.49
Weight, kg												
Model 2: Weight + age	7	0.023	0.56	197	0.009	0.29	133	0.012	0.22	135	0.007	0.46
Model 3: Fully adjusted†	7	0.020	0.59	197	0.008	0.35	131	0.013	0.22	134	0.007	0.44
Overweight BMI at SWAN enrollment												
Body mass index, kg/m^2												
Model 2: BMI + age	20	-0.023	0.74	112	0.019	0.51	35	0.001	0.98	34	0.001	0.98
Model 3: Fully adjusted†	20	-0.013	0.85	111	0.018	0.51	35	-0.014	0.76	34	0.005	0.91
Weight, kg												
Model 2: Weight + age	20	-0.014	0.59	112	0.004	0.67	35	-0.002	0.90	34	0.004	0.81
Model 3: Fully adjusted†	20	-0.012	0.66	111	0.004	0.66	35	-0.006	0.70	34	0.006	0.75
Obese BMI at SWAN enrollment												
Body mass index, kg/m^2												
Model 2: BMI + age	33	-0.106	0.06	83	-0.007	0.82	8	-0.117	0.34	7	0.084	0.24
Model 3: Fully adjusted†	33	-0.100	0.12	82	-0.006	0.86	8	-0.047	0.65	7	0.068	0.34
Weight, kg												
Model 2: Weight + age	33	-0.047	0.01	83	-0.005	0.67	8	-0.036	0.51	7	0.032	0.27
Model 3: Fully adjusted†	33	-0.045	0.04	82	-0.004	0.72	8	0.006	0.90	7	0.024	0.41

*Dense area was square root transformed due to non-normality; percent density was modeled in the natural scale

†Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

Table A.13 Random intercept regression estimates for the outcome of percent density using multiple imputation, by race/ethnicity and BMI category at SWAN enrollment*

	African American			Caucasian			Chinese			Japanese		
	N	**β**	**P Value**	**N**	**β**	**P Value**	**N**	**β**	**P Value**	**N**	**β**	**P Value**
Normal BMI at SWAN enrollment												
Body mass index, kg/m^2												
Model 2: BMI + age	7	0.948	0.61	197	-1.258	<0.001	133	-1.212	0.01	135	-1.225	0.002
Model 3: Fully adjusted†	7	0.227	0.95	197	-1.269	<0.001	132	-1.033	0.03	134	-1.213	0.002
Weight, kg												
Model 2: Weight + age	7	0.571	0.41	197	-0.422	<0.001	133	-0.418	0.01	135	-0.535	<0.001
Model 3: Fully adjusted†	7	0.211	0.89	197	-0.442	<0.001	132	-0.388	0.02	134	-0.492	0.001
Body mass index, kg/m^2												
Model 2: BMI + age	20	-1.499	0.07	112	-0.579	0.06	35	-0.653	0.30	34	-1.040	0.09
Model 3: Fully adjusted†	20	-4.879	<0.001	111	-0.563	0.07	34	-0.820	0.21	33	-0.764	0.25
Weight, kg												
Model 2: Weight + age	20	-0.619	0.05	112	-0.212	0.04	35	-0.251	0.27	34	-0.322	0.18
Model 3: Fully adjusted†	20	-2.028	<0.001	111	-0.210	0.05	34	-0.317	0.19	33	-0.272	0.29
Body mass index, kg/m^2												
Model 2: BMI + age	33	-0.861	0.05	83	-0.295	0.06	8	-0.985	0.17	7	0.102	0.87
Model 3: Fully adjusted†	33	-0.833	0.11	82	-0.287	0.09	8	-0.593	0.53	7	0.130	0.86
Weight, kg												
Model 2: Weight + age	33	-0.350	0.02	83	-0.112	0.06	8	-0.265	0.44	7	0.006	0.98
Model 3: Fully adjusted†	33	-0.399	0.04	82	-0.114	0.06	8	-0.226	0.57	7	0.065	0.83

*Dense area was square root transformed due to non-normality; percent density was modeled in the natural scale

†Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

Table A.14 Random intercept regression estimates for the outcomes of dense area and percent density using multiple imputation restricted to participants with no hormone use throughout SWAN follow-up[*]

	N	β	P Value
Dense breast area			
Body mass index, kg/m^2			
Model 1: BMI	441	-0.006	0.63
Model 2: BMI + age	441	0.011	0.40
Model 3: Fully adjusted[†]	439	0.001	0.94
Weight, kg			
Model 1: Weight	441	-0.002	0.62
Model 2: Weight + age	441	0.003	0.50
Model 3: Fully adjusted[†]	439	-0.002	0.71
Percent breast density			
Body mass index, kg/m^2			
Model 1: BMI	441	-1.504	<0.001
Model 2: BMI + age	441	-1.297	<0.001
Model 3: Fully adjusted[**]	438	-1.171	<0.001
Weight, kg			
Model 1: Weight	441	-0.532	<0.001
Model 2: Weight + age	441	-0.467	<0.001
Model 3: Fully adjusted[**]	438	-0.428	<0.001

[*]Dense area was square root transformed due to non-normality; percent density was modeled in the natural scale
[†]Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit
[**]Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

Table A.15 Random intercept regression estimates for the outcomes of dense area and percent density using multiple imputation, restricted to participants who were always premenopausal/early perimenopausal throughout follow-up[*]

	N	β	P Value
Dense breast area			
Body mass index, kg/m^2			
Model 1: BMI	183	0.016	0.32
Model 2: BMI + age	183	0.021	0.20
Model 3: Fully adjusted[†]	179	0.005	0.81
Weight, kg			
Model 1: Weight	183	0.007	0.31
Model 2: Weight + age	183	0.008	0.21
Model 3: Fully adjusted[†]	179	0.0004	0.96
Percent breast density			
Body mass index, kg/m^2			
Model 1: BMI	183	-1.510	<0.001
Model 2: BMI + age	183	-1.416	<0.001
Model 3: Fully adjusted[**]	180	-1.288	<0.001
Weight, kg			
Model 1: Weight	183	-0.551	<0.001
Model 2: Weight + age	183	-0.521	<0.001
Model 3: Fully adjusted[**]	180	-0.490	<0.001

[*]Dense area was square root transformed due to non-normality

[†]Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

[**]Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

Table A.16 Random intercept regression estimates for the outcomes of dense area and percent density using multiple imputation restricted to participants with ≥80% of mammograms within 90 days of a SWAN visit[*]

	N	β	P Value
Dense breast area			
Body mass index, kg/m^2			
Model 1: BMI	98	-0.013	0.68
Model 2: BMI + age	98	0.003	0.93
Model 3: Fully adjusted[†]	98	-0.014	0.68
Weight, kg			
Model 1: Weight	98	-0.003	0.76
Model 2: Weight + age	98	0.002	0.85
Model 3: Fully adjusted[†]	98	-0.005	0.65
Percent breast density			
Body mass index, kg/m^2			
Model 1: BMI	98	-1.511	<0.001
Model 2: BMI + age	98	-1.326	<0.001
Model 3: Fully adjusted[**]	97	-1.190	<0.001
Weight, kg			
Model 1: Weight	98	-0.504	<0.001
Model 2: Weight + age	98	-0.452	<0.001
Model 3: Fully adjusted[**]	97	-0.411	<0.001

[*]Dense area was square root transformed due to non-normality

[†]Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

[**]Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

Table A.17 Random intercept regression estimates for the outcomes of dense area and percent density using multiple imputation restricted to mammograms rated as good or excellent film quality[*]

	N	β	P Value
Dense breast area			
Body mass index, kg/m^2			
Model 1: BMI	819	-0.017	0.08
Model 2: BMI + age	819	-0.002	0.84
Model 3: Fully adjusted[†]	813	-0.012	0.29
Weight, kg			
Model 1: Weight	819	-0.006	0.10
Model 2: Weight + age	819	-0.001	0.75
Model 3: Fully adjusted[†]	813	-0.006	0.16
Percent breast density			
Body mass index, kg/m^2			
Model 1: BMI	819	-1.488	<0.001
Model 2: BMI + age	819	-1.333	<0.001
Model 3: Fully adjusted[**]	812	-1.201	<0.001
Weight, kg			
Model 1: Weight	819	-0.526	<0.001
Model 2: Weight + age	819	-0.479	<0.001
Model 3: Fully adjusted[**]	812	-0.443	<0.001

[*]Dense area was square root transformed due to non-normality

[†]Model 3 for dense breast area is adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

[**]Model 3 for percent density is adjusted for age, race/site, education, menopausal status, number of previous breast biopsies, age at menarche, age at first birth, number of births, history of oral contraceptive use at baseline, hormone use since previous visit

A.6 EXAMPLES OF INTERPOLATION AND MULTIPLE IMPUTATION OF BREAST DENSITY DATA

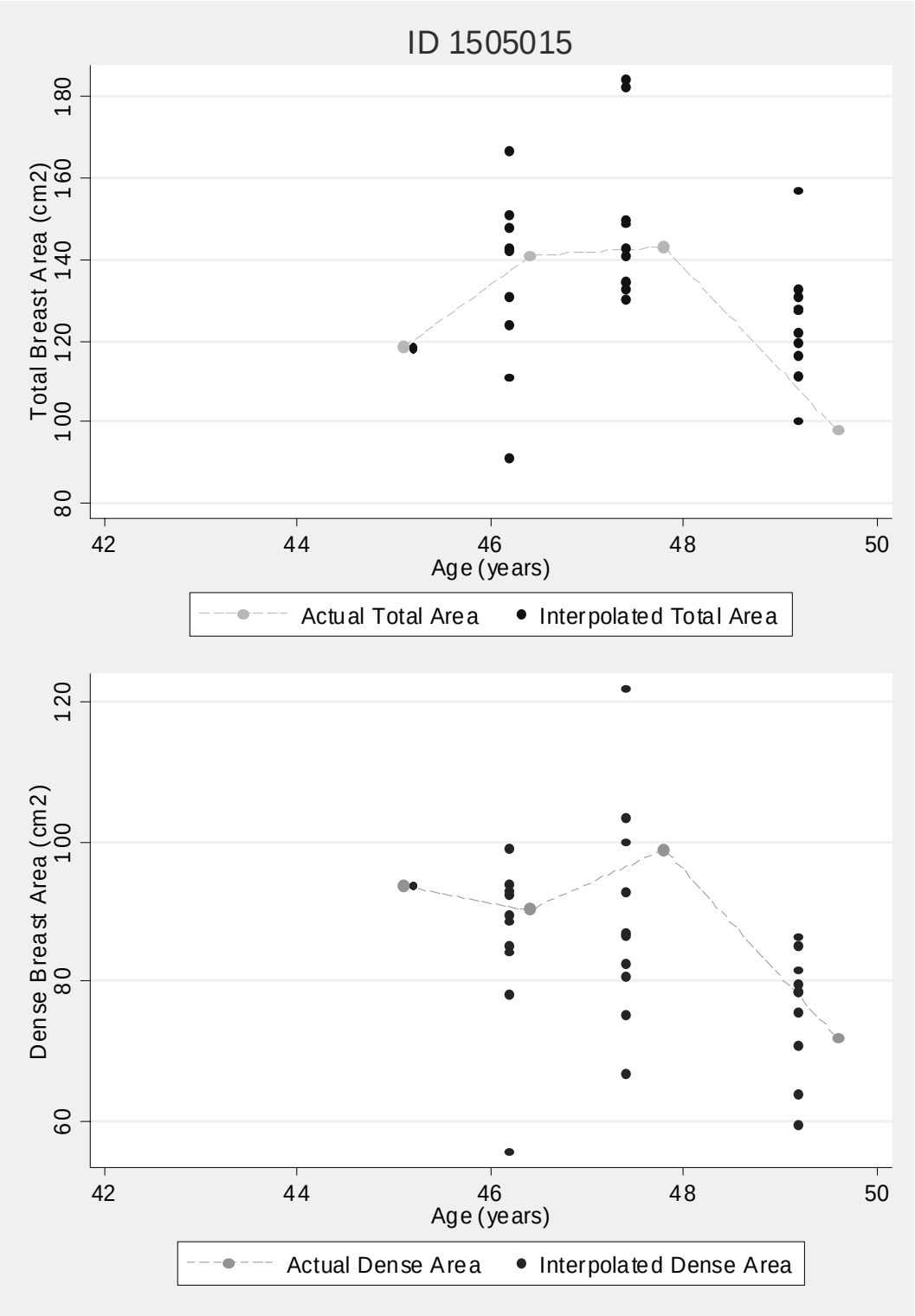


Figure A.1 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant

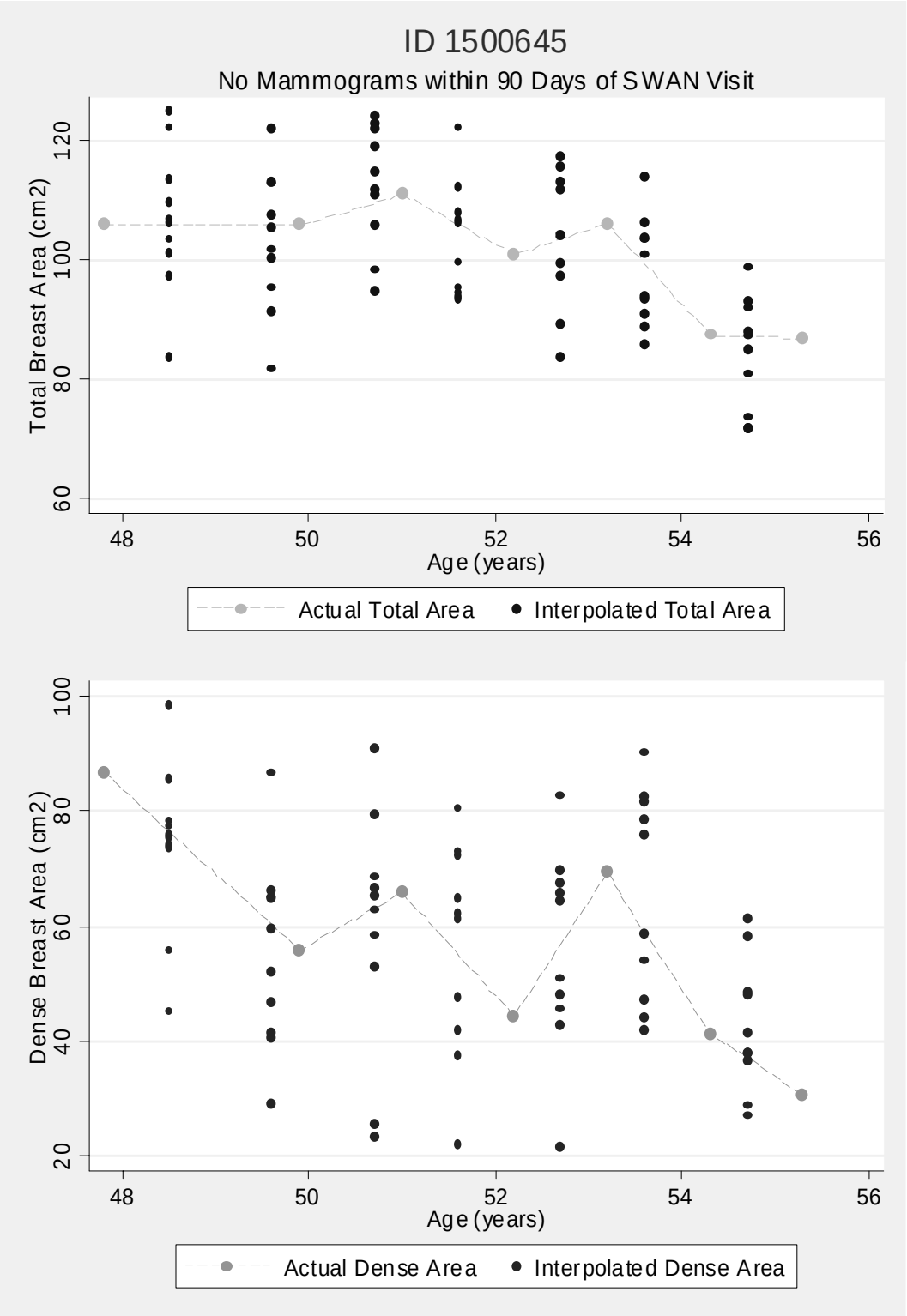


Figure A.2 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant with no mammograms within 90 days of a SWAN visit

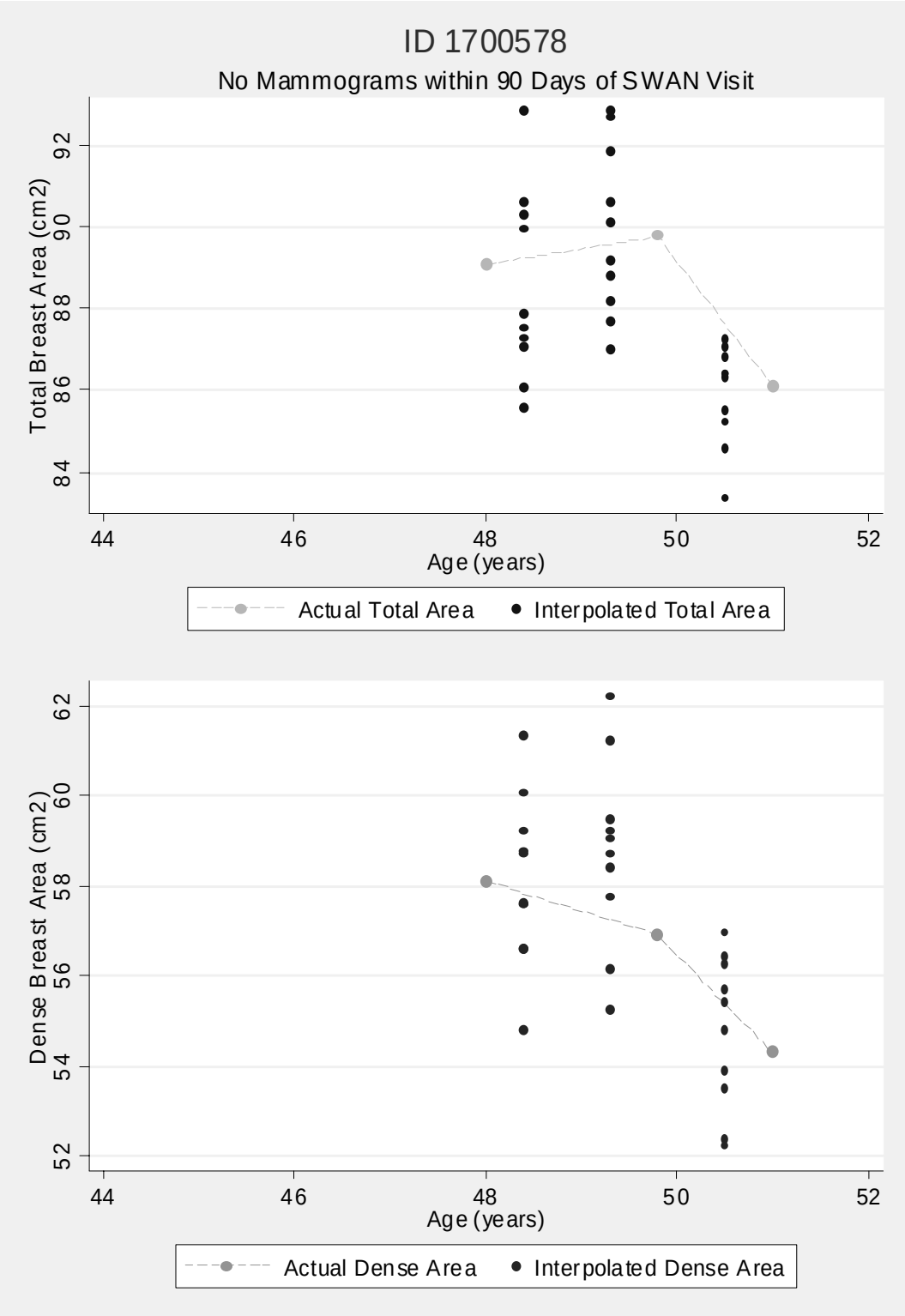


Figure A.3 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant with no mammograms within 90 days of a SWAN visit

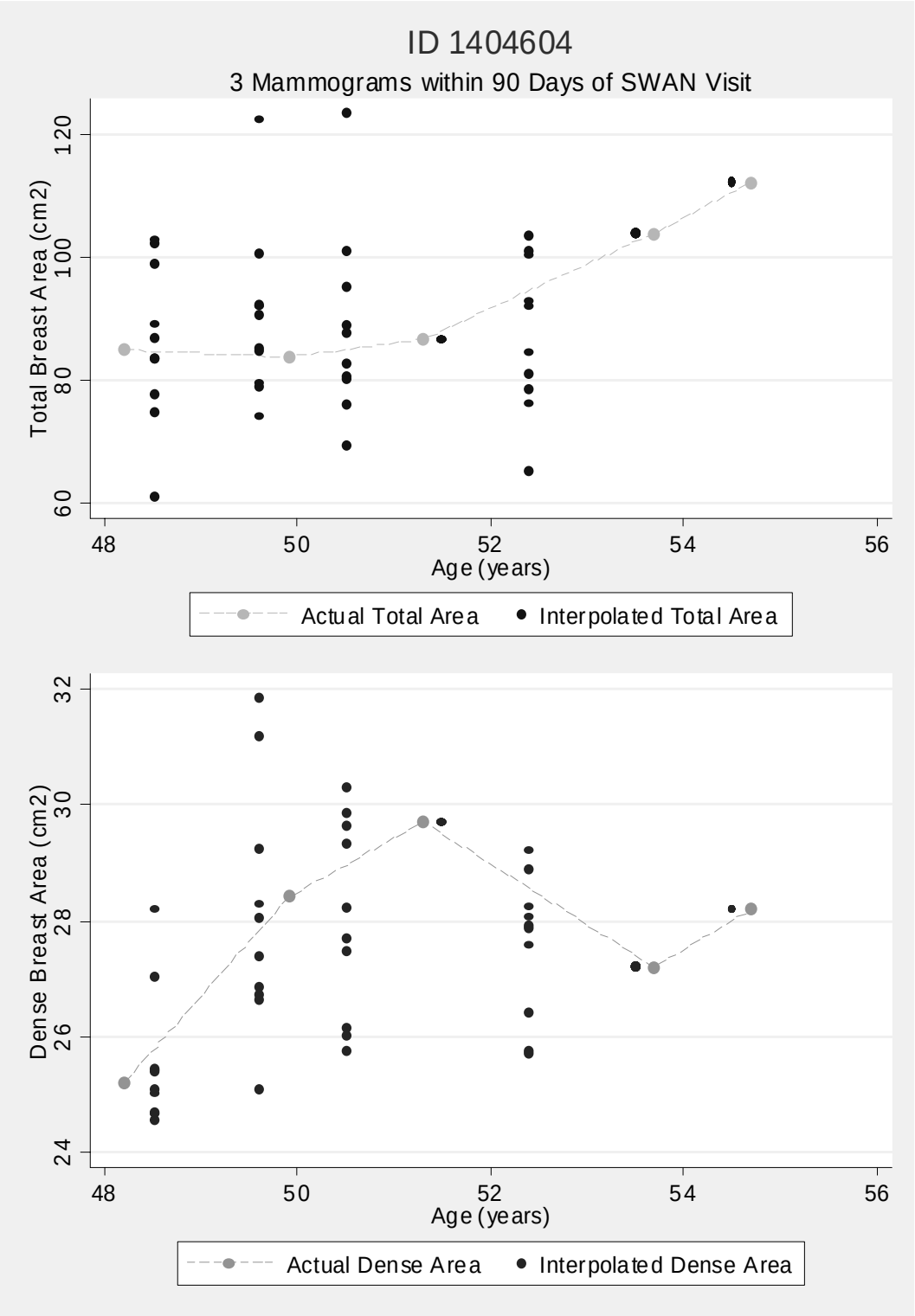


Figure A.4 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant with 3 mammograms within 90 days of a SWAN visit

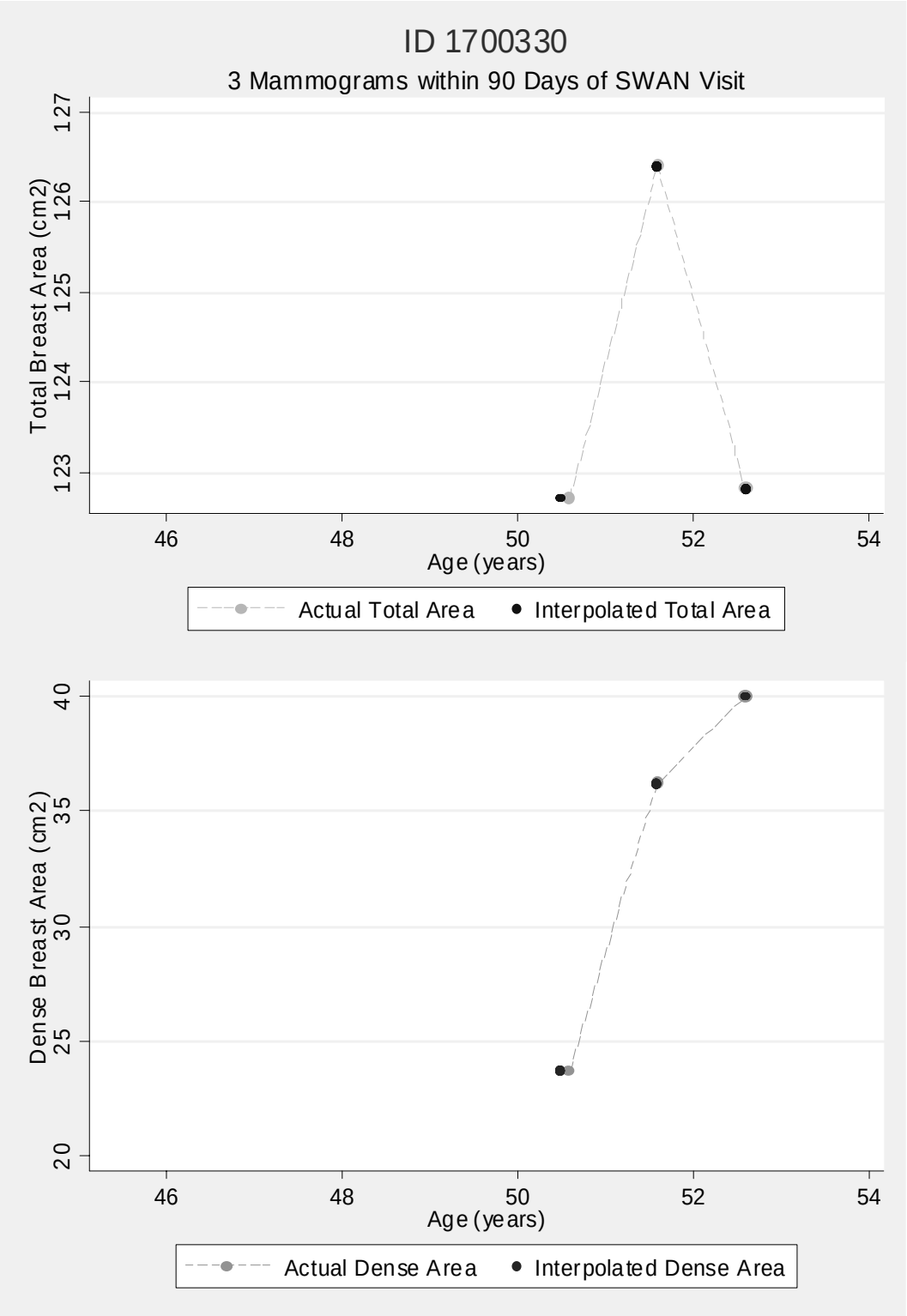


Figure A.5 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant with 3 mammograms within 90 days of a SWAN visit

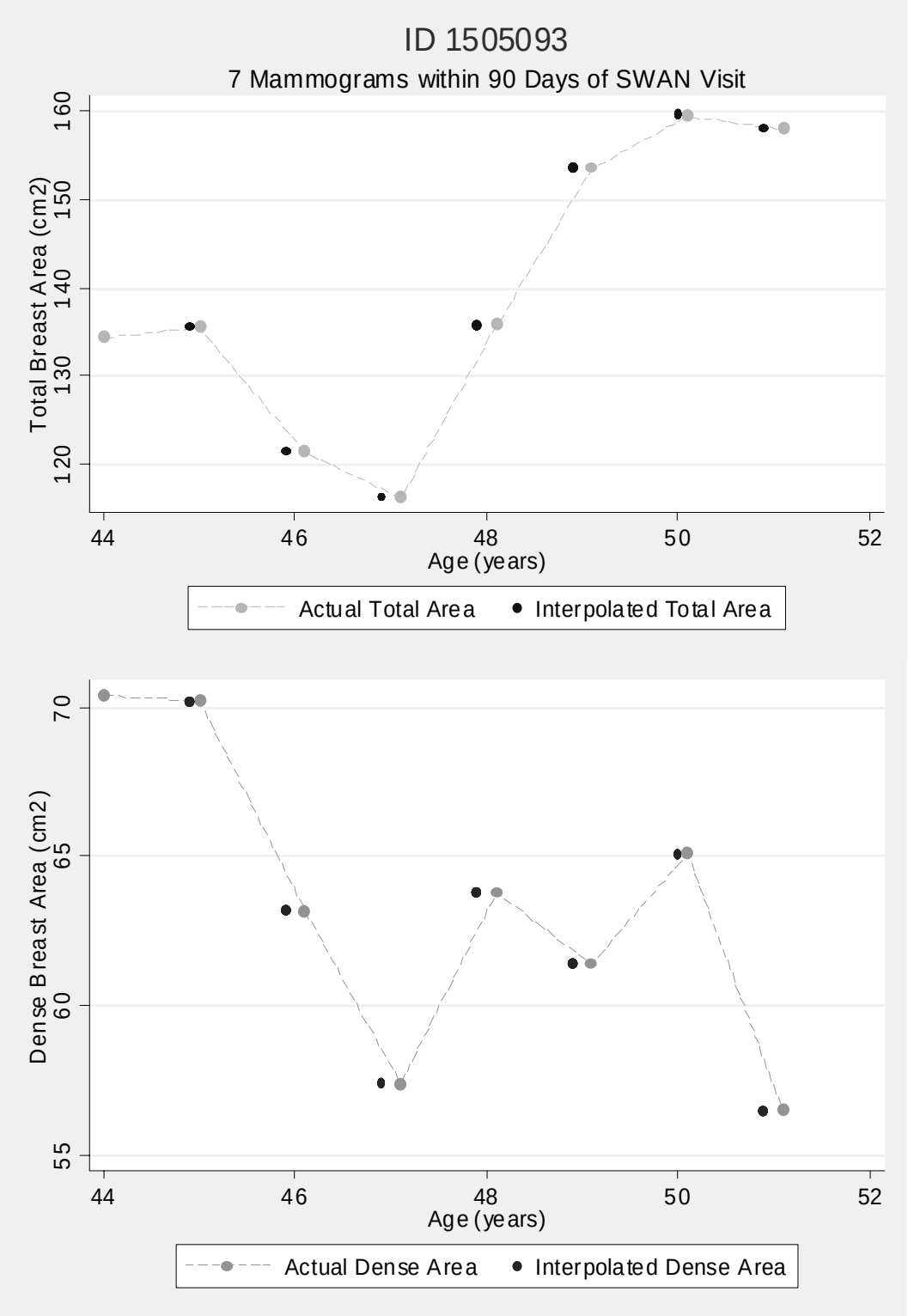


Figure A.6 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant with 7 mammograms within 90 days of a SWAN visit

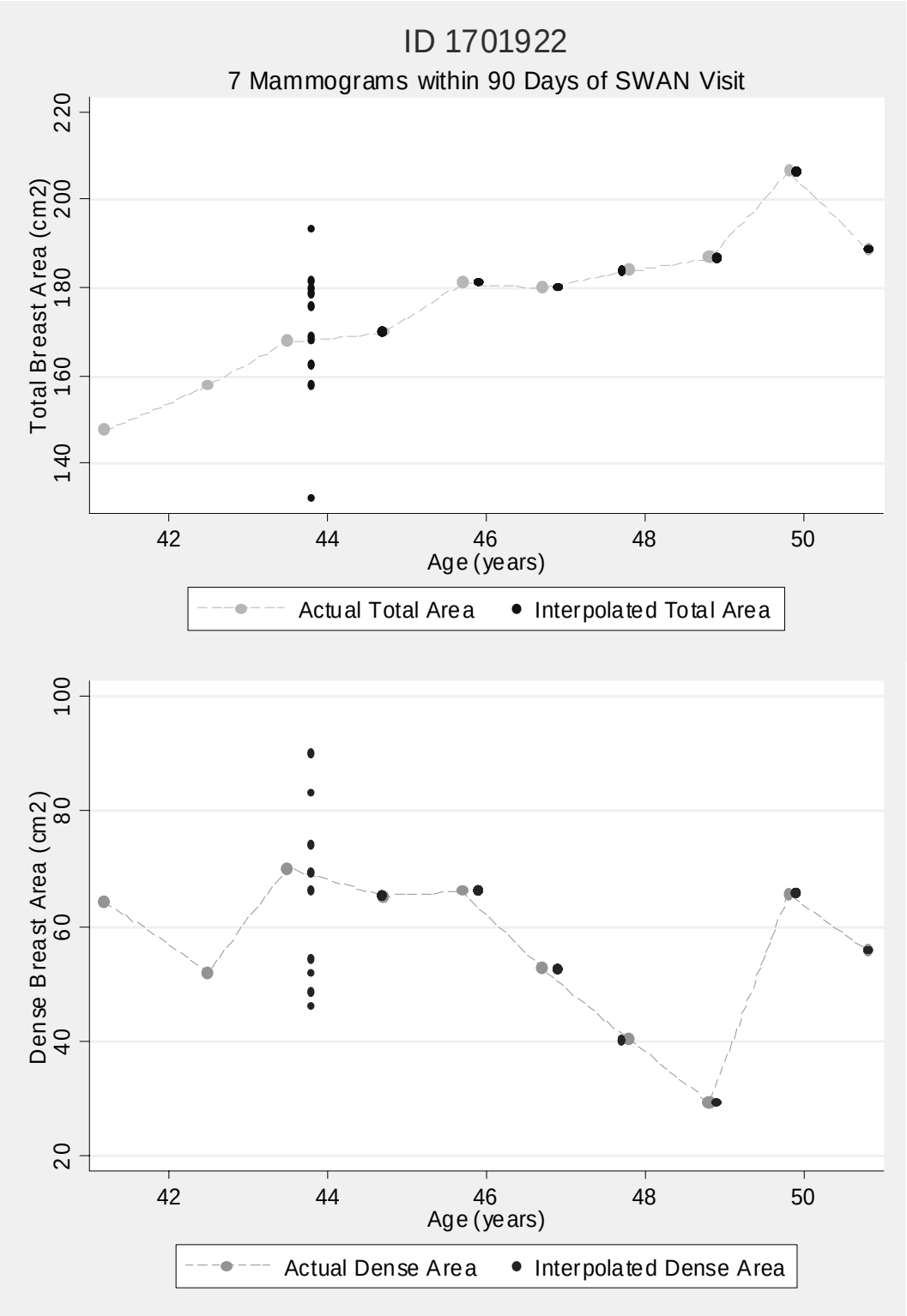
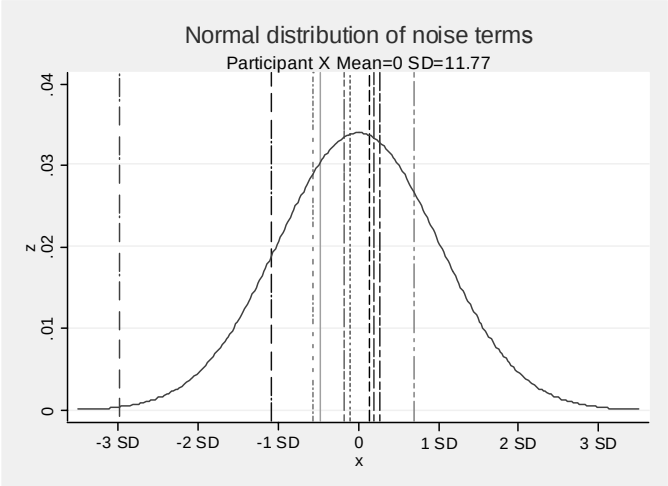


Figure A.7 Actual and interpolated total breast area (top) and dense breast area (bottom) for a randomly selected participant with 7 mammograms within 90 days of a SWAN visit

APPENDIX B

LINEAR INTERPOLATION WITH MULTIPLE IMPUTATION TO ACCOUNT FOR OFF-SCHEDULE OBSERVATIONS IN A LONGITUDINAL STUDY

B.1 ILLUSTRATIONS OF MULTIPLE IMPUTATION METHOD



Imputation	Interpolated Dense Area	Noise Term	Final Dense Area
1	90.72	1.56	92.28
2	90.72	-12.71	78.01
3	90.72	3.06	93.78
4	90.72	2.26	92.98
5	90.72	-35.24	55.48
6	90.72	-1.25	89.48
7	90.72	-2.20	88.52
8	90.72	-6.61	84.11
9	90.72	8.18	98.90
10	90.72	-5.67	85.05

Figure B.1 Selection of noise terms for multiple imputation for visit 4 for a randomly selected participant

B.2 REGRESSION RESULTS IN SUB-COHORT WITH ≥80% OF MAMMOGRAMS WITHIN 90 DAYS OF VISIT

Table B.1 Random intercept regression estimates for the outcome of dense area by estimation approach in sub-cohort

	Body Mass Index				Weight			
Approach to Estimating Dense Breast Area	**N**	**β**	**Variance**	**P value**	**N**	**β**	**Variance**	**P value**
1) Simple matching of mammograms to visits	412	-0.0185	10.9×10^{-5}	0.08	412	-0.0078	14.8×10^{-6}	0.04
2) Linear interpolation without added noise terms	411	-0.0120	15.0×10^{-5}	0.33	411	-0.0066	20.3×10^{-6}	0.14
3) Single imputation of noise terms	411	-0.0114	20.1×10^{-5}	0.42	411	-0.0060	26.5×10^{-6}	0.24
4) Multiple imputation of noise terms	411	-0.0134	26.3×10^{-5}	0.41	411	-0.0069	31.6×10^{-6}	0.22
5) Multiple imputation for all observations, regardless of time between mammogram and visit	410	-0.0167	33.3×10^{-5}	0.36	410	-0.0082	43.3×10^{-6}	0.22

*Density was modeled using a square root transformation due to non-normality; models are adjusted for age, race/site, menopausal status, 1st degree relative with history of breast cancer, number of previous breast biopsies, hormone use since previous visit

APPENDIX C

VASCULAR ENDOTHELIAL GROWTH FACTOR AND BREAST CANCER RISK

C.1 ASSOCIATIONS BETWEEN VEGF AND MEASURED HORMONES

Table C.1 Correlations between VEGF and the measured hormones, N=406[*]

	E2		SHBG		FSH		T	
	r	**P value**	**r**	**P value**	**r**	**P value**	**r**	**P value**
Total Population	-0.09	0.06	-0.07	0.15	0.01	0.77	-0.06	0.23
Control	-0.06	0.38	-0.03	0.63	0.01	0.88	-0.06	0.41
Case	-0.15	0.03	-0.12	0.10	0.01	0.86	-0.08	0.26
Premenopausal	-0.15	0.09	0.02	0.80	-0.05	0.60	-0.08	0.37
Premenopausal Control	-0.23	0.06	0.12	0.34	-0.14	0.25	-0.17	0.17
Premenopausal Case	-0.04	0.78	-0.08	0.53	0.02	0.87	-0.01	0.92
Postmenopausal	-0.02	0.76	-0.10	0.10	-0.08	0.19	-0.04	0.50
Postmenopausal Control	0.10	0.27	-0.08	0.34	-0.05	0.57	-0.01	0.93
Postmenopausal Case	-0.18	0.04	-0.12	0.16	-0.11	0.20	-0.09	0.27

[*]Correlations displayed as Pearson's correlation coefficients calculated on natural log transformed values of VEGF and hormones

C.2 SUMMARY STATISTICS FOR VEGF WITHIN MENOPAUSAL SUBGROUPS

Table C.2 Summary statistics of serum levels of VEGF, E2, FSH, SHBG, and T in the premenopausal women by case/control status, N=132[*]

	N	Mean	SD	Geometric Mean	Median	25^{th} – 75^{th} Percentiles
Controls						
VEGF, pg/mL	66	338.5	230.0	263.0	261.8	166.4 – 475.5
E2, pg/mL	66	104.5	119.1	46.2	69.8	22.1 – 141.2
FSH, mIU/mL	66	27.0	36.8	13.4	11.9	6.4 – 35.3
SHBG, nM	66	66.7	38.4	56.6	57.9	36.9 – 82.7
T, ng/dL	66	34.9	17.2	31.3	32.6	22.9 – 42.3
Cases						
VEGF, pg/mL	66	375.2	247.4	296.4	338.1	170.4 – 526.2
E2, pg/mL	66	76.6	78.6	42.4	52.8	16.0 – 106.6
FSH, mIU/mL	66	43.3	55.9	20.7	16.5	8.9 – 52.0
SHBG, nM	66	70.9	50.9	58.0	53.4	40.1 – 82.3
T, ng/dL	66	40.6	22.5	35.7	31.4	25.2 – 52.5

[*]P values from t tests comparing cases to controls on natural log transformed values: VEGF, p=0.36; E2, p=0.73; FSH, p=0.05; SHBG, p=0.82; T, p=0.12

Table C.3 Summary statistics of serum levels of VEGF, E2, FSH, SHBG, and T in the postmenopausal women by case/control status, N=275[*]

	N	Mean	SD	Geometric Mean	Median	25th – 75th Percentiles
Controls						
VEGF, pg/mL	138	410.9	317.3	306.0	321.1	192.4 – 522.7
E2, pg/mL	138	15.8	29.3	6.8	7.3	2.5 – 16.8
FSH, mIU/mL	138	123.6	53.0	109.5	121.9	83.5 – 157.9
SHBG, nM	138	54.1	30.6	46.1	47.5	33.2 – 68.2
T, ng/dL	138	32.0	19.6	26.7	27.4	17.8 – 42.0
Cases						
VEGF, pg/mL	136[†]	435.4	304.4	334.3	346.0	200.0 – 600.0
E2, pg/mL	137	17.6	24.5	9.5	11.3	5.2 – 17.1
FSH, mIU/mL	137	123.3	54.9	107.4	116.2	84.5 – 154.0
SHBG, nM	137	59.1	41.5	48.8	48.3	32.8 – 71.2
T, ng/dL	137	36.9	21.7	31.4	34.3	23.0 – 44.1

[*]P values from t tests comparing cases to controls on natural log transformed values: VEGF, p=0.37; E2, p=0.02; FSH, p=0.81; SHBG, p=0.45; T, p=0.03
[†]VEGF could not be measured in one case due to insufficient sample volume

Table C.4 Bivariate associations between serum VEGF level and personal characteristics among premenopausal controls, by VEGF level, N=66

Characteristic	VEGF <314.2 pg/mL N (%)	VEGF ≥314.2 pg/mL N (%)	P value[†]
Age, years			0.29
<50	29 (80.6)	27 (90.0)	
50-59	7 (19.4)	3 (10.0)	
Ethnicity			0.65[*]
White	34 (94.4)	27 (90.0)	
Other	2 (5.6)	3 (10.0)	
Education level			0.28[*]
High school	6 (16.7)	2 (6.7)	
Greater than high school	30 (83.3)	28 (93.3)	
Body mass index, kg/m^2			0.23
Normal, <25 kg/m^2	16 (44.4)	14 (46.7)	
Overweight, 25-<30 kg/m^2	12 (33.3)	5 (16.7)	
Obese, ≥30 kg/m^2	8 (22.2)	11 (36.7)	
Physical activity, MET h/wk			0.55[*]
0	1 (2.8)	2 (6.7)	
0.1 – <10	16 (44.4)	10 (33.3)	
≥10	19 (52.8)	18 (60.0)	
Age at menarche, years			0.37
<13	22 (61.1)	15 (50.0)	
≥13	14 (38.9)	15 (50.0)	
Number of live births			0.72[*]
None	8 (22.2)	6 (20.0)	
1	7 (19.4)	3 (10.0)	
2	13 (36.1)	13 (43.3)	
≥3	8 (22.2)	8 (26.7)	
History of breastfeeding			0.49
Not applicable	8 (22.2)	6 (20.0)	
No	10 (28.8)	5 (16.7)	
Yes	18 (50.0)	19 (63.3)	
Previous breast biopsy			0.68[*]
No	32 (88.9)	28 (93.3)	
Yes	4 (11.1)	2 (6.7)	
First degree relative with breast cancer			0.32
No	30 (83.3)	22 (73.3)	
Yes	6 (16.7)	8 (26.7)	

Table C.4 (continued)

Gail score			0.99*
<1.66%	31 (86.1)	26 (86.7)	
≥1.66%	5 (13.9)	4 (13.3)	
Hormone therapy use status			0.09*
Never	35 (97.2)	27 (90.0)	
Former	0 (0.0)	3 (10.0)	
Current (within previous 3 months)	1 (2.8)		
Oral contraceptive use status			0.18*
Never	6 (18.2)	5 (17.9)	
Former	26 (78.8)	18 (64.3)	
Current	1 (3.0)	5 (17.9)	
Serum estradiol level, pg/mL			0.36*
0.0 – <4.1	2 (5.6)	6 (20.0)	
4.1 – <11.2	2 (5.6)	1 (3.3)	
11.2 – <40.9	8 (22.2)	6 (20.0)	
≥40.9	24 (66.7)	17 (56.7)	
Serum testosterone level, ng/dL			0.10*
5.2 – <18.4	4 (11.1)	9 (30.0)	
18.4 – <29.2	8 (22.2)	6 (20.0)	
29.2 – <42.2	16 (44.4)	6 (20.0)	
≥42.2	8 (22.2)	9 (30.0)	

*P value from Fisher's exact test
Abbreviations used: VEGF, vascular endothelial growth factor; MET, metabolic equivalent

Table C.5 Bivariate associations between serum VEGF level and personal characteristics among postmenopausal controls, by VEGF level, N=138

Characteristic	VEGF <314.2 pg/mL N (%)	VEGF ≥314.2 pg/mL N (%)	P value[†]
Age, years			0.75
50-59	32 (48.5)	39 (54.2)	
60-69	21 (31.8)	19 (26.4)	
≥70	13 (19.7)	14 (19.4)	
Ethnicity			0.99[*]
White	61 (92.4)	67 (93.1)	
Other	5 (7.6)	5 (6.9)	
Education level			0.37
High school	16 (24.2)	13 (18.1)	
Greater than high school	50 (75.8)	59 (81.9)	
Body mass index, kg/m^2			0.40
Normal, <25 kg/m^2	19 (28.8)	27 (37.5)	
Overweight, 25-<30 kg/m^2	27 (40.9)	22 (30.6)	
Obese, ≥30 kg/m^2	20 (30.3)	23 (31.9)	
Physical activity, MET h/wk			0.76
0	8 (12.1)	9 (12.5)	
0.1 – <10	24 (36.4)	22 (30.6)	
≥10	34 (51.5)	41 (56.9)	
Age at menarche, years			0.93
<13	33 (50.0)	36 (50.7)	
≥13	33 (50.0)	35 (49.3)	
Age at menopause, years			0.47
<50	24 (37.5)	31 (43.7)	
≥50	40 (62.5)	40 (56.3)	
Number of live births			0.24
None	19 (55.9)	15 (44.1)	
1	5 (7.6)	13 (18.1)	
2	22 (33.3)	20 (27.8)	
≥3	20 (30.3)	24 (33.3)	
History of breastfeeding			0.23
Not applicable	19 (28.8)	15 (20.8)	
No	26 (39.4)	24 (33.3)	
Yes	21 (31.8)	33 (45.8)	
Previous breast biopsy			0.59
No	58 (87.9)	61 (84.7)	
Yes	8 (12.1)	11 (15.3)	
First degree relative with breast cancer			0.07
No	56 (84.9)	67 (94.4)	
Yes	10 (15.2)	4 (5.6)	

Table C.5 (continued)

Gail score			0.48
<1.66%	42 (63.6)	41 (57.8)	
≥1.66%	24 (36.4)	30 (42.3)	
Hormone therapy use status			0.41
Never	24 (36.4)	29 (40.3)	
Former	35 (53.0)	31 (43.1)	
Current (within previous 3 months)	7 (10.6)	12 (16.7)	
Oral contraceptive use status			0.48*
Never	30 (46.2)	25 (38.5)	
Former	35 (53.9)	39 (60.0)	
Current	0 (0.0)	1 (1.5)	
Serum estradiol level, pg/mL			0.33
0.0 – <4.1	19 (28.8)	25 (34.7)	
4.1 – <11.2	27 (40.9)	20 (27.8)	
11.2 – <40.9	17 (25.8)	20 (27.8)	
≥40.9	3 (4.6)	7 (9.7)	
Serum testosterone level, ng/dL			0.77
2.0 – <18.4	17 (25.8)	21 (29.2)	
18.4 – <29.2	16 (24.2)	21 (29.2)	
29.2 – <42.2	16 (24.2)	13 (18.1)	
≥42.2	17 (25.8)	17 (23.6)	

*P values from Fisher's exact test
Abbreviations used: VEGF, vascular endothelial growth factor; MET, metabolic equivalent

BIBLIOGRAPHY

1. American Cancer Society. *Cancer Facts & Figures 2007.* Atlanta: American Cancer Society, Inc.; 2007.

2. American Cancer Society. *Cancer Facts & Figures 2006.* Atlanta: American Cancer Society, Inc.; 2006.

3. American Cancer Society. *Breast Cancer Facts & Figures 2005-2006.* Atlanta: American Cancer Society, Inc.; 2005.

4. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 9 Regs Public-Use, Nov 2005 Sub (1973-2003), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2006, based on the November 2005 submission.

5. Ries LAG HD, Krapcho M, Mariotto A, Miller BA, Feuer EJ, Clegg L, Eisner MP, Horner MJ, Howlader N, Hayat M, Hankey BF, Edwards BK (eds). SEER Cancer Statistics Review, 1975-2003, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2003/, based on November 2005 SEER data submission, posted to the SEER web site, 2006.

6. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 1998;7(12):1133-1144.

7. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Mortality - All COD, Public-Use With State, Total U.S. for Expanded Races/Hispanics (1990-2003), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2006. Underlying mortality data provided by NCHS (www.cdc.gov/nchs).

8. 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(3):284-291.

9. Jatoi I, Becher H, Leake CR. Widening disparity in survival between white and African-American patients with breast carcinoma treated in the U. S. Department of Defense Healthcare system. *Cancer.* 2003;98(5):894-899.

10. Wojcik BE, Spinks MK, Optenberg SA. Breast carcinoma survival analysis for African American and white women in an equal-access health care system. *Cancer.* 1998;82(7):1310-1318.

11. Yood MU, Johnson CC, Blount A, et al. Race and differences in breast cancer survival in a managed care population. *J Natl Cancer Inst.* 1999;91(17):1487-1491.

12. Nanda R, Schumm LP, Cummings S, et al. Genetic testing in an ethnically diverse cohort of high-risk women: a comparative analysis of BRCA1 and BRCA2 mutations in American families of European and African ancestry. *JAMA.* 2005;294(15):1925-1933.

13. Ademuyiwa FO, Olopade OI. Racial differences in genetic factors associated with breast cancer. *Cancer Metastasis Rev.* 2003;22(1):47-53.

14. Pal T, Permuth-Wey J, Holtje T, Sutphen R. BRCA1 and BRCA2 mutations in a study of African American breast cancer patients. *Cancer Epidemiol Biomarkers Prev.* 2004;13(11 Pt 1):1794-1799.

15. Olopade OI, Fackenthal JD, Dunston G, Tainsky MA, Collins F, Whitfield-Broome C. Breast cancer genetics in African Americans. *Cancer.* 2003;97(1 Suppl):236-245.

16. *Health, United States, 2004, with Chartbook on Trends in the Health of Americans.* Hyattsville, MD: National Center for Health Statistics; 2004.

17. Hall IJ, Newman B, Millikan RC, Moorman PG. Body size and breast cancer risk in black women and white women: the Carolina Breast Cancer Study. *Am J Epidemiol.* 2000;151(8):754-764.

18. Zhu K, Caulfield J, Hunter S, Roland CL, Payne-Wilks K, Texter L. Body mass index and breast cancer risk in African American women. *Ann Epidemiol.* 2005;15(2):123-128.

19. Anderson WF, Chatterjee N, Ershler WB, Brawley OW. Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database. *Breast Cancer Res Treat.* 2002;76(1):27-36.

20. Klonoff-Cohen HS, Schaffroth LB, Edelstein SL, Molgaard C, Saltzstein SL. Breast cancer histology in Caucasians, African Americans, Hispanics, Asians, and Pacific Islanders. *Ethn Health.* 1998;3(3):189-198.

21. Kotwall CA, Brinker CC, Covington DL, Hall TL, Maxwell JG. Prognostic indices in breast cancer are related to race. *Am Surg.* 2003;69(5):372-376.

22. Porter PL, Lund MJ, Lin MG, et al. Racial differences in the expression of cell cycle-regulatory proteins in breast carcinoma. *Cancer.* 2004;100(12):2533-2542.

23. Simon MS, Severson RK. Racial differences in breast cancer survival: the interaction of socioeconomic status and tumor biology. *Am J Obstet Gynecol.* 1997;176(6):S233-239.

24. English WP, Cleveland KE, Barber WH. There is no difference in survival between African-American and white women with breast cancer. *Am Surg.* 2002;68(6):594-597.

25. Henson DE, Chu KC, Levine PH. Histologic grade, stage, and survival in breast carcinoma: comparison of African American and Caucasian women. *Cancer.* 2003;98(5):908-917.

26. Chlebowski RT, Chen Z, Anderson GL, et al. Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst.* 2005;97(6):439-448.

27. Cunningham JE, Butler WM. Racial disparities in female breast cancer in South Carolina: clinical evidence for a biological basis. *Breast Cancer Res Treat.* 2004;88(2):161-176.

28. Schairer C, Mink PJ, Carroll L, Devesa SS. Probabilities of death from breast cancer and other causes among female breast cancer patients. *J Natl Cancer Inst.* 2004;96(17):1311-1321.

29. Tammemagi CM, Nerenz D, Neslund-Dudas C, Feldkamp C, Nathanson D. Comorbidity and survival disparities among black and white patients with breast cancer. *JAMA.* 2005;294(14):1765-1772.

30. Ziegler RG, Hoover RN, Pike MC, et al. Migration patterns and breast cancer risk in Asian-American women. *J Natl Cancer Inst.* 1993;85(22):1819-1827.

31. Deapen D, Liu L, Perkins C, Bernstein L, Ross RK. Rapidly rising breast cancer incidence rates among Asian-American women. *Int J Cancer.* 2002;99(5):747-750.

32. Loman N, Johannsson O, Kristoffersson U, Olsson H, Borg A. Family history of breast and ovarian cancers and BRCA1 and BRCA2 mutations in a population-based series of early-onset breast cancer. *J Natl Cancer Inst.* 2001;93(16):1215-1223.

33. Thompson WD. Genetic epidemiology of breast cancer. *Cancer.* 1994;74(1 Suppl):279-287.

34. 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(9291):1389-1399.

35. Colditz GA, Willett WC, Hunter DJ, et al. Family history, age, and risk of breast cancer. Prospective data from the Nurses' Health Study. *JAMA.* 1993;270(3):338-343.

36. London SJ, Connolly JL, Schnitt SJ, Colditz GA. A prospective study of benign breast disease and the risk of breast cancer. *JAMA.* 1992;267(7):941-944.

37. National Cancer Institute. Genetics of Breast and Ovarian Cancer (PDQ®); Health Professional Version. Available:

http://www.cancer.gov/cancertopics/pdq/genetics/breast-and-ovarian/healthprofessional/allpages. Accessed September 18, 2005.

38. Brinton LA, Schairer C, Hoover RN, Fraumeni JF, Jr. Menstrual factors and risk of breast cancer. *Cancer Invest.* 1988;6(3):245-254.

39. Lambe M, Hsieh CC, Chan HW, Ekbom A, Trichopoulos D, Adami HO. Parity, age at first and last birth, and risk of breast cancer: a population-based study in Sweden. *Breast Cancer Res Treat.* 1996;38(3):305-311.

40. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Collaborative Group on Hormonal Factors in Breast Cancer. *Lancet.* 1996;347(9017):1713-1727.

41. Chlebowski RT, Hendrix SL, Langer RD, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial. *JAMA.* 2003;289(24):3243-3253.

42. Stefanick ML, Anderson GL, Margolis KL, et al. Effects of conjugated equine estrogens on breast cancer and mammography screening in postmenopausal women with hysterectomy. *JAMA.* 2006;295(14):1647-1657.

43. van den Brandt PA, Spiegelman D, Yaun SS, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol.* 2000;152(6):514-527.

44. Ellison RC, Zhang Y, McLennan CE, Rothman KJ. Exploring the relation of alcohol consumption to risk of breast cancer. *Am J Epidemiol.* 2001;154(8):740-747.

45. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2006;15(6):1159-1169.

46. Newcomb PA, Storer BE, Longnecker MP, et al. Lactation and a reduced risk of premenopausal breast cancer. *N Engl J Med.* 1994;330(2):81-87.

47. McTiernan A, Kooperberg C, White E, et al. Recreational physical activity and the risk of breast cancer in postmenopausal women: the Women's Health Initiative Cohort Study. *JAMA.* 2003;290(10):1331-1336.

48. Khuder SA, Mutgi AB. Breast cancer and NSAID use: a meta-analysis. *Br J Cancer.* 2001;84(9):1188-1192.

49. Hulka BS, Moorman PG. Breast cancer: hormones and other risk factors. *Maturitas.* 2001;38(1):103-113; discussion 113-106.

50. International Agency for Cancer Research (IARC). *Weight Control and Physical Activity.* Vol 6. Lyon: IARC Press; 2002.

51. Bianchini F, Kaaks R, Vainio H. Weight control and physical activity in cancer prevention. *Obes Rev.* 2002;3(1):5-8.

52. Hamajima N, Hirose K, Tajima K, et al. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Canc.* 2002;87(11):1234-1245.

53. Gail MH, Costantino JP, Bryant J, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst.* 1999;91(21):1829-1846.

54. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst.* 1999;91(18):1541-1548.

55. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81(24):1879-1886.

56. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst.* 2001;93(5):358-366.

57. Rosner B, Colditz GA, Willett WC. Reproductive risk factors in a prospective study of breast cancer: the Nurses' Health Study. *Am J Epidemiol.* 1994;139(8):819-835.

58. Kaaks R, Rinaldi S, Key TJ, et al. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. *Endocr Relat Cancer.* 2005;12(4):1071-1082.

59. Key T, Appleby P, Barnes I, Reeves G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst.* 2002;94(8):606-616.

60. Missmer SA, Eliassen AH, Barbieri RL, Hankinson SE. Endogenous estrogen, androgen, and progesterone concentrations and breast cancer risk among postmenopausal women. *J Natl Cancer Inst.* 2004;96(24):1856-1865.

61. Zeleniuch-Jacquotte A, Shore RE, Koenig KL, et al. Postmenopausal levels of oestrogen, androgen, and SHBG and breast cancer: long-term results of a prospective study. *Br J Cancer.* 2004;90(1):153-159.

62. Zaridze D, Kushlinskii N, Moore JW, Lifanova Ye, Bassalyk L, Wang DY. Endogenous plasma sex hormones in pre- and postmenopausal women with breast cancer: results from a case-control study in Moscow. *Eur J Cancer Prev.* 1992;1(3):225-230.

63. Carr BR, Bradshaw KD. Disorders of the ovary and female reproductive tract. In: Braunwald E, Fauci AS, Kasper DL, Hauser SL, Longo DL, Jameson JL, eds. *Harrison's*

Principles of Internal Medicine. 15th ed. New York, NY: McGraw-Hill; 2001:2154-2168.

64. Eliassen AH, Missmer SA, Tworoger SS, et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. *J Natl Canc Inst.* 2006;98(19):1406-1415.

65. Kabuto M, Akiba S, Stevens RG, Neriishi K, Land CE. A prospective study of estradiol and breast cancer in Japanese women. *Cancer Epidemiol Biomarkers Prev.* 2000;9(6):575-579.

66. Rosenberg CR, Pasternack BS, Shore RE, Koenig KL, Toniolo PG. Premenopausal estradiol levels and the risk of breast cancer: a new method of controlling for day of the menstrual cycle. *Am J Epidemiol.* 1994;140(6):518-525.

67. Helzlsouer KJ, Alberg AJ, Bush TL, Longcope C, Gordon GB, Comstock GW. A prospective study of endogenous hormones and breast cancer. *Cancer Detect Prev.* 1994;18(2):79-85.

68. Kaaks R, Berrino F, Key T, et al. Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC). *J Natl Cancer Inst.* 2005;97(10):755-765.

69. Key TJ. Serum oestradiol and breast cancer risk. *Endocr Relat Cancer.* 1999;6(2):175-180.

70. Sturgeon SR, Potischman N, Malone KE, et al. Serum levels of sex hormones and breast cancer risk in premenopausal women: a case-control study. *Cancer Causes Control.* 2004;15(1):45-53.

71. Wysowski DK, Comstock GW, Helsing KJ, Lau HL. Sex hormone levels in serum in relation to the development of breast cancer. *Am J Epidemiol.* 1987;125(5):791-799.

72. Yu H, Shu X-O, Shi R, et al. Plasma sex steroid hormones and breast cancer risk in Chinese women. *Int J Cancer.* 2003;105(1):92-97.

73. Thomas HV, Key TJ, Allen DS, et al. A prospective study of endogenous serum hormone concentrations and breast cancer risk in premenopausal women on the island of Guernsey. *Br J Canc.* 1997;75(7):1075-1079.

74. Persson I, Weiderpass E, Bergkvist L, Bergstrom R, Schairer C. Risks of breast and endometrial cancer after estrogen and estrogen-progestin replacement. *Cancer Causes Control.* 1999;10(4):253-260.

75. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA.* 2002;288(3):321-333.

76. Anderson GL, Limacher M, Assaf AR, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA.* 2004;291(14):1701-1712.

77. Cui Y, Whiteman MK, Langenberg P, et al. Can obesity explain the racial difference in stage of breast cancer at diagnosis between black and white women? *J Womens Health Gend Based Med.* 2002;11(6):527-536.

78. Jones BA, Kasi SV, Curnen MG, Owens PH, Dubrow R. Severe obesity as an explanatory factor for the black/white difference in stage at diagnosis of breast cancer. *Am J Epidemiol.* 1997;146(5):394-404.

79. Obesity and Overweight. Fact Sheet No. 311, September 2006. http://www.who.int/mediacentre/factsheets/fs311/en/index.html. Accessed January 25, 2007.

80. Bruning PF, Bonfrer JM, Hart AA, et al. Body measurements, estrogen availability and the risk of human breast cancer: a case-control study. *Int J Cancer.* 1992;51(1):14-19.

81. den Tonkelaar I, Seidell JC, Collette HJ, de Waard F. A prospective study on obesity and subcutaneous fat patterning in relation to breast cancer in post-menopausal women participating in the DOM project. *Br J Canc.* 1994;69(2):352-357.

82. Kolonel LN, Nomura AM, Lee J, Hirohata T. Anthropometric indicators of breast cancer risk in postmenopausal women in Hawaii. *Nutr Cancer.* 1986;8(4):247-256.

83. Manjer J, Kaaks R, Riboli E, Berglund G. Risk of breast cancer in relation to anthropometry, blood pressure, blood lipids and glucose metabolism: a prospective study within the Malmo Preventive Project. *Eur J Cancer Prev.* 2001;10(1):33-42.

84. London SJ, Colditz GA, Stampfer MJ, Willett WC, Rosner B, Speizer FE. Prospective study of relative weight, height, and risk of breast cancer. *JAMA.* 1989;262(20):2853-2858.

85. Tehard B, Clavel-Chapelon F. Several anthropometric measurements and breast cancer risk: results of the E3N cohort study. *Int J Obes.* 2006;30(1):156-163.

86. Chow LWC, Lui KL, Chan JCY, et al. Association between body mass index and risk of formation of breast cancer in Chinese women. *Asian J Surg.* 2005;28(3):179-184.

87. Chu SY, Lee NC, Wingo PA, Senie RT, Greenberg RS, Peterson HB. The relationship between body mass and breast cancer among women enrolled in the Cancer and Steroid Hormone Study. *J Clin Epidemiol.* 1991;44(11):1197-1206.

88. den Tonkelaar I, Seidell JC, Collette HJ, de Waard F. Obesity and subcutaneous fat patterning in relation to breast cancer in postmenopausal women participating in the Diagnostic Investigation of Mammary Cancer Project. *Cancer.* 1992;69(11):2663-2667.

89. Hsieh CC, Trichopoulos D, Katsouyanni K, Yuasa S. Age at menarche, age at menopause, height and obesity as risk factors for breast cancer: associations and interactions in an international case-control study. *Int J Cancer.* 1990;46(5):796-800.

90. Key TJ, Appleby PN, Reeves GK, et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *J Natl Canc Inst.* 2003;95(16):1218-1226.

91. La Vecchia C, Negri E, Franceschi S, et al. Body mass index and post-menopausal breast cancer: an age-specific analysis. *Br J Canc.* 1997;75(3):441-444.

92. Lahmann PH, Lissner L, Gullberg B, Olsson H, Berglund G. A prospective study of adiposity and postmenopausal breast cancer risk: the Malmo Diet and Cancer Study. *Int J Cancer.* 2003;103(2):246-252.

93. Pathak DR, Whittemore AS. Combined effects of body size, parity, and menstrual events on breast cancer incidence in seven countries. *Am J Epidemiol.* 1992;135(2):153-168.

94. Sweeney C, Blair CK, Anderson KE, Lazovich D, Folsom AR. Risk factors for breast cancer in elderly women. *Am J Epidemiol.* 2004;160(9):868-875.

95. Trentham-Dietz A, Newcomb PA, Egan KM, et al. Weight change and risk of postmenopausal breast cancer. *Cancer Causes Control.* 2000;11(6):533-542.

96. Yoo K, Tajima K, Park S, et al. Postmenopausal obesity as a breast cancer risk factor according to estrogen and progesterone receptor status. *Cancer Lett.* 2001;167(1):57-63.

97. Carpenter CL, Ross RK, Paganini-Hill A, Bernstein L. Effect of family history, obesity and exercise on breast cancer risk among postmenopausal women. *Int J Cancer.* 2003;106(1):96-102.

98. Sonnenschein E, Toniolo P, Terry MB, et al. Body fat distribution and obesity in pre- and postmenopausal breast cancer. *Int J Epidemiol.* 1999;28(6):1026-1031.

99. Rapp K, Schroeder J, Klenk J, et al. Obesity and incidence of cancer: a large cohort study of over 145,000 adults in Austria. *Br J Canc.* 2005;93(9):1062-1067.

100. Li CI, Malone KE, Daling JR. Interactions between body mass index and hormone therapy and postmenopausal breast cancer risk. *Cancer Causes Control.* 2006;17(5):695-703.

101. Modugno F, Kip KE, Cochrane B, et al. Obesity, hormone therapy, estrogen metabolism and risk of postmenopausal breast cancer. *Int J Cancer.* 2006;118(5):1292-1301.

102. Morimoto LM, White E, Chen Z, et al. Obesity, body size, and risk of postmenopausal breast cancer: the Women's Health Initiative. *Cancer Causes Control.* 2002;13(8):741-751.

103. Huang Z, Hankinson SE, Colditz GA, et al. Dual effects of weight and weight gain on breast cancer risk. *JAMA.* 1997;278(17):1407-1411.

104. Chie WC, Chen CF, Lee WC, Chen CJ. Body size and risk of pre- and post-menopausal breast cancer in Taiwan. *Anticancer Res.* 1996;16(5B):3129-3132.

105. Freni SC, Eberhardt MS, Turturro A, Hine RJ. Anthropometric measures and metabolic rate in association with risk of breast cancer. *Cancer Causes Control.* 1996;7(3):358-365.

106. Huang Z, Willett WC, Colditz GA, et al. Waist circumference, waist:hip ratio, and risk of breast cancer in the Nurses' Health Study. *Am J Epidemiol.* 1999;150(12):1316-1324.

107. 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(2):127-138.

108. Wenten M, Gilliland FD, Baumgartner K, Samet JM. Associations of weight, weight change, and body mass with breast cancer risk in Hispanic and non-Hispanic white women. *Ann Epidemiol.* 2002;12(6):435-434.

109. Radimer KL, Ballard-Barbash R, Miller JS, et al. Weight change and the risk of late-onset breast cancer in the original Framingham cohort. *Nutr Cancer.* 2004;49(1):7-13.

110. Harvie M, Howell A, Vierkant RA, et al. Association of gain and loss of weight before and after menopause with risk of postmenopausal breast cancer in the Iowa women's health study. *Cancer Epidemiol Biomarkers Prev.* 2005;14(3):656-661.

111. Parker ED, Folsom AR. Intentional weight loss and incidence of obesity-related cancers: the Iowa Women's Health Study. *Int J Obes Relat Metab Disord.* 2003;27(12):1447-1452.

112. Verla-Tebit E, Chang-Claude J. Anthropometric factors and the risk of premenopausal breast cancer in Germany. *Eur J Cancer Prev.* 2005;14(4):419-426.

113. Swanson CA, Coates RJ, Schoenberg JB, et al. Body size and breast cancer risk among women under age 45 years. *Am J Epidemiol.* 1996;143(7):698-706.

114. Peacock SL, White E, Daling JR, Voigt LF, Malone KE. Relation between obesity and breast cancer in young women. *Am J Epidemiol.* 1999;149(4):339-346.

115. Key TJ, Allen NE, Verkasalo PK, Banks E. Energy balance and cancer: the role of sex hormones. *Proc Nutr Soc.* 2001;60(1):81-89.

116. Lorincz AM, Sukumar S. Molecular links between obesity and breast cancer. *Endocr Relat Cancer.* 2006;13(2):279-292.

117. Verkasalo PK, Thomas HV, Appleby PN, Davey GK, Key TJ. Circulating levels of sex hormones and their relation to risk factors for breast cancer: a cross-sectional study in 1092 pre- and postmenopausal women. *Cancer Causes Control.* 2001;12(1):47-59.

118. Schnitt SJ. Normal Anatomy and Development. In: Hayes DF, ed. *Atlas of Breast Cancer*. 2nd ed. Philadelphia, PA: Mosby; 2000.

119. Wolfe JN. The prominent duct pattern as an indicator of cancer risk. *Oncology.* 1969;23(2):149-158.

120. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer.* 1976;37(5):2486-2492.

121. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. *Am J Roentgenol.* 1976;126(6):1130-1137.

122. Gravelle IH, Bulstrode JC, Bulbrook RD, Wang DY, Allen D, Hayward JL. A prospective study of mammographic parenchymal patterns and risk of breast cancer. *Br J Radiol.* 1986;59(701):487-491.

123. Saftlas AF, Hoover RN, Brinton LA, et al. Mammographic densities and risk of breast cancer. *Cancer.* 1991;67(11):2833-2838.

124. Gram IT, Funkhouser E, Tabar L. The Tabar classification of mammographic parenchymal patterns. *Eur J Radiol.* 1997;24(2):131-136.

125. American College of Radiology. *The American College of Radiology Breast Imaging Reporting and Data System (BI-RADS ®)*. 3rd ed. Reston, Va: American College of Radiology; 2003.

126. Toniolo P, Bleich AR, Beinart C, Koenig KL. Reproducibility of Wolfe's classification of mammographic parenchymal patterns. *Prev Med.* 1992;21(1):1-7.

127. Gram IT, Funkhouser E, Tabar L. Anthropometric indices in relation to mammographic patterns among peri-menopausal women. *Int J Cancer.* 1997;73(3):323-326.

128. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. *AJR Am J Roentgenol.* 2000;174(6):1769-1777.

129. Haiman CA, Bernstein L, Berg D, Ingles SA, Salane M, Ursin G. Genetic determinants of mammographic density. *Breast Cancer Res.* 2002;4(3):R5.

130. Wolfe JN, Saftlas AF, Salane M. Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. *Am J Roentgenol.* 1987;148(6):1087-1092.

131. Benichou J, Byrne C, Capece LA, et al. Secular stability and reliability of measurements of the percentage of dense tissue on mammograms. *Canc Detect Prev.* 2003;27(4):266-274.

132. Palomares MR, Machia JR, Lehman CD, Daling JR, McTiernan A. Mammographic density correlation with Gail model breast cancer risk estimates and component risk factors. *Cancer Epidemiol Biomarkers Prev.* 2006;15(7):1324-1330.

133. Haars G, van Noord PA, van Gils CH, Grobbee DE, Peeters PH. Measurements of breast density: no ratio for a ratio. *Cancer Epidemiol Biomarkers Prev.* 2005;14(11 Pt 1):2634-2640.

134. Chen Z, Wu AH, Gauderman WJ, et al. Does mammographic density reflect ethnic differences in breast cancer incidence rates? *Am J Epidemiol.* 2004;159(2):140-147.

135. Boyd NF, Byng JW, Jong RA, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. *J Natl Cancer Inst.* 1995;87(9):670-675.

136. Boyd NF, Greenberg C, Lockwood G, et al. Effects at two years of a low-fat, high-carbohydrate diet on radiologic features of the breast: results from a randomized trial. Canadian Diet and Breast Cancer Prevention Study Group. *J Natl Canc Inst.* 1997;89(7):488-496.

137. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med.* 2007;356(3):227-236.

138. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Phys Med Biol.* 1994;39(10):1629-1638.

139. Jakes RW, Duffy SW, Ng FC, Gao F, Ng EH. Mammographic parenchymal patterns and risk of breast cancer at and after a prevalence screen in Singaporean women. *Int J Epidemiol.* 2000;29(1):11-19.

140. Mitchell G, Antoniou AC, Warren R, et al. Mammographic density and breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Cancer Res.* 2006;66(3):1866-1872.

141. Byrne C, Schairer C, Wolfe J, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. *J Natl Canc Inst.* 1995;87(21):1622-1629.

142. Salminen TM, Saarenmaa IE, Heikkila MM, Hakama M. Risk of breast cancer and changes in mammographic parenchymal patterns over time. *Acta Oncologica.* 1998;37(6):547-551.

143. Knight JA, Martin LJ, Greenberg CV, et al. Macronutrient intake and change in mammographic density at menopause: results from a randomized trial. *Cancer Epidemiol Biomarkers Prev.* 1999;8(2):123-128.

144. Chlebowski RT, McTiernan A. Biological significance of interventions that change breast density. *J Natl Cancer Inst.* 2003;95(1):4-5.

145. Vachon CM, Pankratz VS, Scott CG, et al. Longitudinal trends in mammographic percent density and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2007;16(5):921-928.

146. Bergkvist L, Tabar L, Bergstrom R, Adami HO. Epidemiologic determinants of the mammographic parenchymal pattern. A population-based study within a mammographic screening program. *Am J Epidemiol.* 1987;126(6):1075-1081.

147. Brisson J, Sadowsky NL, Twaddle JA, Morrison AS, Cole P, Merletti F. The relation of mammographic features of the breast to breast cancer risk factors. *Am J Epidemiol.* 1982;115(3):438-443.

148. Grove JS, Goodman MJ, Gilbert FI, Jr., Mi MP. Factors associated with mammographic pattern. *Br J Radiol.* 1985;58(685):21-25.

149. Hart BL, Steinbock RT, Mettler FA, Jr., Pathak DR, Bartow SA. Age and race related changes in mammographic parenchymal patterns. *Cancer.* 1989;63(12):2537-2539.

150. Boyd N, Martin L, Stone J, Little L, Minkin S, Yaffe M. A longitudinal study of the effects of menopause on mammographic features. *Cancer Epidemiol Biomarkers Prev.* 2002;11(10 Pt 1):1048-1053.

151. Modugno F, Ngo DL, Allen GO, et al. Breast cancer risk factors and mammographic breast density in women over age 70. *Breast Cancer Res Treat.* 2006;97(2):157-166.

152. Boyd NF, Dite GS, Stone J, et al. Heritability of mammographic density, a risk factor for breast cancer. *N Engl J Med.* 2002;347(12):886-894.

153. Aiello EJ, Tworoger SS, Yasui Y, et al. Associations among circulating sex hormones, insulin-like growth factor, lipids, and mammographic density in postmenopausal women. *Cancer Epidemiol Biomarkers Prev.* 2005;14(6):1411-1417.

154. Boyd NF, Stone J, Martin LJ, et al. The association of breast mitogens with mammographic densities. *Br J Cancer.* 2002;87(8):876-882.

155. Noh JJ, Maskarinec G, Pagano I, Cheung LW, Stanczyk FZ. Mammographic densities and circulating hormones: a cross-sectional study in premenopausal women. *Breast.* 2006;15(1):20-28.

156. Greendale GA, Reboussin BA, Sie A, et al. Effects of estrogen and estrogen-progestin on mammographic parenchymal density. Postmenopausal Estrogen/Progestin Interventions (PEPI) Investigators. *Ann Intern Med.* 1999;130(4 Pt 1):262-269.

157. Greendale GA, Reboussin BA, Slone S, Wasilauskas C, Pike MC, Ursin G. Postmenopausal hormone therapy and change in mammographic density. *J Natl Cancer Inst.* 2003;95(1):30-37.

158. Persson I, Thurfjell E, Holmberg L. Effect of estrogen and estrogen-progestin replacement regimens on mammographic breast parenchymal density. *J Clin Oncol.* 1997;15(10):3201-3207.

159. Stomper PC, Van Voorhis BJ, Ravnikar VA, Meyer JE. Mammographic changes associated with postmenopausal hormone replacement therapy: a longitudinal study. *Radiology.* 1990;174(2):487-490.

160. McNicholas MM, Heneghan JP, Milner MH, Tunney T, Hourihane JB, MacErlaine DP. Pain and increased mammographic density in women receiving hormone replacement therapy: a prospective study. *Am J Roentgenol.* 1994;163(2):311-315.

161. Laya MB, Gallagher JC, Schreiman JS, Larson EB, Watson P, Weinstein L. Effect of postmenopausal hormonal replacement therapy on mammographic density and parenchymal pattern. *Radiology.* 1995;196(2):433-437.

162. Lundstrom E, Wilczek B, von Palffy Z, Soderqvist G, von Schoultz B. Mammographic breast density during hormone replacement therapy: differences according to treatment. *Am J Obstet Gynecol.* 1999;181(2):348-352.

163. Lundstrom E, Wilczek B, von Palffy Z, Soderqvist G, von Schoultz B. Mammographic breast density during hormone replacement therapy: effects of continuous combination, unopposed transdermal and low-potency estrogen regimens. *Climacteric.* 2001;4(1):42-48.

164. Sterns EE, Zee B. Mammographic density changes in perimenopausal and postmenopausal women: is effect of hormone replacement therapy predictable? *Breast Cancer Res Treat.* 2000;59(2):125-132.

165. Colacurci N, Fornaro F, De Franciscis P, Mele D, Palermo M, del Vecchio W. Effects of a short-term suspension of hormone replacement therapy on mammographic density. *Fertil Steril.* 2001;76(3):451-455.

166. Colacurci N, Fornaro F, De Franciscis P, Palermo M, del Vecchio W. Effects of different types of hormone replacement therapy on mammographic density. *Maturitas.* 2001;40(2):159-164.

167. Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S. Changes in breast density associated with initiation, discontinuation, and continuing use of hormone replacement therapy. *JAMA.* 2001;285(2):171-176.

168. Erel CT, Esen G, Seyisoglu H, et al. Mammographic density increase in women receiving different hormone replacement regimens. *Maturitas.* 2001;40(2):151-157.

169. Christodoulakos GE, Lambrinoudaki IV, Panoulis KP, et al. The effect of various regimens of hormone replacement therapy on mammographic breast density. *Maturitas.* 2003;45(2):109-118.

170. Conner P, Svane G, Azavedo E, et al. Mammographic breast density, hormones, and growth factors during continuous combined hormone therapy. *Fertil Steril.* 2004;81(6):1617-1623.

171. Junkermann H, von Holst T, Lang E, Rakov V. Influence of different HRT regimens on mammographic density. *Maturitas.* 2005;50(2):105-110.

172. McTiernan A, Martin CF, Peck JD, et al. Estrogen-plus-progestin use and mammographic density in postmenopausal women: women's health initiative randomized trial. *J Natl Cancer Inst.* 2005;97(18):1366-1376.

173. Harvey JA, Pinkerton JV, Herman CR. Short-term cessation of hormone replacement therapy and improvement of mammographic specificity. *J Natl Cancer Inst.* 1997;89(21):1623-1625.

174. Crandall C, Palla S, Reboussin BA, Ursin G, Greendale GA. Positive association between mammographic breast density and bone mineral density in the Postmenopausal Estrogen/Progestin Interventions Study. *Breast Cancer Res.* 2005;7(6):R922-R928.

175. Jeffreys M, Warren R, Gunnell D, McCarron P, Smith GD. Life course breast cancer risk factors and adult breast density. *Cancer Causes Control.* 2004;15(9):947-955.

176. Bartow SA, Pathak DR, Mettler FA, Key CR, Pike MC. Breast mammographic pattern: a concatenation of confounding and breast cancer risk factors. *Am J Epidemiol.* 1995;142(8):813-819.

177. Brisson J, Morrison AS, Kopans DB, et al. Height and weight, mammographic features of breast tissue, and breast cancer risk. *Am J Epidemiol.* 1984;119(3):371-381.

178. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer. *Cancer Causes Control.* 2000;11(7):653-662.

179. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. *Br J Cancer.* 1998;78(9):1233-1238.

180. Lam PB, Vacek PM, Geller BM, Muss HB. The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont. *Cancer.* 2000;89(2):369-375.

181. McCormack VA, dos Santos Silva I, De Stavola BL, et al. Life-course body size and perimenopausal mammographic parenchymal patterns in the MRC 1946 British birth cohort. *Br J Cancer.* 2003;89(5):852-859.

182. Ursin G, Ma H, Wu AH, et al. Mammographic density and breast cancer in three ethnic groups. *Cancer Epidemiol Biomarkers Prev.* 2003;12(4):332-338.

183. Duffy SW, Jakes RW, Ng FC, Gao F. Interaction of dense breast patterns with other breast cancer risk factors in a case-control study. *Br J Canc.* 2004;91(2):233-236.

184. Boyd NF, Martin LJ, Sun L, et al. Body size, mammographic density, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2006;15(11):2086-2092.

185. Cruz-Korchin N, Korchin L, Gonzalez-Keelan C, Climent C, Morales I. Macromastia: how much of it is fat? *Plast Reconstr Surg.* 2002;109(1):64-68.

186. Maskarinec G, Meng L, Ursin G. Ethnic differences in mammographic densities. *Int J Epidemiol.* 2001;30(5):959-965.

187. Maskarinec G, Pagano I, Lurie G, Kolonel LN. A longitudinal investigation of mammographic density: the multiethnic cohort. *Cancer Epidemiol Biomarkers Prev.* 2006;15(4):732-739.

188. Cooper R, Kaanders JH. Biological surrogate end-points in cancer trials: potential uses, benefits and pitfalls. *Eur J Cancer.* 2005;41(9):1261-1266.

189. Schatzkin A, Gail M. The promise and peril of surrogate end points in cancer research. *Nat Rev Cancer.* 2002;2(1):19-27.

190. Boyd NF, Martin LJ, Li Q, et al. Mammographic density as a surrogate marker for the effects of hormone therapy on risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2006;15(5):961-966.

191. Boyd NF, Lockwood GA, Martin LJ, Byng JW, Yaffe MJ, Tritchler DL. Mammographic density as a marker of susceptibility to breast cancer: a hypothesis. *IARC Sci Publ.* 2001;154:163-169.

192. Folkman J. Tumor angiogenesis: therapeutic implications. *N Engl J Med.* 1971;285(21):1182-1186.

193. Folkman J, Klagsbrun M. Angiogenic factors. *Science.* 1987;235(4787):442-447.

194. Folkman J. What is the evidence that tumors are angiogenesis dependent? *J Natl Cancer Inst.* 1990;82(1):4-6.

195. Kushlinskii NE, Gershtein ES. Role of vascular endothelial growth factor during breast cancer. *Bull Exp Biol Med.* 2002;133(6):521-528.

196. Sledge GW, Jr. Vascular endothelial growth factor in breast cancer: biologic and therapeutic aspects. *Semin Oncol.* 2002;29(3 Suppl 11):104-110.

197. Kurebayashi J, Otsuki T, Kunisue H, et al. Expression of vascular endothelial growth factor (VEGF) family members in breast cancer. *Jpn J Cancer Res.* 1999;90(9):977-981.

198. Ferrara N. The role of vascular endothelial growth factor in pathological angiogenesis. *Breast Cancer Res Treat.* 1995;36(2):127-137.

199. Stimpfl M, Tong D, Fasching B, et al. Vascular endothelial growth factor splice variants and their prognostic value in breast and ovarian cancer. *Clin Cancer Res.* 2002;8(7):2253-2259.

200. Scott PA, Gleadle JM, Bicknell R, Harris AL. Role of the hypoxia sensing system, acidity and reproductive hormones in the variability of vascular endothelial growth factor induction in human breast carcinoma cell lines. *Int J Cancer.* 1998;75(5):706-712.

201. Zhang HT, Scott PA, Morbidelli L, et al. The 121 amino acid isoform of vascular endothelial growth factor is more strongly tumorigenic than other splice variants in vivo. *Br J Canc.* 2000;83(1):63-68.

202. le Noble FA, Stassen FR, Hacking WJ, Struijker Boudier HA. Angiogenesis and hypertension. *J Hypertens.* 1998;16(11):1563-1572.

203. Sharp PS, Al-Mrayat M, Valabhji J, Kearney TM, Wright D. Serum levels of vascular endothelial growth factor in diabetic subjects: the relationship with blood pressure. *Diabetologia.* 1998;41(8):984-985.

204. Iacobellis G, Cipriani R, Gabriele A, Di Mario U, Morano S. High circulating vascular endothelial growth factor (VEGF) is related to a better systolic function in diabetic hypertensive patients. *Cytokine.* 2004;27(1):25-30.

205. Felmeden DC, Spencer CGC, Belgore FM, Blann AD, Beevers DG, Lip GYH. Endothelial damage and angiogenesis in hypertensive patients: relationship to cardiovascular risk factors and risk factor management. *Am J Hypertens.* 2003;16(1):11-20.

206. Hahn RG, Yin L, Ekengren J, Sandfeldt L. Vascular endothelial growth factor in serum indicates cardiovascular risk in urology patients. *Scand J Urol Nephrol.* 2006;40(2):144-148.

207. Chung NAY, Lydakis C, Belgore F, Blann AD, Lip GYH. Angiogenesis in myocardial infarction. An acute or chronic process? *Eur Heart J.* 2002;23(20):1604-1608.

208. Khurana R, Simons M, Martin JF, Zachary IC. Role of angiogenesis in cardiovascular disease: a critical appraisal. *Circulation.* 2005;112(12):1813-1824.

209. Gu J-W, Fortepiani LA, Reckelhoff JF, Adair TH, Wang J, Hall JE. Increased expression of vascular endothelial growth factor and capillary density in hearts of spontaneously hypertensive rats. *Microcirculation.* 2004;11(8):689-697.

210. Pawlak K, Pawlak D, Myliwiec M. Possible association between circulating vascular endothelial growth factor and oxidative stress markers in hemodialysis patients. *Medical Science Monitor.* 2006;12(4):CR181-185.

211. Felmeden DC, Spencer CGC, Blann AD, Beevers DG, Lip GYH. Physical activity in relation to indices of endothelial function and angiogenesis factors in hypertension: a substudy of the Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT). *J Intern Med.* 2003;253(1):81-91.

212. Patel JV, Lim HS, Nadar S, Tayebjee M, Hughes EA, Lip GY. Abnormal soluble CD40 ligand and C-reactive protein concentrations in hypertension: relationship to indices of angiogenesis. *J Hypertens.* 2006;24(1):117-121.

213. Belgore FM, Blann AD, Li-Saw-Hee FL, Beevers DG, Lip GY. Plasma levels of vascular endothelial growth factor and its soluble receptor (SFlt-1) in essential hypertension. *Am J Cardiol.* 2001;87(6):805-807.

214. Sane DC, Anton L, Brosnihan KB. Angiogenic growth factors and hypertension. *Angiogenesis.* 2004;7(3):193-201.

215. Beckman M. Pregnancy Problem Yields Insight Into Cancer Drug–Induced High Blood Pressure *J Natl Canc Inst.* 2007;99:1288-1289.

216. Ratain MJ, Eisen T, Stadler WM, et al. Phase II placebo-controlled randomized discontinuation trial of sorafenib in patients with metastatic renal cell carcinoma.[see comment]. *J Clin Oncol.* 2006;24(16):2505-2512.

217. Lichtenbeld HC, Barendsz-Janson AF, van Essen H, Struijker Boudier H, Griffioen AW, Hillen HF. Angiogenic potential of malignant and non-malignant human breast tissues in an in vivo angiogenesis model. *Int J Cancer.* 1998;77(3):455-459.

218. Bachelder RE, Wendt MA, Mercurio AM. Vascular endothelial growth factor promotes breast carcinoma invasion in an autocrine manner by regulating the chemokine receptor CXCR4. *Canc Res.* 2002;62(24):7203-7206.

219. Price DJ, Miralem T, Jiang S, Steinberg R, Avraham H. Role of vascular endothelial growth factor in the stimulation of cellular invasion and signaling of breast cancer cells. *Cell Growth Differ.* 2001;12(3):129-135.

220. Donovan D, Harmey JH, Toomey D, Osborne DH, Redmond HP, Bouchier-Hayes DJ. TGF beta-1 regulation of VEGF production by breast cancer cells. *Ann Surg Oncol.* 1997;4(8):621-627.

221. Shibata A, Nagaya T, Imai T, Funahashi H, Nakao A, Seo H. Inhibition of NF-kappaB activity decreases the VEGF mRNA expression in MDA-MB-231 breast cancer cells. *Breast Cancer Res Treat.* 2002;73(3):237-243.

222. Heffelfinger SC, Yan M, Gear RB, Schneider J, LaDow K, Warshawsky D. Inhibition of VEGFR2 prevents DMBA-induced mammary tumor formation. *Lab Invest.* 2004;84(8):989-998.

223. Wang G, Dong Z, Xu G, et al. The effect of antibody against vascular endothelial growth factor on tumor growth and metastasis. *J Cancer Res Clin Oncol.* 1998;124(11):615-620.

224. Wu Y, Hooper AT, Zhong Z, et al. The vascular endothelial growth factor receptor (VEGFR-1) supports growth and survival of human breast carcinoma. *Int J Cancer.* 2006;119(7):1519-1529.

225. Yoshiji H, Harris SR, Thorgeirsson UP. Vascular endothelial growth factor is essential for initial but not continued in vivo growth of human breast carcinoma cells. *Canc Res.* 1997;57(18):3924-3928.

226. Bachelder RE, Crago A, Chung J, et al. Vascular endothelial growth factor is an autocrine survival factor for neuropilin-expressing breast carcinoma cells. *Canc Res.* 2001;61(15):5736-5740.

227. Pidgeon GP, Barr MP, Harmey JH, Foley DA, Bouchier-Hayes DJ. Vascular endothelial growth factor (VEGF) upregulates BCL-2 and inhibits apoptosis in human and murine mammary adenocarcinoma cells. *Br J Canc.* 2001;85(2):273-278.

228. Biroccio A, Candiloro A, Mottolese M, et al. Bcl-2 overexpression and hypoxia synergistically act to modulate vascular endothelial growth factor expression and in vivo angiogenesis in a breast carcinoma line. *FASEB Journal.* 2000;14(5):652-660.

229. Landberg G, Nilsson K, Jirstrom K, et al. Cyclin I is expressed in human breast cancer and closely associated with VEGF and KDR expression. *Breast Cancer Res Treat.* 2005;89(3):313-316.

230. Seeger H, Wallwiener D, Kraemer E, Mueck AO. Comparison of possible carcinogenic estradiol metabolites: effects on proliferation, apoptosis and metastasis of human breast cancer cells. *Maturitas.* 2006;54(1):72-77.

231. Tang Y, Nakada MT, Kesavan P, et al. Extracellular matrix metalloproteinase inducer stimulates tumor angiogenesis by elevating vascular endothelial cell growth factor and matrix metalloproteinases. *Canc Res.* 2005;65(8):3193-3199.

232. Zetser A, Bashenko Y, Edovitsky E, Levy-Adam F, Vlodavsky I, Ilan N. Heparanase induces vascular endothelial growth factor expression: correlation with p38 phosphorylation levels and Src activation. *Canc Res.* 2006;66(3):1455-1463.

233. Chelouche-Lev D, Miller CP, Tellez C, Ruiz M, Bar-Eli M, Price JE. Different signalling pathways regulate VEGF and IL-8 expression in breast cancer: implications for therapy. *Eur J Cancer.* 2004;40(16):2509-2518.

234. Azenshtein E, Meshel T, Shina S, Barak N, Keydar I, Ben-Baruch A. The angiogenic factors CXCL8 and VEGF in breast cancer: regulation by an array of pro-malignancy factors. *Cancer Lett.* 2005;217(1):73-86.

235. Klos KS, Wyszomierski SL, Sun M, et al. ErbB2 increases vascular endothelial growth factor protein synthesis via activation of mammalian target of rapamycin/p70S6K leading to increased angiogenesis and spontaneous metastasis of human breast cancer cells. *Canc Res.* 2006;66(4):2028-2037.

236. Petit AM, Rak J, Hung MC, et al. Neutralizing antibodies against epidermal growth factor and ErbB-2/neu receptor tyrosine kinases down-regulate vascular endothelial growth factor production by tumor cells in vitro and in vivo: angiogenic implications for signal transduction therapy of solid tumors. *Am J Pathol.* 1997;151(6):1523-1530.

237. Kang HJ, Kim HJ, Rih J-K, et al. BRCA1 plays a role in the hypoxic response by regulating HIF-1alpha stability and by modulating vascular endothelial growth factor expression. *J Biol Chem.* 2006;281(19):13047-13056.

238. Hyder SM, Nawaz Z, Chiappetta C, Stancel GM. Identification of functional estrogen response elements in the gene coding for the potent angiogenic factor vascular endothelial growth factor. *Cancer Res.* 2000;60(12):3183-3190.

239. Guo P, Fang Q, Tao H-Q, et al. Overexpression of vascular endothelial growth factor by MCF-7 breast cancer cells promotes estrogen-independent tumor growth in vivo. *Canc Res.* 2003;63(15):4684-4691.

240. Dabrosin C, Margetts PJ, Gauldie J. Estradiol increases extracellular levels of vascular endothelial growth factor in vivo in murine mammary cancer. *Int J Cancer.* 2003;107(4):535-540.

241. Buteau-Lozano H, Ancelin M, Lardeux B, Milanini J, Perrot-Applanat M. Transcriptional regulation of vascular endothelial growth factor by estradiol and tamoxifen in breast cancer cells: a complex interplay between estrogen receptors alpha and beta. *Canc Res.* 2002;62(17):4977-4984.

242. Garvin S, Nilsson UW, Dabrosin C. Effects of oestradiol and tamoxifen on VEGF, soluble VEGFR-1, and VEGFR-2 in breast cancer and endothelial cells. *Br J Canc.* 2005;93(9):1005-1010.

243. Lee JE, Chung K-W, Han W, et al. Effect of estrogen, tamoxifen and epidermal growth factor on the transcriptional regulation of vascular endothelial growth factor in breast cancer cells. *Anticancer Res.* 2004;24(6):3961-3964.

244. Maity A, Sall W, Koch CJ, Oprysko PR, Evans SM. Low pO2 and beta-estradiol induce VEGF in MCF-7 and MCF-7-5C cells: relationship to in vivo hypoxia. *Breast Cancer Res Treat.* 2001;67(1):51-60.

245. Ruohola JK, Valve EM, Karkkainen MJ, Joukov V, Alitalo K, Harkonen PL. Vascular endothelial growth factors are differentially regulated by steroid hormones and antiestrogens in breast cancer cells. *Mol Cell Endocrinol.* 1999;149(1-2):29-40.

246. Takei H, Lee E-S, Jordan VC. In vitro regulation of vascular endothelial growth factor by estrogens and antiestrogens in estrogen-receptor positive breast cancer. *Breast Cancer.* 2002;9(1):39-42.

247. Mirkin S, Wong BC, Archer DF. Effect of 17 beta-estradiol, progesterone, synthetic progestins, tibolone, and tibolone metabolites on vascular endothelial growth factor mRNA in breast cancer cells. *Fertil Steril.* 2005;84(2):485-491.

248. Bogin L, Degani H. Hormonal regulation of VEGF in orthotopic MCF7 human breast cancer. *Canc Res.* 2002;62(7):1948-1951.

249. Elkin M, Orgel A, Kleinman HK. An angiogenic switch in breast cancer involves estrogen and soluble vascular endothelial growth factor receptor 1. *J Natl Canc Inst.* 2004;96(11):875-878.

250. Stoner M, Wormke M, Saville B, et al. Estrogen regulation of vascular endothelial growth factor gene expression in ZR-75 breast cancer cells through interaction of estrogen receptor alpha and SP proteins. *Oncogene.* 2004;23(5):1052-1063.

251. Sengupta K, Banerjee S, Saxena NK, Jonathan NB, Campbell DR, Banerjee SK. Differential expression of VEGF-A mRNA by 17beta-estradiol in breast tumor cells lacking classical ER-alpha may be mediated through a variant form of ER-alpha. *Mol Cell Biochem.* 2004;262(1-2):215-224.

252. Garvin S, Dabrosin C. Tamoxifen inhibits secretion of vascular endothelial growth factor in breast cancer in vivo. *Canc Res.* 2003;63(24):8742-8748.

253. Ramakrishnan R, Zell JA, Malave A, Rathinavelu A. Expression of vascular endothelial growth factor mRNA in GI-101A and HL-60 cell lines. *Biochem Biophys Res Commun.* 2000;270(3):709-713.

254. Hyder SM, Chiappetta C, Stancel GM. Pharmacological and endogenous progestins induce vascular endothelial growth factor expression in human breast cancer cells. *Int J Cancer.* 2001;92(4):469-473.

255. Liang Y, Hyder SM. Proliferation of endothelial and tumor epithelial cells by progestin-induced vascular endothelial growth factor from human breast cancer cells: paracrine and autocrine effects. *Endocrinology.* 2005;146(8):3632-3641.

256. Hyder SM, Stancel GM. Inhibition of progesterone-induced VEGF production in human breast cancer cells by the pure antiestrogen ICI 182,780. *Cancer Lett.* 2002;181(1):47-53.

257. Gasparini G. Prognostic value of vascular endothelial growth factor in breast cancer. *Oncologist.* 2000;5 Suppl 1:37-44.

258. Ryden L, Linderholm B, Nielsen NH, Emdin S, Jonsson P-E, Landberg G. Tumor specific VEGF-A and VEGFR2/KDR protein are co-expressed in breast cancer. *Breast Cancer Res Treat.* 2003;82(3):147-154.

259. Bando H, Weich HA, Brokelmann M, et al. Association between intratumoral free and total VEGF, soluble VEGFR-1, VEGFR-2 and prognosis in breast cancer. *Br J Canc.* 2005;92(3):553-561.

260. Nicolini A, Campani D, Miccoli P, et al. Vascular endothelial growth factor (VEGF) and other common tissue prognostic indicators in breast cancer: a case-control study. *Int J Biol Markers.* 2004;19(4):275-281.

261. Linderholm BK, Lindh B, Beckman L, et al. Prognostic correlation of basic fibroblast growth factor and vascular endothelial growth factor in 1307 primary breast cancers. *Clin Breast Cancer.* 2003;4(5):340-347.

262. Zhao J, Yan F, Ju H, Tang J, Qin J. Correlation between serum vascular endothelial growth factor and endostatin levels in patients with breast cancer. *Cancer Lett.* 2004;204(1):87-95.

263. Byrnes K, White S, Chu Q, et al. High eIF4E, VEGF, and microvessel density in stage I to III breast cancer. *Ann Surg.* 2006;243(5):684-690; discussion 691-682.

264. Ludovini V, Sidoni A, Pistola L, et al. Evaluation of the prognostic role of vascular endothelial growth factor and microvessel density in stages I and II breast cancer patients. *Breast Cancer Res Treat.* 2003;81(2):159-168.

265. Linderholm BK, Lindahl T, Holmberg L, et al. The expression of vascular endothelial growth factor correlates with mutant p53 and poor prognosis in human breast cancer. *Canc Res.* 2001;61(5):2256-2260.

266. Toi M, Bando H, Ogawa T, Muta M, Hornig C, Weich HA. Significance of vascular endothelial growth factor (VEGF)/soluble VEGF receptor-1 relationship in breast cancer. *Int J Cancer.* 2002;98(1):14-18.

267. Tsutsui S, Yasuda K, Suzuki K, Tahara K, Higashi H, Era S. Macrophage infiltration and its prognostic implications in breast cancer: the relationship with VEGF expression and microvessel density. *Oncol Rep.* 2005;14(2):425-431.

268. Konecny GE, Meng YG, Untch M, et al. Association between HER-2/neu and vascular endothelial growth factor expression predicts clinical outcome in primary breast cancer patients. *Clin Cancer Res.* 2004;10(5):1706-1716.

269. Nishimura R, Nagao K, Miyayama H, et al. Higher plasma vascular endothelial growth factor levels correlate with menopause, overexpression of p53, and recurrence of breast cancer. *Breast Cancer.* 2003;10(2):120-128.

270. Mohammed RA, Green A, El-Shikh S, Paish EC, Ellis IO, Martin SG. Prognostic significance of vascular endothelial cell growth factors -A, -C and -D in breast cancer and their relationship with angio- and lymphangiogenesis. *Br J Cancer.* 2007;96(7):1092-1100.

271. Meo S, Dittadi R, Peloso L, Gion M. The prognostic value of vascular endothelial growth factor, urokinase plasminogen activator and plasminogen activator inhibitor-1 in node-negative breast cancer. *Int J Biol Markers.* 2004;19(4):282-288.

272. Foekens JA, Peters HA, Grebenchtchikov N, et al. High tumor levels of vascular endothelial growth factor predict poor response to systemic therapy in advanced breast cancer. *Canc Res.* 2001;61(14):5407-5414.

273. Linderholm B, Tavelin B, Grankvist K, Henriksson R. Does vascular endothelial growth factor (VEGF) predict local relapse and survival in radiotherapy-treated node-negative breast cancer? *Br J Canc.* 1999;81(4):727-732.

274. Coradini D, Boracchi P, Daidone MG, et al. Contribution of vascular endothelial growth factor to the Nottingham prognostic index in node-negative breast cancer. *Br J Canc.* 2001;85(6):795-797.

275. Gasparini G, Toi M, Gion M, et al. Prognostic significance of vascular endothelial growth factor protein in node-negative breast carcinoma. *J Natl Canc Inst.* 1997;89(2):139-147.

276. Linderholm B, Tavelin B, Grankvist K, Henriksson R. Vascular endothelial growth factor is of high prognostic value in node-negative breast carcinoma. *J Clin Oncol.* 1998;16(9):3121-3128.

277. Manders P, Beex LVAM, Tjan-Heijnen VCG, et al. The prognostic value of vascular endothelial growth factor in 574 node-negative breast cancer patients who did not receive adjuvant systemic therapy. *Br J Canc.* 2002;87(7):772-778.

278. De Paola F, Granato AM, Scarpi E, et al. Vascular endothelial growth factor and prognosis in patients with node-negative breast cancer. *Int J Cancer.* 2002;98(2):228-233.

279. Linderholm B, Grankvist K, Wilking N, Johansson M, Tavelin B, Henriksson R. Correlation of vascular endothelial growth factor content with recurrences, survival, and first relapse site in primary node-positive breast carcinoma after adjuvant treatment. *J Clin Oncol.* 2000;18(7):1423-1431.

280. MacConmara M, O'Hanlon DM, Kiely MJ, Connolly Y, Jeffers M, Keane FBV. An evaluation of the prognostic significance of vascular endothelial growth factor in node positive primary breast carcinoma. *Int J Oncol.* 2002;20(4):717-721.

281. Lim S-C. Role of COX-2, VEGF and cyclin D1 in mammary infiltrating duct carcinoma. *Oncol Rep.* 2003;10(5):1241-1249.

282. Bachelot T, Ray-Coquard I, Menetrier-Caux C, Rastkha M, Duc A, Blay JY. Prognostic value of serum levels of interleukin 6 and of serum and plasma levels of vascular endothelial growth factor in hormone-refractory metastatic breast cancer patients. *Br J Canc.* 2003;88(11):1721-1726.

283. Coradini D, Biganzoli E, Pellizzaro C, et al. Vascular endothelial growth factor in node-positive breast cancer patients treated with adjuvant tamoxifen. *Br J Canc.* 2003;89(2):268-270.

284. Lu H, Shu X-O, Cui Y, et al. Association of genetic polymorphisms in the VEGF gene with breast cancer survival. *Canc Res.* 2005;65(12):5015-5019.

285. Balasubramanian SP, Cox A, Cross SS, Higham SE, Brown NJ, Reed MW. Influence of VEGF-A gene variation and protein levels in breast cancer susceptibility and severity. *Int J Cancer.* 2007.

286. Yoshiji H, Gomez DE, Shibuya M, Thorgeirsson UP. Expression of vascular endothelial growth factor, its receptor, and other angiogenic factors in human breast cancer. *Canc Res.* 1996;56(9):2013-2016.

287. Anan K, Morisaki T, Katano M, et al. Vascular endothelial growth factor and platelet-derived growth factor are potential angiogenic and metastatic factors in human breast cancer. *Surgery.* 1996;119(3):333-339.

288. Brown LF, Berse B, Jackman RW, et al. Expression of vascular permeability factor (vascular endothelial growth factor) and its receptors in breast cancer. *Hum Pathol.* 1995;26(1):86-91.

289. Anan K, Morisaki T, Katano M, et al. Assessment of c-erbB2 and vascular endothelial growth factor mRNA expression in fine-needle aspirates from early breast carcinomas: pre-operative determination of malignant potential. *Eur J Surg Oncol.* 1998;24(1):28-33.

290. Greb RR, Maier I, Wallwiener D, Kiesel L. Vascular endothelial growth factor A (VEGF-A) mRNA expression levels decrease after menopause in normal breast tissue but not in breast cancer lesions. *Br J Cancer.* 1999;81(2):225-231.

291. Guidi AJ, Schnitt SJ, Fischer L, et al. Vascular permeability factor (vascular endothelial growth factor) expression and angiogenesis in patients with ductal carcinoma in situ of the breast. *Cancer.* 1997;80(10):1945-1953.

292. Lewis JS, Landers RJ, Underwood JC, Harris AL, Lewis CE. Expression of vascular endothelial growth factor by macrophages is up-regulated in poorly vascularized areas of breast carcinomas. *J Pathol.* 2000;192(2):150-158.

293. Chhieng DC, Tabbara SO, Marley EF, Talley LI, Frost AR. Microvessel density and vascular endothelial growth factor expression in infiltrating lobular mammary carcinoma. *Breast J.* 2003;9(3):200-207.

294. Fuckar D, Dekanic A, Stifter S, et al. VEGF expression is associated with negative estrogen receptor status in patients with breast cancer. *Int J Surg Pathol.* 2006;14(1):49-55.

295. Leek RD, Hunt NC, Landers RJ, Lewis CE, Royds JA, Harris AL. Macrophage infiltration is associated with VEGF and EGFR expression in breast cancer. *J Pathol.* 2000;190(4):430-436.

296. Obermair A, Kucera E, Mayerhofer K, et al. Vascular endothelial growth factor (VEGF) in human breast cancer: correlation with disease-free survival. *Int J Cancer.* 1997;74(4):455-458.

297. Ryden L, Jirstrom K, Bendahl P-O, et al. Tumor-specific expression of vascular endothelial growth factor receptor 2 but not vascular endothelial growth factor or human epidermal growth factor receptor 2 is associated with impaired response to adjuvant tamoxifen in premenopausal breast cancer. *J Clin Oncol.* 2005;23(21):4695-4704.

298. Ryden L, Stendahl M, Jonsson H, Emdin S, Bengtsson NO, Landberg G. Tumor-specific VEGF-A and VEGFR2 in postmenopausal breast cancer patients with long-term follow-up. Implication of a link between VEGF pathway and tamoxifen response. *Breast Cancer Res Treat.* 2005;89(2):135-143.

299. Sancak B, Coskun U, Gunel N, et al. No association between serum levels of insulin-like growth factor-I, vascular endothelial growth factor, prolactin and clinicopathological characteristics of breast carcinoma after surgery. *Intern Med J.* 2004;34(6):310-315.

300. Gasparini G, Toi M, Miceli R, et al. Clinical relevance of vascular endothelial growth factor and thymidine phosphorylase in patients with node-positive breast cancer treated with either adjuvant chemotherapy or hormone therapy. *Cancer J Sci Am.* 1999;5(2):101-111.

301. Valkovic T, Dobrila F, Melato M, Sasso F, Rizzardi C, Jonjic N. Correlation between vascular endothelial growth factor, angiogenesis, and tumor-associated macrophages in invasive ductal breast carcinoma. *Virchows Archiv.* 2002;440(6):583-588.

302. Yamasaki T, Tsuda H, Imazeki N, Matsubara O. Vascular endothelial growth factor mRNA levels quantified by reverse transcription-polymerase chain reaction in microdissected breast carcinoma tissues are correlated with histological type and grade of both invasive and intraductal components. *Pathol Int.* 2005;55(5):255-263.

303. Coradini D, Pellizzaro C, Veneroni S, Ventura L, Daidone MG. Infiltrating ductal and lobular breast carcinomas are characterised by different interrelationships among markers related to angiogenesis and hormone dependence. *Br J Canc.* 2002;87(10):1105-1111.

304. Linderholm B, Lindh B, Tavelin B, Grankvist K, Henriksson R. p53 and vascular-endothelial-growth-factor (VEGF) expression predicts outcome in 833 patients with primary breast carcinoma. *Int J Cancer.* 2000;89(1):51-62.

305. Zaman K, Driscoll R, Hahn D, et al. Monitoring multiple angiogenesis-related molecules in the blood of cancer patients shows a correlation between VEGF-A and MMP-9 levels before treatment and divergent changes after surgical vs. conservative therapy. *Int J Cancer.* 2006;118(3):755-764.

306. Soufla G, Porichis F, Sourvinos G, Vassilaros S, Spandidos DA. Transcriptional deregulation of VEGF, FGF2, TGF-beta1, 2, 3 and cognate receptors in breast tumorigenesis. *Cancer Lett.* 2006;235(1):100-113.

307. Lantzsch T, Hefler L, Krause U, et al. The correlation between immunohistochemically-detected markers of angiogenesis and serum vascular endothelial growth factor in patients with breast cancer. *Anticancer Res.* 2002;22(3):1925-1928.

308. Howard EM, Lau SK, Lyles RH, et al. Correlation and expression of p53, HER-2, vascular endothelial growth factor (VEGF), and e-cadherin in a high-risk breast-cancer population. *Int J Clin Oncol.* 2004;9(3):154-160.

309. Blackwell KL, Dewhirst MW, Liotcheva V, et al. HER-2 gene amplification correlates with higher levels of angiogenesis and lower levels of hypoxia in primary breast tumors. *Clin Cancer Res.* 2004;10(12 Pt 1):4083-4088.

310. Benoy I, Salgado R, Colpaert C, Weytjens R, Vermeulen PB, Dirix LY. Serum interleukin 6, plasma VEGF, serum VEGF, and VEGF platelet load in breast cancer patients. *Clin Breast Cancer.* 2002;2(4):311-315.

311. Coskun U, Gunel N, Sancak B, et al. Significance of serum vascular endothelial growth factor, insulin-like growth factor-I levels and nitric oxide activity in breast cancer patients. *Breast.* 2003;12(2):104-110.

312. Hoar FJ, Lip GYH, Belgore F, Stonelake PS. Circulating levels of VEGF-A, VEGF-D and soluble VEGF-A receptor (sFIt-1) in human breast cancer. *Int J Biol Markers.* 2004;19(3):229-235.

313. Dirix LY, Vermeulen PB, Pawinski A, et al. Elevated levels of the angiogenic cytokines basic fibroblast growth factor and vascular endothelial growth factor in sera of cancer patients. *Br J Canc.* 1997;76(2):238-243.

314. Salven P, Perhoniemi V, Tykka H, Maenpaa H, Joensuu H. Serum VEGF levels in women with a benign breast tumor or breast cancer. *Breast Cancer Res Treat.* 1999;53(2):161-166.

315. Adams J, Carder PJ, Downey S, et al. Vascular endothelial growth factor (VEGF) in breast cancer: comparison of plasma, serum, and tissue VEGF and microvessel density and effects of tamoxifen. *Canc Res.* 2000;60(11):2898-2905.

316. Granato AM, Nanni O, Falcini F, et al. Basic fibroblast growth factor and vascular endothelial growth factor serum levels in breast cancer patients and healthy women: useful as diagnostic tools? *Breast Cancer Res.* 2004;6(1):R38-45.

317. Heer K, Kumar H, Read JR, Fox JN, Monson JR, Kerin MJ. Serum vascular endothelial growth factor in breast cancer: its relation with cancer type and estrogen receptor status. *Clin Cancer Res.* 2001;7(11):3491-3494.

318. Salven P, Manpaa H, Orpana A, Alitalo K, Joensuu H. Serum vascular endothelial growth factor is often elevated in disseminated cancer. *Clin Cancer Res.* 1997;3(5):647-651.

319. Wu Y, Saldana L, Chillar R, Vadgama JV. Plasma vascular endothelial growth factor is useful in assessing progression of breast cancer post surgery and during adjuvant treatment. *Int J Oncol.* 2002;20(3):509-516.

320. Yamamoto Y, Toi M, Kondo S, et al. Concentrations of vascular endothelial growth factor in the sera of normal controls and cancer patients. *Clin Cancer Res.* 1996;2(5):821-826.

321. Teh SH, Hill ADK, Lee AWS, et al. Raised plasma endostatin levels correlate inversely with breast cancer angiogenesis. *J Surg Res.* 2004;116(1):165-171.

322. Jacobs EJ, Feigelson HS, Bain EB, et al. Polymorphisms in the vascular endothelial growth factor gene and breast cancer in the Cancer Prevention Study II cohort. *Breast Cancer Res.* 2006;8(2):R22.

323. Krippl P, Langsenlehner U, Renner W, et al. A common 936 C/T gene polymorphism of vascular endothelial growth factor is associated with decreased breast cancer risk. *Int J Cancer.* 2003;106(4):468-471.

324. Jin Q, Hemminki K, Enquist K, et al. Vascular endothelial growth factor polymorphisms in relation to breast cancer development and prognosis. *Clin Cancer Res.* 2005;11(10):3647-3653.

325. Eppenberger-Castori S, Moore DH, Jr., Thor AD, et al. Age-associated biomarker profiles of human breast cancer. *Int J Biochem Cell Biol.* 2002;34(11):1318-1330.

326. Tarnowski B, Chudecka-Glaz A, Gorski B, Rzepka-Gorska I. Vascular endothelial growth factor (VEGF) levels and mutation of the BRCA1 gene in breast cancer patients. *Breast Cancer Res Treat.* 2004;88(3):287-288.

327. Dabrosin C. Positive correlation between estradiol and vascular endothelial growth factor but not fibroblast growth factor-2 in normal human breast tissue in vivo. *Clin Cancer Res.* 2005;11(22):8036-8041.

328. Artini PG, Monti M, Fasciani A, Tartaglia ML, D'Ambrogio G, Genazzani AR. Correlation between the amount of follicle-stimulating hormone administered and plasma and follicular fluid vascular endothelial growth factor concentrations in women undergoing in vitro fertilization. *Gynecol Endocrinol.* 1998;12(4):243-247.

329. McIlhenny C, George WD, Doughty JC. A comparison of serum and plasma levels of vascular endothelial growth factor during the menstrual cycle in healthy female volunteers. *Br J Canc.* 2002;86(11):1786-1789.

330. Malamitsi-Puchner A, Tziotis J, Tsonou A, Protonotariou E, Sarandakou A, Creatsas G. Changes in serum levels of vascular endothelial growth factor in males and females throughout life. *J Soc Gynecol Investig.* 2000;7(5):309-312.

331. Torry DS, Holt VJ, Keenan JA, Harris G, Caudle MR, Torry RJ. Vascular endothelial growth factor expression in cycling human endometrium. *Fertil Steril.* 1996;66(1):72-80.

332. Unkila-Kallio L, Vuorela-Vepsalainen P, Tiitinen A, Halmesmaki E, Ylikorkala O. No cyclicity in serum vascular endothelial growth factor during normal menstrual cycle but significant luteal phase elevation during an in vitro fertilization program. *Am J Reprod Immunol.* 2000;43(1):25-30.

333. Heer K, Kumar H, Speirs V, et al. Vascular endothelial growth factor in premenopausal women--indicator of the best time for breast cancer surgery? *Br J Canc.* 1998;78(9):1203-1207.

334. Kusumanto YH, Hospers GAP, Sluiter WJ, Dam WA, Meijer C, Mulder NH. Circulating vascular endothelial growth factor during the normal menstrual cycle. *Anticancer Res.* 2004;24(6):4237-4241.

335. Agrawal R, Conway GS, Sladkevicius P, et al. Serum vascular endothelial growth factor (VEGF) in the normal menstrual cycle: association with changes in ovarian and uterine Doppler blood flow. *Clin Endocrinol.* 1999;50(1):101-106.

336. Dabrosin C. Variability of vascular endothelial growth factor in normal human breast tissue in vivo during the menstrual cycle. *J Clin Endocrinol Metab.* 2003;88(6):2695-2698.

337. Dabrosin C. Sex steroid regulation of angiogenesis in breast tissue. *Angiogenesis.* 2005;8(2):127-136.

338. Habel LA, Capra AM, Oestreicher N, et al. Mammographic density in a multiethnic cohort. *Menopause.* 2007;14(5):891-899.