INTEGRATED MOLECULAR AND HISTOLOGIC SUBTYPES IN BREAST CANCER RISK AND SURVIVAL

Lindsay Almquist Williams

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

Chapel Hill
2017

Approved by:

Melissa Troester

Hazel Nichols

Andrew Olshan

Katherine Hoadley

Michael Love

ABSTRACT

Lindsay Almquist Williams: Integrated Molecular and Histologic Subtypes in Breast Cancer Risk and Survival
(Under the direction of Melissa Troester and Hazel Nichols)

Background: Intrinsic breast cancer subtypes and histologic subtypes have distinct risk factor profiles. Ductal carcinoma, diagnosed in 60-80% of cases, shows intrinsic subtype diversity. Lobular and mixed ductal-lobular carcinomas, each diagnosed in up to 15% of cases, are predominantly Luminal A subtype. It is unclear whether reported risk factor and survival profiles by histologic subtype will persist in analyses restricted to Luminal A subtype.

Methods: Using 4,359 invasive breast cancer cases from the Carolina Breast Cancer Study (1993-2013) we estimated tumor characteristic-histologic subtype associations and compared results to those in The Cancer Genome Atlas (TCGA). We estimated associations between reproductive risk factors and histologic subtypes. Finally, we evaluated associations between histologic subtype and survival. All association analyses were performed across all tumors regardless of molecular subtype and among tumors of Luminal A subtype (clinical and RNA-based) only.

Results: Tumor and molecular characteristics of lobular and mixed carcinomas were quantitatively different from ductal carcinomas, overall and among Luminal A subtype. Associations between histologic subtypes and tumor and genomic characteristics were similar in CBCS and TCGA. Lobular tumors were predominantly Luminal A subtype and had fewer TP53 pathway defects than ductal tumors. In the analysis of reproductive risk factors by histologic subtype, case-case analyses suggested significant differences between lobular and ductal tumors for older age (≥26 years) at first birth, lactation duration >12 months,

and oral contraceptive use. These associations among all tumors did not vary by race and were similar in direction and magnitude among Luminal A tumors only. Compared to ductal tumors, lobular and mixed tumors were inversely associated with breast cancer-specific death from 0-10 years after diagnosis and positively associated with breast cancer-specific death at great than 10 years since diagnosis, even among Luminal A tumors only.

Conclusions: We observed that histology was strongly associated with molecular tumor characteristics in the CBCS and TCGA, showing that the vast majority of lobular tumors are Luminal A subtype. Risk factor profiles and breast cancer-specific survival associations with histologic subtype among Luminal A tumors were similar in direction and magnitude to those observed among all tumors regardless of molecular subtype.

To Brendan, I would not be where I am today without your ineffable love and support.
To my mother, thank you for always encouraging me to achieve my dreams.
To my daughter, Caroline, you inspire me to be a better person every day. May you always know that you are deeply loved by so many.

ACKNOWLEDGEMENTS

I would like to thank Dr. Melissa Troester for her tremendous personal and professional support. Her passion for research inspires me to approach cancer epidemiology with creativity. I cannot thank her enough for the countless hours of input and guidance. I am so thankful for our friendship.

I am extremely grateful to my dissertation committee members. I would like to thank Dr. Hazel Nichols for her thoughtful feedback on my dissertation and other projects and for serving as a great role model; Dr. Andrew Olshan for his invaluable mentorship and always serving as a great source of support in my professional pursuits; Dr. Katherine Hoadley for her insightful feedback, expertise, and guidance on this dissertation; Dr. Michael Love for providing support and expertise on items critical to the success of this dissertation.

In the Department of Epidemiology, I would link to thank Dr. David Richardson and Dr. Charles Poole who have contributed greatly to my epidemiologic methods training.

In the Department of Medicine, I would like to thank Dr. Jen Jen Yeh for her years of personal and professional support. I feel privileged to have worked in her lab for 3 years.

I would not have made it through these years without the friendship, support, and good humor of my dear friends at UNC: Humberto Parada, Laura McGuinn, Melissa Furlong, Kristin Voltzke, Richard Moffitt, Gabby Herrera, Erin Kirk, and the Troester lab members.

I would like to thank my family and friends, particularly Christine Scheuring and Kristen Beyer, who make this life worth living!

Finally, I would like to honor Karen Jarvi and Maureen Fontaine. You are missed daily. Thank you for inspiring me to pursue a career focused on understanding cancer and

for serving as a reminder that lives are affected daily by this terrible disease that we are working so tirelessly to understand and stop.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
AJCC	American Joint Committee on Cancer
BMI	body mass index
CBCS	Carolina Breast Cancer Study
CDH1	E-cadherin
CK5/6	Cytokeratin 5/6
DNA	deoxyribonucleic acid
EGFR	Epidermal Growth Factor Receptor
EPC	expert pathology committee
E+P	estrogen plus progesterone menopausal hormone therapy
ER	estrogen receptor
FFPE	formalin fixed paraffin embedded
HER2	Human Epidermal Growth Factor Receptor 2
H&E	Hematoxylin and Eosin stain
HT	menopausal hormone therapy
HR	hazard ratio
IARC	International Agency for Research on Cancer
ICD-9	International Classification of Diseases version 9
ICD-10	International Classification of Diseases version 10
IHC	immunohistochemistry
LRT	Likelihood Ratio Test
NCI	National Cancer Institute
NDI	National Death Index
OR	odds ratio
PAM50	Prediction Analysis of Microarray 50-gene set for intrinsic breast cancer subtype
PR	progesterone receptor
PRS	pathology reports
RFD	relative frequency difference

RNA	ribonucleic acid
SEER	Surveillance, Epidemiology, and End Results Program
TCGA	The Cancer Genome Atlas
TDLU	terminal ductal lobular units
95% CIs	95% Confidence Interval

LIST OF SYMBOLS

α	alpha
&	and
~	approximately
β	beta
=	equals
>	greater than
≥	greater than or equal to
<	less than
≤	less than or equal to
#	number
/	per
%	percent
+	plus
-	to

CHAPTER 1: BACKGROUND

1.1 Overview of molecular and histologic heterogeneity of breast cancer

Studying heterogeneity in breast tumors most often refers to the well-characterized molecular and intrinsic subtype heterogeneity [1–4]. Histologic heterogeneity is less often the focus of breast cancer research and few studies simultaneously consider molecular and histologic heterogeneity in risk factor and survival analysis. Work by The Cancer Genome Atlas (TCGA) using RNA-based intrinsic subtype shows heterogeneity by histologic subtype with nearly all lobular and mixed ductal-lobular tumors classified as Luminal A subtype while ductal tumors show heterogeneity by subtype and nearly all Basal-like and HER2-enriched tumors occur among ductal tumors [5]. These findings are mirrored in two epidemiologic studies using IHC-based subtyping methods, the Carolina Breast Cancer Study (CBCS) (Phase 1) and the Nurses' Health Study (NHS) [3, 6]. Similarly, a number of clinical studies have determined IHC-based intrinsic subtype among the histologic subtypes and the resulting trends are similar to those observed in TCGA, the CBCS (Phase 1), and the NHS [3, 5, 7–16].

A vast majority of epidemiologic breast cancer risk factor analyses do not examine differences in risk by histologic subtype; yet, studies conducted where histologic subtype stratification has been performed suggest differences in breast cancer risk by histologic subtype for reproductive risk factors. Without stratification by histologic subtype, estimates of invasive breast cancer risk are biased towards estimates for ductal carcinoma, the most common subtype [17–30]. In stratified analyses, there is an increased risk for lobular versus ductal disease among women with younger age at menarche, older age at first birth, older age at menopause, and combined menopausal hormone therapy use

[6, 19, 24, 27–29, 31–36]. These mirror risk factor profiles reported for Luminal A intrinsic subtype [3, 4, 6]. There are no reports from population-based studies presenting risk factor analyses by intrinsic and histologic subtype, which would allow for a clearer assessment of differences in risk factor profiles.

Concerning survival by histologic subtype, results are inconsistent. There is some suggestion of better survival for lobular tumors relative to ductal tumors, but when stratification by ER status is performed a different pattern emerges [15, 17, 20, 26, 37–40]. When stratifying by ER/PR status and histologic subtype, the risk of breast cancer specific death appears to be highest among patients with lobular tumors that are ER-/PR- compared to ER+/PR+ tumors [22, 41]. No survival analyses have stratified by intrinsic and histologic subtype. The aim of this dissertation was to estimate the associations between molecular tumor characteristics and histologic subtype, to study risk factor profiles by histologic subtype, to examine differences in survival by histologic subtype using the population-based Carolina Breast Cancer study, and to estimate these associations overall and among Luminal A tumors only. By restricting to Luminal A subtype, we can more clearly define differences in both risk and survival by histologic subtype. The next sections discuss the biology of breast tissue, breast cancer heterogeneity defined by both molecular and histologic subtype, and risk factor and survival profiles by histologic and molecular subtype.

1.2 Biology of the breast

The breast is a dynamic organ with maturation occurring in various stages over the female life course. Breast tissue begins to differentiate into ducts and lobules after exposure to estrogen and progesterone during adolescence and early adulthood [42]. During pregnancy, the hormone prolactin stimulates development of ductal epithelium and initiates lactation postpartum [42]. Development of the ducts and lobules during pregnancy and lactation signifies peak maturation of the breast tissue, and is most relevant to breast cancer development. At this point, there are 15-20 lobes in each breast, referred to as terminal ductal lobular units (TDLUs). Each TDLU is made up of several lobules connected to

a duct, which opens through the nipple allowing milk to flow from the alveoli [42]. Following completion of lactation, a majority of TDLU epithelial cells undergo apoptosis and the remaining breast tissue is remodeled to the pre-pregnant, less differentiated state, a process termed post-lactational involution [43, 44]. The cycle of differentiation and post-lactational involution occurs with each pregnancy and lactation period.

Age-related involution, a distinct biological phenomenon from post-lactational involution, is characterized by glandular tissue loss due to decreasing epithelial content and replacement of stromal components with fat [43, 44]. Since a vast majority of breast tumors occur in the TDLUs, it follows that an increasing degree of age-related involution is associated with decreased risk of breast cancer [43–47]. In general, age-related involution starts when a woman is in her 40s and reaches completion when she is in her 70s, but variation in initiation and completion depends on her reproductive history. Most notably, there is a strong association between the number of live births and menopausal hormone therapy use and a delay in the completion of age-related involution [43, 48]. Interestingly, there is evidence that lobular breast cancer is more strongly associated with reproductive risk factors such as nulliparity [27, 29, 33, 35] and menopausal hormone therapy use [24, 28, 49, 50] than ductal carcinoma; suggesting that there could be an association with degree of age-related involution and risk of breast cancer by histologic subtype that has yet to be established.

1.3 Histologic heterogeneity of breast cancer

Breast tumors are classified by a pathologist based on several histologic features. First, tumors are classified as *in situ* or invasive. *In situ* disease is confined to the epithelial component of the breast and does not invade the basement membrane of the TDLU [51]. In contrast, invasive disease shows evidence of stromal infiltration through the TDLU basement membrane [51, 52]. Next, tumors are classified by a pathologist to determine histologic subtype, or the appearance of the tumor cells within the breast tissue. The three most common histologic subtypes of breast cancer, comprising up to 90% of tumors, are ductal,

lobular, and mixed ductal-lobular [52, 53]. Invasive ductal carcinoma represents 60-80% of cases [18, 51, 52]. Histologically, ductal carcinoma is characterized by sheets, cords or nests of epithelial cells embedded in the surrounding stromal tissue [51, 52].

Invasive lobular carcinoma is the second most common histologic subtype diagnosed in 5-15% of cases [51]. Invasive lobular carcinoma is characterized by small, round epithelial cells infiltrating the stroma in non-cohesive, single-file strands [17, 51]. It is important to note that histologic classifications of ductal or lobular are based on the appearance of tumor cells within the tissue, not the location of the tumor within the architecture of the breast.

The third most common histologic subtype of breast cancer, comprising up to 15% of cases, is mixed ductal-lobular (henceforward referred to as mixed) [5, 21]. While the exact threshold used to classify a tumor as mixed varies by pathologist, the International Agency for Research on Cancer (IARC) defines mixed tumors as having 10-49% of one histology present alongside a second histologic subtype [54].

There are a number of less common histologic subtypes that comprise approximately 10-20% of breast cancers and are each diagnosed in up to 5% of cases including: tubular, cribriform, papillary, mucinous, medullary, pleomorphic lobular, micropapillary, metaplastic, neuroendocrine, and mixed-other (containing various combinations of the listed histologic subtypes). Due to the low frequency of these less common subtypes, it is difficult to study them with respect to epidemiologic risk factors and survival [51].

1.4 Intrinsic subtype heterogeneity of breast cancer

Molecular heterogeneity of breast cancer is mostly commonly defined by immunohistochemistry (IHC)-based markers and RNA-based gene expression patterns [1–3, 55]. In the early 2000s, four main intrinsic subtypes were identified using gene expression, and later refined to: Luminal A, Luminal B, HER2-enriched, Basal-like [2, 56]. Prior to the discovery of RNA-based intrinsic subtypes, which represented a major advance in tumor classification, tumors were classified based on a limited number of IHC markers: estrogen

receptor (ER), progesterone receptor (PR), and more recently human epidermal growth factor 2 receptor (HER2) [57, 58]. However, the use of RNA-based intrinsic subtypes in clinical studies remains challenging because high throughput, multi-gene assays are not well suited for formalin-fixed paraffin embedded (FFPE) tissue samples. Eventually, IHC-based surrogates for the RNA-based intrinsic subtypes were identified using five IHC markers (ER, PR, HER2, CK5/6, EGFR) and were suitable for FFPE tissues [3]. The definitions are Luminal A (ER+ and/or PR+, HER2-), Luminal B (ER+ and/or PR+, HER2+), Basal-like [ER-, PR-, HER2-, cytokeratin 5/6+ (CK5/6), and/or epidermal growth factor receptor+ (EGFR)], HER2+ (HER2+, ER-, PR-), unclassified (ER-, PR-, HER2-, CK5/6- and EGFR-) [3]. Three-marker clinical subtypes serve as proxies to the 5-marker subtypes with Triple Negative (ER-, PR-, HER2-) tumors encompassing most Basal-like and unclassified tumors. More recently, cellular proliferation maker, Ki67, has also been incorporated into IHC-based definitions [59–61].

Luminal A is the most common intrinsic subtype at 50-80% prevalence. Studies suggest lower prevalence of Basal-like (10-30%), Luminal B (10-15%), and HER2-enriched (5-10%) tumor types [3, 4, 62, 63]. Significant variation in subtype distributions by age, race, and menopausal status, has been observed and may be etiologic in nature arising from different patterns of biological and lifestyle factors between women of different races [3, 4, 62, 63].

1.5 Epidemiologic risk factor profiles by histologic subtype of breast cancer

A small number of studies have evaluated risk factor profiles between ductal, lobular, and mixed breast tumors. The study of risk among the less common histologic subtypes remains challenging due to sample size limitations. Most risk factor associations for overall disease mirror those of ductal carcinoma since it is the most frequently diagnosed histologic subtype. However, when analyses are stratified by histologic subtype, stronger associations emerge for lobular compared to ductal carcinoma for younger age at menarche, older age at menopause, premenopausal status, combined estrogen and progesterone (E+P) hormone

therapy (HT) use, and later age at first birth. Reproductive risk factors that do not show heterogeneity in risk by histologic subtype include breastfeeding duration and parity. Mixed tumors are less frequently studied and tend to show risk factor profiles intermediate to those of ductal and lobular subtypes. Therefore, the next section will focus on the differences and similarities of associations between reproductive risk factors and ductal and lobular disease (Tables 1.1-1.6).

Early age at menarche and later age at menopause both increase risk of invasive breast cancer, likely by increasing exposure to endogenous estrogen and progesterone, which impacts epithelial cell growth and differentiation [6, 19, 31]. Lobular carcinoma is more strongly associated with younger age at menarche and older age at menopause than ductal carcinoma leading to the hypothesis that lobular carcinomas are more sensitive to hormonal exposures (Table 1.1) [19, 27–29, 33, 35, 64–68]. When controlling for age at diagnosis, the risk of lobular carcinoma is higher than ductal among premenopausal women in two out of three studies [19, 30, 69].

Table 1.1. Association between age at menarche, age at menopause, menopausal status and histologic subtype of breast cancer

	Variable Definition	Author, Year	Ductal	Lobular	Mixed
Age at Menarche	8-11 vs ≥14	Beaber, 2009	1.1 (0.7-1.5)	0.9 (0.6-1.4)	1.5 (0.9-2.3)
	12-13 vs ≥14		1.2 (0.8-1.6)	0.9 (0.6-1.3)	1.7 (1.0-3.0)
	12-13 vs <11	Li, 2006	1.1 (1.0-1.2)	0.8 (0.6-1.1)	1.2 (0.9-1.6)
	14+ vs <11		1.0 (0.9-1.1)	0.5 (0.4-0.8)	1.0 (0.7-1.4)
	≤11 vs ≥14	Li, 2007	1.2 (0.8-1.7)	2.2 (1.1-4.4)	
	12 vs ≥14		0.9 (0.6-1.2)	1.7 (0.9-3.3)	
	13 vs ≥14		0.9 (0.7-1.2)	1.5 (0.8-2.9)	
	≤11 vs ≥14	Nyante, 2008	1.2 (0.9-1.5)	1.3 (0.7-2.6)	
	12 vs ≥14		1.3 (1.1-1.7)	1.5 (0.8-2.9)	
	13 vs ≥14		1.0 (0.8-1.3)	1.4 (0.7-2.6)	
	<12 vs 12-13	Reeves, 2009	1.0 (1.0-1.1)	1.2 (1.1-1.3)	1.1 (1.0-1.3)
	14+ vs 12-13		1.0 (0.9-1.0)	0.9 (0.9-1.0)	0.9 (0.8-1.1)
	Trend per year		0.9 (0.9-1.0)	0.7 (0.6-0.8)	0.7 (0.6-1.0)
	Risk per year, ER+	CGHFBC, 2012	1.0 (1.0-1.1)	1.1 (1.1-1.1)	
	Risk per year, ER-		1.0 (1.0-1.0)	1.1(1.00-1.2)	
Menopausal Status	Post vs Pre	Reeves, 2009	0.9 (0.8-1.0)	0.6 (0.5-0.8)	1.0 (0.6-1.5)
	Post vs Pre	Nyante, 2008	0.7 (0.6-0.9)	0.8 (0.4-1.7)	
	Pre vs Post	CGHFBC, 2012	1.3 (0.03)	2.0 (0.1)	
Age at Menopause	<45 vs 50-54	Reeves, 2009	0.7 (0.7-0.9)	0.5 (0.3-0.7)	0.7 (0.4-1.4)
	45-49 vs 50-54		0.9 (0.8-0.9)	0.8 (0.6-1.9)	1.1 (0.8-1.5)
	55+ vs 50-54		1.2 (1.1-1.4)	1.0 (0.8-1.4)	1.2 (0.7-2.3)
	Trend per 5 years		1.2 (1.1-1.3)	1.2 (1.1-1.5)	1.2 (0.9-1.5)
	Trend per year	Kotsopoulos, 2010	1.0 (1.0-1.0)	1.0 (1.0-1.0)	
	48-50 vs ≤47	Li, 2006	1.1 (0.8-1.4)	1.8 (0.9-3.6)	1.9 (0.8-4.4)
	≥51 vs ≤47		1.1 (0.8-1.4)	1.4 (0.8-2.9)	1.3 (0.5-3.2)
	Risk per year, ER+	CGHFBC, 2012	1.0 (1.0-1.0)	1.1 (1.0-1.1)	
	Risk per year, ER-		1.0 (1.0-1.0)	1.0 (1.0-1.1)	

In addition to an increased risk of breast cancer due to endogenous hormone exposures, the use of combined estrogen and progesterone (E+P) hormone therapy (HT) increases the risk of invasive breast cancer [6, 70, 71]. Breast cancer incidence rates decreased in the early 2000s following the Women's Health Initiative trial of HT versus placebo, which showed a 24% increased risk of breast cancer by women on HT; however, the incidence rates have since increased suggesting the contribution of other risk factors remains important [49, 70, 72]. Among users of HT, there is evidence of modification by histologic subtype for current users, increasing duration of use, and type of therapy used with higher risk estimates among lobular compared to ductal tumors (Table 1.2) [28, 30, 49, 50, 73–76].

Table 1.2. Association between menopausal hormone therapy use and histologic subtype of breast cancer

HT Use	Author, Year	Ductal	Lobular	Mixed
Yes (no known hysterectomy) vs No	Phipps, 2010	1.2 (1.1-1.2)	1.5 (1.2-1.7)	1.3 (1.1-1.5)
Yes (hysterectomy reported) vs No		1.0 (0.9-1.0)	1.3 (1.0-1.5)	0.7 (0.6-0.9)
Former vs Never	Li, 2013	0.7 (0.6-1.0)	0.9 (0.7-1.2)	
Current E Only vs Never		0.9 (0.6-1.2)	1.6 (1.1-2.2)	
Current Combined vs Never		1.1 (0.8-1.5)	2.3 (1.7-3.2)	
E+P <5 years	Kotsopoulos, 2010	1.6 (1.4-1.8)	2.2 (1.5-3.1)	
E+P 6-10 years		1.8 (1.6-2.0)	3.1 (2.4-4.1)	
Current E+P Use	Calle, 2009	1.8 (1.5-2.0)	2.1 (1.6-2.8)	
E+P 1-<5 years		1.5 (1.2-1.8)	1.9 (1.3-2.8)	
E+P 5-9 years		1.8 (1.5-2.2)	2.2 (1.5-3.2)	
E+P >10 years		2.1 (1.7-2.5)	2.2 (1.5-3.2)	

Another reproductive risk factor associated with an increased risk of breast cancer is later age at first birth [6]. Later age at first birth, particularly 30 years of age and older, is more strongly associated with lobular and mixed carcinoma than ductal carcinoma (Table 1.3) [24, 27, 29, 33–36].

Table 1.3. Association between age at first birth and histologic subtype of breast cancer

Age at first birth	Author, Year	Ductal	Lobular	Mixed
20-24 vs ≤19	Li, 2006	1.0 (0.9-1.2)	1.9 (1.3-2.9)	1.0 (0.7-1.5)
25-29 vs ≤19		1.2 (1.0-1.4)	2.6 (1.6-4.0)	1.4 (0.9-2.2)
≥30 vs ≤19		1.1 (0.9-1.3)	1.8 (1.1-3.1)	1.1 (0.7-1.9)
20-24 vs ≤19	Beaber, 2009	0.8 (0.6-1.1)	0.9 (0.6-1.4)	1.0 (0.6-1.7)
25-29 vs ≤19		0.8 (0.5-1.2)	1.0 (0.6-1.6)	0.9 (0.5-1.5)
≥30 vs ≤19		1.4 (0.8-2.5)	1.7 (0.9-3.2)	2.1 (1.0-4.3)
20-24 vs <20	Nyante, 2008	1.2 (0.9-1.5)	0.9 (0.4-1.9)	
25-29 vs <20		1.3 (1.0-1.7)	1.5 (0.7-3.1)	
≥30 vs <20		1.6 (1.2-2.1)	2.5 (1.2-5.1)	
Nulliparous vs ≤19	Li, 2007	1.8 (1.2-2.7)	2.8 (1.1-7.0)	
20-24 vs ≤19		1.2 (0.9-1.7)	2.6 (1.2-5.8)	
25-29 vs ≤19		1.2 (0.8-1.8)	2.5 (1.1-6.1)	
30-34 vs ≤19		1.4 (0.8-2.4)	2.4 (0.8-7.2)	
≥35 vs ≤19		1.9 (0.9-4.0)	4.0 (1.0-15.6)	
Nulliparous vs <20	Newcomb, 2013	1.2 (1.1-1.3)	1.7 (1.3-2.2)	1.1 (0.8-1.5)
20-24 vs <20		1.1 (1.1-1.1)	1.3 (1.1-1.5)	1.0 (0.8-1.2)
25-29 vs <20		1.2 (1.2-1.3)	1.7 (1.4-2.0)	1.4 (1.1-1.7)
≥30 vs <20		1.3 (1.2-1.4)	2.4 (1.9-2.9)	1.6 (1.2-2.1)
Trend per 5 years		1.1 (1.1-1.1)	1.3 (1.2-1.4)	1.2 (1.1-1.3)
≥30 vs <30	Phipps, 2010	1.2 (1.2-1.3)	1.7 (1.4-2.1)	1.7 (1.5-2.1)
20-24 vs <20	Reeves, 2009	1.1 (1.0-1.1)	1.2 (1.2-1.3)	1.2 (1.1-1.4)
25-29 vs <20		1.2 (1.1-1.2)	1.5 (1.4-1.6)	1.4 (1.2-1.6)
30+ vs <20		1.3 (1.2-1.3)	1.8 (1.6-2.0)	1.7 (1.4-2.1)
Trend per 5 years		1.1 (1.1-1.1)	1.2 (1.2-1.3)	1.2 (1.1-1.3)
Continuous per year	Kotsopoulos, 2010	1.1	1.5	

Having one or more children has been associated with a decreased risk of breast cancer [6]. Supported by work in animal models, parity is thought to protect against breast cancer by causing differentiation of breast tissue, ridding the breast tissue of so-called "susceptible" cells that could lead to tumor initiation [77]. Parity modeled dichotomously and continuously shows that bearing one or more children decreases the risk of all three histologic subtypes with ductal and lobular disease showing the strongest associations (Table 1.4) [27, 29, 33, 35].

Table 1.4. Association between parity and histologic subtype of breast cancer

Parity/Number of live births	Author, Year	Ductal	Lobular	Mixed
Parous vs Nulliparous	Beaber, 2009	0.5 (0.3-0.8)	0.5 (0.3-0.7)	0.5 (0.3-0.8)
	Li, 2006	0.9 (0.8-1.0)	0.8 (0.6-1.1)	0.7 (0.5-1.0)
	Kotsopoulos, 2010	1.3 (1.1-1.4)	1.2 (0.9-1.7)	
	Nyante, 2008	0.8 (0.6-0.9)	0.8 (0.5-1.4)	
	Phipps, 2010	0.8 (0.8-0.8)	1.0 (0.8-1.2)	0.8 (0.7-0.9)
1 vs 0	Li, 2006	1.0 (0.8-1.2)	0.3 (0.2-0.6)	0.8 (0.5-1.3)
2 vs 0		0.9 (0.8-1.1)	0.5 (0.3-0.8)	0.8 (0.5-1.2)
3+ vs 0		0.8 (0.7-0.9)	0.5 (0.3-0.7)	0.6 (0.4-0.9)
2 vs 1	Beaber, 2009	0.6 (0.4-0.9)	0.7 (0.4-1.2)	0.8 (0.4-1.5)
3 vs 1		0.6 (0.4-1.0)	0.7 (0.4-1.3)	0.8 (0.4-1.4)
4+ vs 1		0.7 (0.4-1.1)	0.6 (0.3-1.0)	0.7 (0.4-1.3)
2-3 vs 1	Nyante, 2008	0.9 (0.8-1.1)	1.1 (0.7-1.9)	
4+ vs 1		0.6 (0.4-0.8)	0.3 (0.1-1.2)	
1 vs 0	Reeves, 2009	0.8 (0.9-0.9)	1.0 (0.9-1.1)	0.9 (0.8-1.2)
2 vs 0		0.8 (0.8-0.8)	1.0 (0.9-1.0)	0.8 (0.7-0.9)
3+ vs 0		0.7 (0.7-0.7)	0.8 (0.8-0.9)	0.7 (0.6-0.8)
Trend per birth		0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.8-1.0)

Longer breastfeeding duration has been found to reduce the risk of breast cancer overall [6, 78, 79]. Breastfeeding is thought protect against breast cancer by lengthening the fully-differentiated state of the breast tissue beyond pregnancy and by decreasing ovulation frequency resulting in reduced endogenous estrogen exposure [77, 80]. Increased breastfeeding duration shows a greater reduction in risk of ductal carcinoma than lobular or mixed carcinoma (Table 1.5) [27, 35].

Table 1.5. Association between breastfeeding duration and histologic subtype of breast cancer

Breastfeeding duration (months)	Author, Year	Ductal	Lobular	Mixed
1-11 vs Never, <1	Li, 2006	0.8 (0.7-0.9)	0.9 (0.6-1.2)	0.8 (0.6-1.2)
≥12 vs Never, <1		0.7 (0.6-0.8)	1.0 (0.7-1.5)	0.8 (0.6-1.2)
Ever vs Never	Beaber, 2009	0.7 (0.5-0.9)	0.9 (0.7-1.3)	0.9 (0.6-1.4)

Previous epidemiologic work has shown that longer duration and recency of oral contraceptive use increases the risk of breast cancer with the highest risk observed among current, long-term (4 or more years of use) users at the time of diagnosis [80–82]. Oral contraceptives are thought to work by increasing the levels of bioavailable hormones in the body and breast tissue, subsequently impacting cell differentiation. Results regarding current oral contraceptive use and the risk of ductal carcinoma suggest an increased risk among current users with conflicting results among current users diagnosed with lobular carcinoma (Table 1.6) [33, 83]. The risk among former and ever users is similar by subtype [33, 83].

Table 1.6. Association between oral contraceptive use and histologic subtype of breast cancer

Oral Contraceptive Use	Author, Year	Ductal	Lobular	Mixed
Ever vs Never	Newcomer, 2003	1.0 (0.9-1.1)	1.2 (0.9-1.6)	
Former vs Never		1.0 (0.9-1.1)	1.2 (0.9-1.6)	
Current vs Never		1.2 (0.8-1.9)	2.6 (1.0-7.1)	
Ever vs Never	Nyante, 2008	1.2 (1.0-1.5)	1.1 (0.7-1.8)	
Former vs Never		1.2 (1.0-1.4)	1.2 (0.7-1.9)	
Current vs Never		1.5 (1.1-2.0)	0.3 (0.1-1.5)	

Few studies have reported racial differences in histologic subtype [21, 27, 35, 37, 84–87]. Some reports suggest that black women experience a greater frequency of ductal tumors and a lower frequency of lobular tumors than white women [37, 87, 88]. For a clear picture of breast cancer risk by histologic subtype among women of different races, it is also important to account for reproductive patterns and intrinsic subtype, both of which differ by race [4].

In summary, the observed risk factor differences between ductal and lobular tumors are largely confined to reproductive and hormone-modulating exposures; however, these differences may be largely driven by intrinsic subtype. There is limited evidence of differences in breast cancer risk by histologic subtype and race. There are few studies that characterize risk factor profiles for mixed tumors, yet it remains an important public health

concern as it is the third most common histologic subtype and risk reduction strategies remain to be identified.

1.6 Comparison of risk factor profiles for intrinsic subtype and histologic subtype

This section will compare and contrast the associations between the previously described reproductive risk factors and histologic subtype to the risk factor associations observed for intrinsic subtype. These comparisons may help to obtain a clearer picture of the suspected confounding by intrinsic subtype of the reproductive risk factor-histologic subtype associations. As will be discussed in greater detail in section 1.10, previous research suggests that nearly 90% of lobular tumors are Luminal A subtype, but ductal tumors show diversity in intrinsic subtypes with nearly all Basal-like and HER2-enriched tumors classified as ductal histologically [3, 5, 7–16]. Therefore, Luminal A intrinsic subtype tumors and lobular tumors may be similar for risk factor and survivorship associations.

The associations observed for lobular tumors and younger age at menarche, older age at menopause, older age at first birth, nulliparity, and HT use mirror a recent report from the NHS on these risk factors being more strongly associated with Luminal A intrinsic subtype than HER2+, Basal-like, or unclassified tumor types [6]. Reports on breastfeeding and the risk of ductal carcinoma suggest a decreased risk of similar magnitude to that observed for Basal-like carcinomas, which comprise an estimated 25% of ductal carcinomas [4, 6, 11]. However, it is difficult to draw any firm risk factor comparisons between any of the intrinsic subtypes and ductal tumors due intrinsic subtype diversity among ductal tumors.

1.7 Clinical characteristics by histologic subtype of breast cancer

Previous research has shown that compared to ductal carcinoma, lobular carcinoma is diagnosed among women at older ages, presents as larger tumors, is more frequently low-intermediate grade, more frequently higher stage at diagnosis, and is associated with a higher rate of contralateral disease [8, 9, 11, 12, 15, 17, 18, 20–26, 37, 54, 89–102]. Using IHC, lobular and mixed tumors are more frequently ER positive, PR positive, and HER2

negative [5, 9, 11, 13–15, 17, 18, 20–26, 35, 37, 38, 89, 90, 92, 102–105]. In contrast, ductal carcinomas are more frequently of higher grade, have higher proportions of ER negative, PR negative, and HER2 positive tumors than lobular carcinomas [5, 11, 13, 17, 18, 20–26, 37, 38, 89]. Mixed tumors do not display clinical characteristics distinct from ductal or lobular tumors except for a greater frequency of lymph node metastasis [18, 21, 24, 26, 37, 95].

1.8 Survival by histologic subtype of breast cancer

There is substantial analytic variation across studies of disease specific and overall survival by histologic subtype (Table 1.7). The literature on survival by histologic subtype is largely based on clinical studies that use retrospective medical records abstraction, and may not be population-based or racially diverse. Overall survival at 5, 10, and greater than 10-years tends to be fairly similar between the histologic subtypes [15, 17, 20, 26, 37–40, 95]. The risk of death from any cause is higher among patients diagnosed with lobular tumors, particularly in studies with long-term follow up [12, 23, 106, 107]. One clinical trial reported that <10 years after diagnosis, patients with lobular tumors had a reduced risk of death from any cause relative to patients with ductal carcinoma; however, the trend was reversed after 10 years, and risk of death was highest among ER-, lobular tumors, suggesting that the risk of death may vary by histologic subtype, molecular subtype, and time [23].

The findings for risk of breast cancer-specific death are complex and hard to compare across studies as methods for stratification by histologic and molecular subtype as well as age vary greatly. Overall, studies have shown a reduced risk of breast cancer-specific death for both ductal and lobular tumors compared to mixed tumors and a reduced risk of breast cancer specific death for lobular versus ductal tumors [18, 108]. When Li (2010) adjusted for ER and PR status, lobular tumors showed a reduced risk of breast cancer-specific death [22]. When stratifying by ER status and age, the risk of breast cancer-specific death for both lobular and mixed tumors, compared to ductal tumors, among women aged 30-49 years was highest among ER-/PR- tumors and was reversed among

women age 50 and older [22]. Dunnwald *et al.* (2007) observed little variation in the risk of death by histologic subtype when comparing tumors that were ER+/PR- or ER-/PR- to ER+/PR+, suggesting that molecular subtype rather than histologic subtype may be a stronger risk factor for breast cancer specific death [41]. There are no survival analyses accounting for intrinsic subtype and histologic subtype from population-based, racially diverse study.

Table 1.7. Association between histologic subtype and survival

Variable Definition*	Author, Year	Ductal (D)	Lobular (L)	Mixed (M)
Overall				
5 year (%)	Arpino, 2004	84 (83-85)	86 (84-87)	
	Bharat, 2009	80	87	84
	Cristofanilli, 2005	70	90	
	Korhonen, 2004	77	80	
	Warren, 2013	89	90	92
	Jung, 2010	93	94	
	Zengel, 2015	88	82	92
10 year (%)	Bharat, 2009	61	68	69
	Ellis, 1992	47	54	40
	Korhonen, 2004	50	59	
	Zengel, 2015	71	64	83
> 10 years (%)	Suryadevara, 2010	39		46
L vs D	Garcia-Fernandez, 2015		1.2 (0.8-1.8)	
L vs D	Korhonen, 2013		1.6 (1.2-2.2)	
L vs D	Boyle, 2014		0.8 (0.2-3.2)	
M vs D				0.9 (0.3-3.1)
≤10 years; L vs D	Pestalozzi, 2008		0.8 (0.7-1.0)	
≤10 years ER+; L vs D			0.9 (0.8-1.1)	
≤10 years ER- ; L vs D			0.7 (0.6-1.0)	
>10 years; L vs D			1.5 (1.2-1.9)	
>10 years ER+; L vs D			1.3 (1.0-1.6)	
>10 years ER-; L vs D			2.7 (1.6-4.5)	
Breast Cancer-Specific				
D vs M	Arps, 2013	0.6 (0.3-1.0)		
L vs M			0.7 (0.4-1.4)	
L vs D	Campbell, 2015		0.7 (0.6-1.0)	
L vs D	Garcia-Fernandez, 2015		1.2 (0.7-2.0)	
30-49 yo: L vs D	Li, 2010		0.9 (0.8-1.0)	
30-49 yo ER+/PR+: L vs D			1.0 (0.8-1.1)	
30-49 yo ER-/PR-: L vs D			1.3 (1.1-1.7)	
30-49 yo: M vs D				0.9 (0.8-1.0)
30-49 yo ER+/PR+: M vs D				1.1 (1.0-1.2)
30-49 yo ER-/PR-: M vs D				1.2 (1.0-1.4)
50+ yo: L vs D			0.9 (0.8-0.9)	
50+ yo ER+/PR+: L vs D			1.0 (0.9-1.0)	
50+ yo ER-/PR-: L vs D			0.9 (0.8-1.1)	
50+ yo: M vs D				0.8 (0.8-0.9)
50+ yo ER+/PR+: M vs D				0.9 (0.9-1.0)
50+ yo ER-/PR-: M vs D				1.1 (1.0-1.3)
ER+/PR- vs ER+/PR+	Dunnwald, 2007	1.5 (1.4-1.6)	1.4 (1.0-1.9)	1.3 (0.9-1.7)
ER-/PR+ vs ER+/PR+		1.8 (1.6-2.0)	0.4 (0.1-1.4)	1.1 (0.6-2.0)
ER-/PR- vs ER+/PR+		2.3 (2.2-2.5)	1.9 (1.4-2.7)	2.7 (2.1-3.5)

*Multivariate-adjusted HR (95% CI) unless otherwise noted

1.9 Comparison of survival by intrinsic subtype and histologic subtype

Survival by IHC-based intrinsic subtype has been well characterized using the CBCS Phases 1 and 2 [3, 109]. In general, survival for HER2+ and Basal-like breast cancer is

worse than that observed for Luminal A and B subtypes [3, 109]. The 5-year survival of women diagnosed with Luminal A or B tumors of approximately 90% matches what has been reported from clinical trials among lobular tumors [3, 15, 17, 20, 26, 37–39, 109]. In contrast, the 10-year survival estimates of approximately 75-80% for Luminal A or B subtype from CBCS are nearly 20% higher than those reported for lobular tumors from clinical studies (Table 1.7) [3, 26, 37, 38, 40, 109]. While estimates of 5- and 10-year survival for patients with ductal tumors are lower than estimates for lobular tumors, these are likely confounded by molecular subtype as survival at 5 years is approximately 70% for patients with Basal-like or HER2-enriched tumors, which is about 20% lower than that for Luminal A and B tumors [3, 15, 17, 20, 26, 37–39, 109]. As such, it would be expected that lobular, ductal, and mixed tumors would show similar survival patterns when restricted to Luminal A intrinsic subtype. Unfortunately, due to variation in study-specific stratification for histologic subtype and the risk of death and intrinsic subtype and risk of death, comparisons between histologic subtype and intrinsic subtype cannot be inferred.

1.10 Molecular heterogeneity of histologic subtypes of breast cancer

There is tremendous molecular heterogeneity between histologic breast cancer subtypes. This heterogeneity extends beyond the traditionally discussed intrinsic subtypes and includes single cell markers. The most recent and comprehensive examination of the molecular differences between histologic subtypes was carried out by Ciriello *et al.* (2015) in TCGA [5].

Evidence from both RNA-based and IHC-based intrinsic subtype analyses indicates that lobular and mixed tumors have little molecular variability compared to that observed for ductal tumors. Using the RNA-based PAM50 intrinsic subtype classification [110], TCGA found lobular tumors were predominantly Luminal A (87%) (Table 1.8) [5]. A clinical trial in the Netherlands that used RNA-based intrinsic subtype methods reported that 71% of lobular tumors were Luminal A intrinsic subtype and reported higher proportions of HER2-enriched and Basal-like lobular tumors than in TCGA [16]. Among ductal tumors, 48-65%

are Luminal A or B, 10-17% are HER2-enriched, 21-26% are Basal-like, and 3-8% are Normal-like using RNA-based methods [5, 16]. TCGA reported RNA-based intrinsic subtype distributions for mixed tumors that were more variable than those for lobular tumors with 68% of mixed tumors classified as Luminal A, 23% as Luminal B, 5% each were Basal-like HER2-enriched subtypes [5].

Two epidemiologic studies have reported IHC-based intrinsic subtype among histologic subtypes and the distributions are similar those from TCGA. In work with the CBCS Phase 1 by Carey *et al.* (2006) and a recent report from the NHS, both observed approximately 80% of lobular cases were Luminal A (Table 1.8) [3, 6]. Among ductal cases, the distributions of intrinsic subtype were fairly similar between the CBCS and the NHS, but CBCS had less Luminal B cases (CBCS:16% vs NHS: 27%) and double the frequency of Basal-like tumors (CBCS: 22% vs NHS:11%) [3, 6]. For mixed tumors, the NHS reported less variability than the CBCS with 70% classified as Luminal A whereas the CBCS reported 51% were Luminal A, 13% were Basal-like, and 11% were unclassified [3, 6]. The percentages from the CBCS and the NHS for lobular, Luminal A tumors are slightly higher than results from clinical studies that also display considerable interstudy variability in intrinsic subtype distributions among ductal and mixed tumors using IHC-based methods [8–10, 12–15, 111].

Table 1.8. Distribution of intrinsic subtypes by histologic subtype of breast cancer

			Histologic Subtype		
Technical Method	Intrinsic subtype	Author, Year	Ductal (%)	Lobular (%)	Mixed (%)
RNA-based	Luminal A	Ciriello, 2015 (TCGA)	41	87	68
	Luminal B		25	4	23
	HER2-enriched		10	2	5
	Basal		21	1	5
	Normal-like		3	6	0
	Luminal A	Lips, 2012	27	71	
	Luminal B		21	5	
	HER2-enriched		17	5	
	Basal-like		26	10	
	Normal-like		8	10	
IHC-Based	Luminal A	Carey, 2006 (CBCS Phase 1)	47	82	51
	Luminal B		16	13	20

HER2/ER-		8	0	4
Basal-like		22	0	13
Unclassified		6	5	11
Luminal A	Sisti, 2016 (NHS)	52	75	70
Luminal B		27	21	27
HER2+		6	1	0
Basal-like		11	1	1
Unclassified		3	2	1
Luminal A	Braunstein, 2015	52	91	80
Luminal B		19	8	16
HER2+		5	1	1
Luminal-HER2 +		7	0	3
Triple Negative		17	0	1
Luminal A	Caldarella, 2013	27	42	41
Luminal B HER2-		24	30	36
Luminal B HER2+		14	7	8
HER2+		14	17	3
Triple Negative		21	3	23
Luminal A	Cha, 2015	42	75	
Luminal B		23	19	
HER2		10	1	
Triple Negative		25	4	
Luminal A	Azim, 2014	55	78	
Luminal B		17	8	
HER2+		12	3	
Triple Negative		17	11	
Luminal A	Lim, 2014	56	83	
Luminal B		11	5	
HER2+		11	1	
Triple Negative		22	11	
Luminal A	Engstrom, 2015	48	54	
Luminal B		37	34	
HER2+		8	1	
Basal-like		6	2	
Luminal A	Garcia-Fernandez, 2015	45	67	
Luminal B		35	27	
HER2 +		5	0	
Triple Negative		13	5	
Luminal A	Jung, 2010	51	91	
Non-Luminal A		49	9	

Intrinsic subtype markers cannot reliably predict histology; however, there may be other markers that can. TCGA found that 95% of lobular tumors were characterized by alterations of E-cadherin (CDH1) activity where only 2% of ductal tumors were

characterized as such [5]. This is consistent with IHC studies showing that E-cadherin loss is more common among lobular tumors and may be a discriminatory marker between lobular and ductal tumors [5, 14, 25, 105, 112–114]. Loss of CDH1, an epithelial cell-cell adhesion molecule, is hypothesized to play a role in the single-file epithelial cells infiltrating the stroma in lobular tumors, which is thought to contribute to difficulty in clinical detection, increased cell invasion, tumor metastasis, more aggressive disease, and an increased risk of death in breast and other cancers in human and animal models [14, 18, 51, 115–125].

In addition to CDH1, other single markers have been studied with respect to histologic subtype including, tumor suppressor TP53, CK5/6, EGFR, ER, and Ki67. Lobular tumors are less frequently TP53 mutant by IHC and DNA sequencing than ductal tumors [5, 15–17, 25, 26, 126]. Low levels of CK5/6 and EGFR (~5% positivity each) among lobular tumors is consistent with lobular tumors being predominantly Luminal A or B subtype, both of which are CK5/6 and EGFR negative [4, 14, 17, 25, 127, 128]. Additionally, TCGA hypothesized that ER regulation occurs through different molecular mechanisms in ductal and lobular tumors based on differences in protein levels and mutation frequency in ER regulating genes, including CDH1 [5, 114, 129, 130]. This hypothesis coincides with clinical trial work using chemotherapies that target either the ERα or ERβ subunits, which are differentially expressed by histologic subtype, resulting in varying chemotherapeutic responses between ductal and lobular tumors [89, 131–133]. Concerning cellular proliferation maker, Ki67, lobular tumors display lower expression than ductal tumors (16-34% vs 30-62%, respectively) as measured by IHC [11, 14, 26, 91, 92, 134]. Low cellular proliferation is thought to contribute to chemotherapeutic resistance among lobular tumors for therapies that rely on rapid cell division for drug uptake and are often more effective in treating patients with ductal tumors [5, 7, 14, 20, 38, 91, 92, 135]. TCGA found that mixed tumor gene expression profiles clustered with ductal or lobular tumors suggesting that mixed tumors do not represent a unique molecular entity [5]. The molecular profile differences of mixed tumors could lead to the varying epidemiologic risk factor profiles

observed for mixed tumors. In summary, there is considerable molecular heterogeneity between histologic subtypes concerning intrinsic subtypes and other commonly studied molecular markers. As such, the molecular variability observed for histologic subtype is compelling evidence for the stratification of both risk and survival estimates by both tumor characteristics.

1.11 Significance

Recent decades have established that breast cancer represents many diseases with distinct molecular phenotypes. There are countless analyses that stratify by various predictive and prognostic markers for breast cancer including ER status, intrinsic subtype, and even histologic subtype, to identify differences in risk and survival allowing for more tailored public health messages to reduce incidence and mortality. However, there is little epidemiologic research focusing on risk and survival of breast cancer that integrates both histologic and molecular subtype, both of which have unique risk factor and survival profiles.

When considering risk by histologic subtype, there is evidence that hormone regulating exposures including younger age at menarche and older age at menopause, older age at first birth, nulliparity, and combined E+P HT use more strongly increase the risk of lobular carcinoma than ductal or mixed carcinomas. Molecular characterization of histologic subtypes has indicated that lobular and mixed tumors are predominantly Luminal A intrinsic subtype and ductal tumors display intrinsic subtype diversity. To determine whether the strongest differences in risk come from histologic or intrinsic subtype, both must be considered simultaneously. Identifying the impact of modifiable breast cancer risk factors by histologic and molecular subtype can help to ascertain etiologic differences in tumor heterogeneity and can serve to clarify public health messages aimed at reducing the breast cancer burden.

Similarly, for survival, stratification by various predictive and prognostic markers highlights the heterogeneity of breast cancer. Based on the current literature, there is no

clear pattern of differential survival by histologic subtype as seen with risk factor analyses. There appears to be modification by ER status and time since diagnosis for risk of death between lobular and ductal tumors, but consistency in both analysis strategy and populations studied is lacking. Compared to histologic survival analyses, molecular survival analyses have been more clear and consistent in showing differences in survival by subtype. This leads to the question of whether incorporation of intrinsic breast cancer subtype information and histologic subtype will allow for a more complete and stable picture of survival by both tumor phenotypes.

CHAPTER 2: SPECIFIC RESEARCH AIMS

Breast cancer incidence rates have slowly increased since 2003 following a drop in incidence likely related to reduced menopausal hormone therapy use, which was found to increase breast cancer risk [72, 136]. There has not been a report describing changes in incidence of all stages of disease by histologic subtype in the past decade; however, SEER data (1993-2011) indicate that among patients with stage 4 disease, rates of both lobular and ductal carcinoma steadily increased (+3.0% vs +1.7% APC, respectively) [72]. These findings underscore the need to identify risk factors for each histologic subtype.

Previous research suggests that compared to ductal, lobular carcinoma is more strongly associated with younger age at menarche, later age at first birth and menopause, and use of menopausal hormone therapy [24, 27–29, 33, 35, 36, 49, 50, 64, 79]. Less is known about the risk of mixed ductal-lobular carcinoma, but associations appear intermediate to those of ductal and lobular carcinoma [24, 27, 29, 35, 36, 51]. Recent molecular analyses by The Cancer Genome Atlas (TCGA) showed intrinsic subtype heterogeneity between histologic subtypes identifying nearly 90% of lobular and mixed tumors as Luminal A subtype while ductal carcinomas display intrinsic subtype diversity similar to breast cancer overall (41% Luminal A, 25% Luminal B, 21% Basal-like, 10% HER2, and 3% Normal-like) [5]. These findings are similar to those from Phase 1 of the Carolina Breast Cancer Study [3] and a number of clinical studies [6, 11–16]. Therefore, the observed differences in risk factor profiles by histologic subtype may be confounded by intrinsic subtype, each of which has unique reproductive risk factor associations [3, 4, 6].

Similarly, there are different patterns of clinical characteristics for ductal, lobular, and mixed carcinomas. Compared to ductal carcinomas, lobular and mixed carcinomas occur

more frequently in older women, are larger, lower grade tumors, and are predominantly estrogen (ER) and progesterone (PR) receptor positive [5, 9, 11, 13–15, 17–26, 35, 37, 38, 89–92, 98–105]. Patterns of survival by histologic subtype are not as clear as risk factor profiles. Overall 5- and 10-year survival appears to be similar among the three histologic subtypes [15, 17, 20, 26, 37–40, 95]. The risk of breast cancer-specific death varies between studies making differences by histologic subtype difficult to discern. The influence of intrinsic and histologic subtype on breast cancer survival remains unexplored.

Using the Carolina Breast Cancer Study (CBCS) Phases 1-3 (1993-2013) the proposed study will investigate molecular profiles, epidemiologic risk-factor profiles, and survival differences between ductal (2,856, 66%), lobular (326, 7%), and mixed ductal-lobular (473, 11%) tumors.

Aim 1. To estimate the association between molecular and clinical characteristics and histologic subtype, overall and among Luminal A tumors, while comparing the CBCS findings to those observed in TCGA. Approach: The association of molecular and clinical characteristics and histologic subtype will be estimated by relative frequency differences between ductal, lobular, and mixed ductal-lobular tumors in the CBCS and TCGA. Then, the associations between clinical characteristics (tumor size, tumor grade, stage of disease, and lymph node status), age at diagnosis, race, and histologic subtype, overall and among Luminal A tumors, will be estimated. **Hypothesis:** We hypothesize that lobular and mixed tumors will be predominantly Luminal A intrinsic subtype and ductal tumors will show variability in intrinsic subtype. Further, we hypothesize that lobular and mixed tumors will be lower grade, larger in size, and more frequently higher stage disease than ductal tumors, but differences will decrease when restricted to Luminal A tumors.

Aim 2: To estimate the association between risk factors (race, age, menopausal status, recency of last birth, age at first birth, parity, lactation duration, and exogenous hormone use) and histologic subtype, overall and among Luminal A tumors.

Approach: We will use polytomous logistic regression in case-control (CBCS 1-2) and case-case analyses (CBCS 1-3) to estimate odds ratios and 95% confidence intervals as the measure of association between risk factors and histologic subtype. **Rationale:** Integrated analysis will reveal patterns of risk associated with histologic subtype among Luminal A tumors. **Hypothesis:** Differences in risk by histologic subtype will be diminished in integrated analyses among Luminal A tumors.

Aim 3: To estimate the association between histologic subtype and breast cancer-specific survival, overall and among Luminal A tumors.

Approach: We will perform survival analysis to estimate the association between histologic subtype and breast cancer-specific survival (CBCS 1-2). **Rationale:** Associations between histologic subtype and survival have been inconsistent. No population-based studies have characterized survival by histologic subtype while accounting for Luminal A subtype. **Hypothesis:** We hypothesize that among Luminal A tumors, lobular and mixed tumors will be associated with a higher risk of death than ductal tumors**.**

CHAPTER 3: RESEARCH STRAGEGY

3.1. Study populations

3.1.1 The Carolina Breast Cancer Study

The Carolina Breast Cancer Study (CBCS) was initiated in 1993 and enrolled participants in 3 phases, ending enrollment in 2013. CBCS Phases 1-2 is a population-based, case-control study that recruited participants from 24 of the 100 North Carolina counties [137]. CBCS was designed to oversample women under 50 years of age and black women, who comprised approximately 40% of the study population [82]. Cases of invasive breast cancer were identified using rapid case ascertainment via the North Carolina Central Cancer Registry. For CBCS 1-2, breast cancer cases were enrolled following informed consent and fulfillment of the eligibility criteria: being female, first diagnosis of breast cancer [(invasive or *in situ* (CBCS Phase 2 participants only)], aged 20-74 years at diagnosis, and residence in the specified county regions. Controls for CBCS 1-2 were identified from Division of Motor Vehicle lists (women under age 65) and the Health Care Financing Administration lists (women 65 years or older). Controls were frequency matched to cases based on 5-year age categories and race. Cooperation rates in CBCS 1-2 among eligible women were 78% for cases and 70% for controls [138].

CBCS Phase 3 (2008-2013) is a case-only study with case ascertainment carried out as in CBCS 1-2 with the oversampling of young women and black women, but expanding to enrollment to women from 44 of the 100 North Carolina counties. Following informed consent, participants in all 3 phases completed in-person interviews by a trained study nurse to obtain study questionnaire data, including risk factor data, height, weight, waist circumference, and a 30mL blood draw. Cases gave consent to obtain medical records and

tumor tissue blocks and/or tumor tissue slides from their treatment centers. Clinical data including tumor size, stage of disease, and lymph node status were abstracted from the medical records. Tumor grade was determined using combined mitotic and clinical tumor grade for CBCS 1 and 3. The study maintains Institutional Review Board approval at the University of North Carolina.

This dissertation utilizes data from CBCS 1-3 to investigate the association between molecular and histologic subtypes of breast cancer in relation to established breast cancer risk factors and survival. Phases 1 (1993-1996), 2 (1998-2001), and 3 (2008-2013) enrolled 4,806 women with invasive breast cancer. To be eligible for these analyses, CBCS participants must have contributed hematoxylin and eosin (H&E) immunohistochemistry (IHC) stained tumor tissue that underwent centralized pathology review to determine the histologic subtype of invasive breast cancer. For ductal, lobular, and mixed tumors, this results in 3,655 of the 4,806 cases from CBCS 1-3 available for the main analyses. The remaining cases (n=704) are made up by the less common histologic subtypes [mixed ductal/non-lobular (n=285), mucinous (n=89), mixed ductal/metaplastic (n=63), metaplastic (n=44), DCIS w/focal invasion (n=44), undifferentiated high grade (n=29), tubular (n=23), micropapillary (n=21), papillary (n=19), medullary (n=18), pleomorphic lobular (n=17), anaplastic (n=14), apocrine (n=11), cribriform (n=9), neuroendocrine (n=3), others (n=15)] or have unknown or missing histology data (n=447) [unknown (n=99) or missing (n=376)].

3.1.2. The Cancer Genome Atlas

The Cancer Genome Atlas (TCGA) breast cancer study population consists of patients with invasive breast cancer who were undergoing surgical resection and had received no prior chemotherapy or radiation [1, 5]. Participants were enrolled from various medical institutions around the country, which all had IRB approval to obtain tumor tissue and adjacent normal tissue when possible. A total of 808 females with newly diagnosed invasive breast cancer were used in this dissertation, which includes all female cases (9 male cases

were excluded) used in the molecular characterization of ductal, lobular, and mixed ductal-lobular invasive breast tumors carried out by Ciriello *et al.* (2015) [5]. Tumor specimens were used for histologic subtype classification, RNA extraction, and RNA sequencing among a number of other molecular assays. Age at diagnosis, race, tumor size, lymph node status, and stage of disease were abstracted from the medical records. All stages of disease were mapped to the AJCC 7th edition of breast cancer stages of disease. TCGA data is publicly available [1, 5].

3.2. Breast cancer subtyping methods and survival data

3.2.1. Histologic subtype of breast cancer

CBCS. The histologic subtypes ductal, lobular, and mixed ductal-lobular, comprise 84% of invasive breast cancers in the CBCS with the remaining cases classified as: mixed ductal/non-lobular, mucinous, mixed ductal/metaplastic, metaplastic, DCIS w/focal invasion, undifferentiated high grade, tubular, micropapillary, papillary, medullary, pleomorphic lobular, anaplastic, apocrine, cribriform, neuroendocrine, and other. Histologic subtype for CBCS 1-3 was determined via centralized pathology review. To be classified as a single histologic subtype the tumor was at least 80% representative of a single histology. Mixed tumors contained <80% of one histologic subtype and ≥20% of another subtype.

TCGA. Histologic subtype was available for all 808 women and was determined using clinical diagnostic criteria for histologic subtype applied by an expert pathologist committee [5]. A consensus ruling on histologic subtype was reached using agreement between pathology reports (PRS) and the expert pathologist committee (ERC) classification. Discrepancies were resolved using the following rules [5]:

- If (EPC=Ductal AND PRS= Ductal) or (EPC= Ductal AND PRS=Mixed), then histologic subtype = Ductal.
- If (EPC=Lobular AND PRS = Lobular) OR (EPC= Lobular AND PRS = Mixed) OR (EPC=Mixed AND PRS =Lobular), then histologic subtype = Lobular.
- If (EPC=Lobular AND PRS=Ductal) OR (EPC=Mixed AND PRS=Mixed) OR (EPC=Ductal

AND PRS =Lobular) OR (EPC=Mixed AND PRS=Ductal), then histologic subtype = Mixed.

- If (EPC=Other OR PRS =Other), then histologic subtype = Other.

3.2.2. Clinical and intrinsic breast cancer subtypes

Clinical subtype

CBCS. For CBCS 1-2, estrogen receptor (ER) and progesterone receptor (PR) status originated from medical records for 80% of cases and IHC was performed at the University of North Carolina (UNC) for ER and PR on tumors from the remaining 20% of participants with tissue available [3]. For tumors that underwent IHC staining at UNC, a study pathologist determined ER and PR positivity defined as any staining >5% (CBCS 1) or by using contemporaneous clinical cut points (CBCS 2) [3]. Human epidermal growth factor receptor 2 (HER2) staining was performed at UNC for all CBCS 1-2 cases with available tumor tissue. HER2 positivity was defined as membrane or membrane plus cytoplasmic staining classified as weak or greater intensity in ≥10% of tumor cells as determined by a study pathologist [3].

In CBCS 3, 98% of cases had ER, PR, and HER2 information, which was abstracted from medical records serving as the primary source of molecular subtype. For the remaining 2% of cases without ER, PR, and HER2 status from the medical records, IHC staining was performed at UNC with positivity cut points of ≥10% for ER and PR. HER2 positivity was defined as 3+ staining intensity [negativity was defined as 0/1+ (cases with equivocal staining, 2+ staining intensity, were excluded)]. As described by Allott *et al.* (2016) a core-to-case collapsing method was used for the CBCS 3 tissue microarrays (TMA) where cases could have 1-4 tumor tissue cores available for IHC staining [139]. An automated analysis approach was used to determine positivity for ER, PR, and HER2. Agreement between the clinical record and automated analysis was 93% for ER and HER2 and 88% for PR [139].

TCGA. ER and PR status was provided from Tissue Source Sites using contemporaneous positivity cut points for immunohistochemistry. HER2 status was also available from the

Tissue Source Sites for a majority of cases and data was supplemented by TCGA using HER2 copy number rather than fluorescence in situ hybridization (FISH) data, where necessary [1, 5].

For CBCS 1-3 and TCGA, 3-marker clinical subtypes were defined as follows: Luminal A (ER+ or PR+ and HER2-), Luminal B (ER+ or PR+ and HER2+), Triple Negative (TN) (ER- and PR- and HER2-), and HER2+ (ER- and PR- and HER2+).

Intrinsic subtype

CBCS. For CBCS 3, Nanostring assays were carried out on a randomly sampled subset of available formalin fixed paraffin embedded (FFPE) tumor tissue cores (n=1,122) [110, 139]. RNA was isolated from 2, 1.0-mm cores from the same FFPE block using the Qiagen RNeasy FFPE kit (catalogue # 73504). Nanostring assays, which use RNA counting as a measure of gene expression, were conducted. Samples lacking sufficient quality data (n=101) or cases with >1 tumor block were excluded (n=8). RNA-based intrinsic subtype was determined using the PAM50 gene signature described by Parker *et al.* (2009) [110]. Based on the highest Pearson correlation with a subtype-defined centroid, each tumor was categorized into one of five intrinsic subtypes (Luminal A, Luminal B, HER2, Basal-like, Normal-like), using the 50 gene, PAM50 signature described Parker *et al.* (2009) [1, 110].

TCGA. For TCGA, RNA was extracted from flash frozen tumor samples as previously described [1, 5]. Whole genome microarrays were used in RNA sequencing, which was performed at UNC [5]. PAM50 intrinsic subtype was assigned by first selecting a sample of cases from the TCGA samples used in Ciriello *et al.* (2015) with RNA sequencing data and ER status to match the distribution of ER status in the original PAM50 training set [5]. Then, the entire sample of tumors was assigned PAM50 intrinsic subtypes by adjusting gene expression to the median values calculated for the PAM50 genes from the "ER balanced" subset [1, 5, 110].

3.2.3 TP53 status

CBCS. TP53 status determined by IHC was available for CBCS 3 cases. IHC staining conditions were optimized using breast tissue sections and cell lines with established *TP53* mutation status [(wild type: MCF-7, SUM102), (mutant: SKBR3)]. IHC was carried out at the UNC Translational Pathology Laboratory using a Bond Autostainer (Leica Microsystems Inc. Norwell, MA 02061). Slides were dewaxed in Bond Dewax solution (AR9222) and hydrated in Bond Wash solution (AR9590). Antigen retrieval was performed for 20 minutes in Bond-Epitope Retrieval Solution 1 pH-6.0 (AR9961). Slides were incubated for 15 min with a mouse monoclonal anti-TP53 antibody (BioGenex, Fremont, CA; clone D07 [catalog # MU239-UC], 1:7200). Detection was performed using the Bond Intense R Detection System (DS9263) supplemented with Dako EnVision Mouse (Carpinteria, CA, K4001). Stained slides were counterstained with hematoxylin, dehydrated, and coverslipped. A control TMA containing TP53 positive and negative breast tissue and cell lines was included in each run along with a negative control (no primary antibody). CBCS 3 TMA construction, has been previously described [139]. TMAs were constructed with 1-4, 1mm cores per participant.

TP53-stained TMA slides were scanned using the Aperio ScanScope XT at 20x magnification. Details of the scoring algorithm have been described previously [140]. Briefly, TP53 staining was measured with the Aperio Nuclearv9 algorithm by quantifying tumor cellularity and was combined with the Genie Histology Pattern Recognition tool to correctly classify the number of tumor and normal epithelial cells per core allowing for enrichment of tumor cells. Algorithm parameters including nuclear size and nuclear compactness were optimized to achieve the best nuclear segmentation. The algorithm returned a total number of nuclei per core and the number of nuclei positive for TP53. To determine the average percent positivity, a method of core-to-case collapsing developed by *Allott et al*. (2016) [139] was used by summing the total number of nuclei per core (1-4 cores/participant). Each core was given a weight equal to the number of core nuclei divided by the total nuclei for the participant. For the core weighted percent positivity, the core's

TP53 percent positivity was multiplied by the core's weight. The weighted core values were summed to obtain the participant's overall weighted percent TP53 positivity. Weighted percent TP53 positivity was dichotomized to classify patients as negative or positive (<10% for negative/wild-type, ≥10% positive/mutant). The ≥10% cut point was chosen to ensure similar relative frequencies of TP53 mutation to those observed in CBCS 1-2 [140–142]. However, the *TP53* protocol for CBCS 3 differed from that used for CBCS 1-2 [moderate staining intensity (>2+), ≥50% TP53 positive tumor cells, or average H score of 60 was previously used] because the CBCS 3 staining protocol resulted in more intense staining with increased sensitivity [140, 142].

RNA-based TP53 status was also determined for cases in CBCS 3. In addition to the PAM50 genes, the Nanostring probe set contained 52 genes for a previously validated TP53 signature that was used to classify CBCS 3 tumors as TP53 mutant-like or wild-type-like [143]. For the TP53 signature, the mutant-like versus wild-type-like class was determined based on a similarity-to-centroid approach (Pearson coefficient) for each case [143]. The TP53 signature is independent of intrinsic subtype and can be used to detect deficiencies in the TP53 pathway.

TCGA. RNA-based TP53 status was available in TCGA. TP53 status, as determined by the aforementioned TP53 gene signature was determined using RNA gene expression data from microarrays in the same manner used for CBCS 3 as described above [110, 143]. Tumors were classified as TP53 mutant-like or wild-type-like.

3.2.4 Survival data

Breast cancer-specific survival is the outcome of interest in the third aim of this dissertation using CBCS 1-2 participants. Survival time was defined from the date of diagnosis to the date of death from breast cancer or any cause or loss to follow-up. To identify breast cancer specific deaths International Classification of Death codes, ICD-9: 174.9 and ICD-10: 50.9, were used to determine the primary or secondary causes of death in the National Death Index (NDI) database. If a participant had an ICD-9 or ICD-10 code

other than those listed for breast cancer as the primary cause of death, the cause of death was classified as 'other' and was censored in the breast cancer-specific survival analyses. The most recent CBCS NDI linkage was 12/31/2011 resulting in up to maximum of 18 years of survival data (median 13.5 years).

3.3 Statistical analysis

3.3.1. Aim 1. To estimate the association between molecular and clinical characteristics and histologic subtype, overall and among Luminal A tumors, while comparing the CBCS findings to those observed in TCGA.

Generalized linear models were used to estimate relative frequency differences (RFDs) and 95% confidence intervals (95% CIs) as the measure of association between histologic subtype and age, race, tumor characteristics, clinical subtype, intrinsic subtype, and TP53 status [144]. For CBCS 1-3, unweighted sample size counts are presented alongside weighted percentages to account for the sampling design of CBCS. The following variables were studied in association with histologic subtype in CBCS and TCGA: age (≤50, >50), race [CBCS: self-reported black, non-black (>98% white, 2% other (referred to as white); TCGA: black and white (other races were excluded in race-specific analyses due to low sample sizes)], combined mitotic and clinical tumor grade (CBCS 1 and 3 only) (low-intermediate, high), AJCC stage of disease (I/II, III/IV), lymph node status (positive, negative), tumor size (≤2cm, >2cm), ER, PR, HER2 (negative/positive), 3-marker IHC-based clinical subtype (Luminal A, Luminal B, Triple Negative, HER2+) , PAM50 intrinsic breast cancer subtype [Luminal A, Non-Luminal A (Luminal B, Basal-like, HER2-enriched) (excluding Normal-like subtype, which generally reflects insufficient tumor cellularity in the sampled biospecimens)], TP53 status (IHC: <10% for negative/wild-type, ≥10% for positive/mutant; RNA: mutant-like, wild-type-like). Sample proportions and generalized linear model analyses were done in SAS version 9.4 (SAS Institute, Cary, NC). Graphs were constructed using GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla,

CA). P-values were produced for a two-sided test with an alpha of 0.05 for statistical significance.

3.3.2 Aim 2: To estimate the association between reproductive breast cancer risk factors and histologic subtypes, overall and among Luminal A tumors.

Patient characteristics.

The associations between histologic subtype and race, age, menopausal status, and clinical subtype were estimated using generalized linear models that were adjusted for age, race, and study phase (1, 2, 3), where appropriate. Relative frequency differences (RFDs) and 95% confidence intervals (95% CIs) were estimated as the measure of association [144]. To account for the CBCS sampling design, weighted percentages are presented alongside unweighted sample size counts. Patient characteristics were defined as: race [self-report: black, non-black (>98% white, henceforth referred to as white)], age (years) (<40, 40-49, 50-59, ≥60), menopausal status (pre-, post-), and clinical subtype (Luminal A, Luminal B, Triple Negative, HER2+) as defined above.

Reproductive risk factor analyses

Polytomous logistic regression was used to obtain odds ratios (ORs) and 95% confidence intervals (95% CIs) as the measure of association between each reproductive breast cancer risk factor and histologic subtype in case-control (CBCS 1-2) and case-case (CBCS 1-3) analyses. Weighted percentages to account for the study sampling design are presented with unweighted sample size counts. The following risk factors were studied in association with histologic breast cancer subtype: parity (nulliparous, 1, 2, ≥3), years since last birth (defined as: age at diagnosis minus age at last birth; among parous women only) (0-≤10, 10-≤20, >20), age at first live birth (years) (parous women only) (<26, ≥26), lifetime lactation duration (months) (parous women only) (never, 0-≤12, >12), oral contraceptive use (never, former, current), and hormone therapy (HT) use [never, estrogen alone, combined estrogen + progesterone (E+P)]. Additional variables used case-control analyses included: study phase (1, 2, 3), age (continuous) (20-74), family history of breast

cancer (yes, no), alcohol intake (ever, never), smoking duration (years) (never, <10, 11-19, ≥20), oral contraceptive use (ever, never), breastfeeding (ever, never), age at menarche (years) (<13, ≥13), and the CBCS offset term to account for the study sampling design. Case-case analyses are presented for CBCS 1-3 to assess etiologic heterogeneity by histologic subtype [4, 145]. In case-case analyses, ductal served as the referent group for lobular and mixed tumor types. Case-case models are adjusted for age, race, and study phase.

Tests of trend for the null hypothesis that the slope of the line equals zero for age at diagnosis, parity, and lactation duration, were conducted with the variable modeled continuously and the beta-coefficient p-value reported. We conducted a race-stratified sensitivity analysis of the pregnancy-related risk factor associations with histologic subtype, but we were unable to assess racial differences in oral contraceptive and HT use due to low sample size among black women for current OC use and combined E+P HT use. All analyses were done in SAS version 9.4 (SAS Institute, Cary, NC). P-values were produced for a two-sided test with an alpha of 0.05 for statistical significance.

3.3.3 Aim 3: To estimate the association between histologic subtype and breast cancer-specific survival, overall and among Luminal A tumors.

First, descriptive analyses for ductal, lobular, and mixed histologic subtypes from CBCS 1-2 were carried out for race [self-report: black, non-black (>98% white, henceforth referred to as white)], age (years) (<40, 40-49, 50-59, ≥60), menopausal status (pre-, post-), tumor size (≤2cm, >2cm), combined mitotic and clinical tumor grade (CBCS 1 only) (low-intermediate, high), lymph node status (positive, negative), AJCC stage of disease (I/II, III/IV), ER, PR, HER2 status (negative/positive), and 3-marker clinical subtype (Luminal A, Luminal B, Triple Negative, HER2+) as defined above. Weighted sample percentages were presented with unweighted sample counts.

Survival proportions at 5-, 10-, and 15-year time points were estimated by histologic subtype. Kaplan-Meier curves were created to visually assess differences in breast

cancer-specific survival by histologic subtype, overall and among Luminal A tumors. Non-parametric Log-Rank tests were carried out for each set of Kaplan-Meier survival curves. Log-log plots were examined to assess deviation from the proportional hazards assumption, which appeared to occur around the 10-year timepoint. Therefore, time-stratified Cox Proportional Hazards models were used to estimate hazards ratios (HRs) and 95% confidence intervals (95% CI) as the measure of association between histologic subtype and breast cancer-specific death adjusting for age and race for 0-10 years and >10 years post-diagnosis. To further assess violations of the proportional hazards assumption, 1^{st}-, 2^{nd}-degree polynomial and log(time) interaction terms between time and histologic subtype were examined. Based on Likelihood Ratio Tests (LRT) and Akaike Information Criterion (AIC) an interaction term between time modeled as a quadratic term and histologic subtype was included in all Cox regression models. All analyses were done in SAS version 9.4 (SAS Institute, Cary, NC). Graphs were constructed using GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla, CA). P-values were produced for a two-sided test with an alpha of 0.05 for statistical significance.

3.4 Strengths and limitations

Using the Carolina Breast Cancer Study, we were able to examine the molecular characteristics, risk factor profiles, and survival profiles by histologic breast cancer subtype in a population-based, racially-diverse study. We used centralized pathology review to determine histologic subtype allowing for consistency in histologic subtype categorization between all three phases of CBCS. CBCS has rich molecular data allowing us to draw direct comparisons to the TCGA for associations between histologic subtype and clinical subtype, PAM50 intrinsic breast cancer subtype, and TP53 status. CBCS also has data on a number of reproductive risk factors and exogenous hormone use allowing us to characterize these associations by histologic subtype. We were able to restrict our analyses to Luminal A tumors to determine whether the observed risk factor associations by histologic subtype persist among Luminal A tumors only, which has yet to be reported.

While CBCS had a large proportion of mixed ductal-lobular tumors, we were unable to classify the mixed tumors as ductal-like or lobular-like as illustrated by the TCGA [5]. Therefore, we focused on the differences between ductal and lobular tumors as the mixed tumor associations for risk factors and survival were intermediate to those for ductal and lobular tumors with no unique patterns emerging. Additionally, for the survival analyses in aim 3, sample sizes were too small for lobular and mixed tumors to draw firm conclusions. As CBCS 3 finished recruitment in 2013, survival data is not yet mature for these participants and as such could not be included in these analyses.

3.5 Summary

In summary, dissertation was well poised to characterize the impact of both molecular and histologic subtype on breast cancer risk and survival using a number of statistical methods in the population-based CBCS. Using relative frequency differences as an absolute measure of association helped in comparing and contrasting molecular differences between ductal, lobular, and mixed tumors in the CBCS and TCGA. Understanding these associations then helped us to account for intrinsic subtype in our risk factor analyses. Failing to account for intrinsic subtype and histologic subtype may not give a clear picture of breast cancer risk, especially if these risk factors were more strongly associated with molecular or histologic subtype. Additionally, this dissertation sought to identify etiologic differences between histologic subtypes that may help to further inform the biologic development of invasive breast cancer. While our survival analyses were underpowered, we were able to account for intrinsic and histologic subtype to determine if both tumor phenotypes contribute to breast cancer-specific survival.

CHAPTER 4: DIFFERENCES IN RACE, MOLECULAR, AND TUMOR CHARACTERISTICS BY HISTOLOGIC SUBTYPE OF INVAIVSE BREAST CANCER

4.1 Introduction

Invasive breast cancer is composed of several distinct histologic subtypes. Ductal carcinoma is most commonly diagnosed, representing 60-80% of tumors; however, lobular and mixed ductal/lobular (henceforth referred to as mixed) carcinoma represent a substantial portion of the breast cancer burden, each diagnosed in up to 15% of cases [18, 51, 52]. Previous clinical research has shown that compared to ductal carcinoma, lobular carcinoma tends to be diagnosed in women at older ages, presents as larger, lower grade tumors, is more frequently diagnosed at higher stage, and is associated with a higher rate of contralateral disease [8, 9, 11, 12, 15, 17, 18, 20–26, 37, 54, 89–102]. While studies have shown that lobular and mixed tumors are predominantly Luminal A intrinsic subtype, it is unclear if the observed associations between histologic subtype and tumor characteristics will be similar in magnitude and direction when restricted to Luminal A subtype.

A recent analysis from The Cancer Genome Atlas project (TCGA) found lobular tumors show distinct molecular differences from ductal tumors [5]. However, TCGA tumors are larger in size, diagnosed at later stages of disease, and are predominantly from white women, which necessitates evaluating these associations in racially diverse populations among women with earlier stage of disease and smaller tumors. To explore the relationship between histologic and molecular subtype in a population-based sample with larger numbers of younger women (<50) and black women, we evaluated associations between patient and tumor characteristics, PAM50 subtype, a validated TP53 gene signature, and histologic subtype among participants in the Carolina Breast Cancer Study (1993-2013) and placed them in context of the same associations estimated in TCGA.

4.2 Methods

4.2.1 Study populations

The present analysis includes 4,359 cases of invasive breast cancer from the Carolina Breast Cancer Study (CBCS), Phases 1-3 (1993-2013). The CBCS is a population-based study among women from North Carolina [137], designed to oversample younger women (<50 years at diagnosis) and black women [82]. Initiated in 1993, the CBCS recruited participants from 24 (Phase 1-2) and from 44 (Phase 3) of the 100 North Carolina counties using rapid case ascertainment via the North Carolina Central Cancer Registry. After giving informed consent, CBCS breast cancer cases were enrolled under an Institutional Review Board protocol approved at the University of North Carolina. CBCS eligibility criteria included being female, a first diagnosis of breast cancer [invasive or (*in situ*: Phase 2 only)], aged 20-74 years at diagnosis, and residence in specified counties. Cases provided consent to access tumor tissue blocks/slides and medical records from treatment centers.

The Cancer Genome Atlas (TCGA) study population has been described previously [1, 5]. A total of 808 females with newly diagnosed invasive breast cancer were used in this analysis. Cases were enrolled at numerous medical institutions and provided informed consent to obtain tumor tissue specimens used for histologic subtype classification, RNA extraction, RNA sequencing, and other molecular assays. Age at diagnosis, race, tumor size, lymph node status, and stage of disease were abstracted from the medical records. TCGA data is publicly available [1, 5].

4.2.2 Histologic subtype

Eligible CBCS cases were those with invasive tumor tissue available for centralized pathology review. Single histologic subtype tumors were ≥80% representative of a single histology and mixed tumors contained ≥20% of a second histologic subtype in a tumor of another dominant (<80%) histologic subtype. The following subtypes were included in the main analysis: ductal (n=2,856), lobular (n=326), and mixed ductal/lobular (n=473) histologic subtype. The following histologic s7ubtypes were included in the 'other' category

(n=704) for select analyses: mixed ductal/non-lobular (n=285), mucinous (n=89), mixed ductal/metaplastic (n=63), metaplastic (n=44), DCIS w/focal invasion (n=44), undifferentiated high grade (n=29), tubular (n=23), micropapillary (n=21), papillary (n=19), medullary (n=18), pleomorphic lobular (n=17), anaplastic (n=14), apocrine (n=11), cribriform (n=9), neuroendocrine (n=3), others (n=15). Cases with unknown (n=99) or missing (n=376) histologic subtype were excluded.

In the TCGA, histologic subtype was available for all 808 women and was determined using clinical diagnostic criteria for histology applied by an expert pathologist committee [5]. A consensus ruling on histologic subtype was reached using the pathology reports and pathologist committee classification.

4.2.3 IHC-based clinical subtypes

For CBCS 1-2, estrogen receptor (ER) and progesterone receptor (PR) status was abstracted from medical records for approximately 80% of cases and the remaining cases with available tumor tissue had whole slide immunohistochemistry (IHC) staining performed at UNC on tumor tissue samples. The percent positivity for ER and PR was determined by a study pathologist using contemporaneous clinical definitions. Human epidermal growth factor receptor 2 (HER2) IHC staining was performed at UNC for cases with available tissue (positivity defined as membrane/membrane plus cytoplasmic staining classified as weak or greater intensity in ≥10% of tumor cells) [3, 109].

In CBCS3, 98% of cases had ER, PR, and HER2 data abstracted from the medical records, which served as the primary data source used to determine the clinical subtypes for CBCS3. For the remaining 2% of cases without medical record data on ER, PR, and HER2, IHC staining was performed at UNC. For these 2% of cases, positivity cut points of ≥10% were used for ER and PR. HER2 positivity was defined as 3+ staining intensity [negative was defined as 0/1+ (equivocal cases with 2+ staining were excluded)] [139]. As Allott *et al.* (2016) have described, for CBCS3 multiple tissue microarray (TMA) cores per case were stained for ER, PR, and HER2 and a core-to-case collapsing method was applied to classify

the case as positive/negative for each marker [139]. Percent positivity and staining intensity was determined by the Genie and NuclearV9 digital algorithm (Aperio Technologies).

In TCGA, ER and PR status was provided from Tissue Source Sites using contemporaneous positivity cut points. HER2 data was available for a majority of cases, but where unavailable was supplemented by TCGA using HER2 copy number rather than FISH data when needed [1, 5].

Across all studies, ER, PR and HER2 status, predominantly from the medical records for CBCS 1-3 and from IHC performed by Tissue Source Sites for TCGA, was used to create 3-marker IHC-based clinical subtypes defined as: Luminal A (ER+ or PR+, HER2-), Luminal B (ER+ or PR+, HER2+), Triple Negative (TN) (ER-, PR-, HER2-), and HER2+ (ER-, PR, HER2+).

4.2.4 RNA-based intrinsic subtypes

For CBCS3, RNA counting (Nanostring) assays were carried out on a randomly sampled subset of available FFPE tumor tissue cores (n=1,122) [110, 139]. RNA was isolated from 2, 1.0-mm cores from the same FFPE block using the Qiagen RNeasy FFPE kit (cat# 73504). Samples lacking sufficient quality data (n=101) or cases with >1 tumor block (n=8) were excluded. RNA-based intrinsic subtype was determined using the PAM50 gene signature described by Parker *et al.* (2009) [110]. Based on the highest Pearson correlation with a subtype-defined centroid and each tumor was categorized into one of five intrinsic subtypes (Luminal A, Luminal B, HER2, Basal-like, Normal-like).

For TCGA, RNA was extracted from flash frozen tumor samples as previously described [1, 5]. PAM50 intrinsic subtype was determined using RNA gene expression data from microarrays or RNA sequencing data and categorized into one of the 5 intrinsic subtypes using a similar algorithm as applied to CBCS 3 data [110, 143].

4.2.5 TP53 status

TP53 status was determined by immunohistochemistry (IHC) and by RNA expression for CBCS 3 cases. IHC staining conditions were optimized using breast tissue sections and cell lines with established *TP53* mutation status [(wild type: MCF-7, SUM102), (mutant: SKBR3)]. *TP53* status determined IHC was available for CBCS 3 cases. IHC staining conditions were optimized using breast tissue sections and cell lines with established *TP53* mutation status [(wild type: MCF-7, SUM102), (mutant: SKBR3)]. IHC was carried out at the UNC Translational Pathology Laboratory using a Bond Autostainer (Leica Microsystems Inc. Norwell, MA 02061). Slides were dewaxed in Bond Dewax solution (AR9222) and hydrated in Bond Wash solution (AR9590). Antigen retrieval was performed for 20 minutes in Bond-Epitope Retrieval Solution 1 pH-6.0 (AR9961). Slides were incubated for 15 minutes with a mouse monoclonal anti-TP53 antibody (BioGenex, Fremont, CA; clone D07 [catalog # MU239-UC], 1:7200). Detection was performed using the Bond Intense R Detection System (DS9263) supplemented with Dako EnVision Mouse (Carpinteria, CA, K4001). Stained slides were counterstained with hematoxylin, dehydrated, and coverslipped. A control TMA containing TP53 positive and negative breast tissue and cell lines was included in each run along with a negative control (no primary antibody). CBCS 3 TMA construction, has been previously described [139]. TMAs were constructed with 1-4, 1mm cores per participant.

TP53-stained TMA slides were scanned using the Aperio ScanScope XT at 20x magnification. Details of the scoring algorithm have been described previously [140]. Briefly, TP53 staining was measured with the Aperio Nuclearv9 algorithm by quantifying tumor cellularity and was combined with the Genie Histology Pattern Recognition tool to correctly classify the number of tumor and normal epithelial cells per core allowing for enrichment of tumor cells. Algorithm parameters including nuclear size and nuclear compactness were optimized to achieve the best nuclear segmentation. The algorithm returned a total number of nuclei per core and the number of nuclei positive for TP53. To determine the average percent positivity, a method of core-to-case collapsing developed by

Allott et al. (2016) [139] was used by summing the total number of nuclei/core (1-4 cores/participant). Each core was given a weight equal to the number of core nuclei divided by the total nuclei for the participant. For the core weighted percent positivity, the core's TP53 percent positivity was multiplied by the core's weight. The weighted core values were summed to obtain the participant's overall weighted percent TP53 positivity. Weighted percent TP53 positivity was dichotomized to classify patients as negative or positive (<10% for negative/wild-type, ≥10% positive/mutant). The ≥10% cut point was chosen to ensure similar relative frequencies of TP53 mutation to those observed in CBCS 1-2 [140–142]. However, the *TP53* protocol for CBCS 3 differed from that used for CBCS 1-2 [moderate staining intensity (>2+), ≥50% TP53 positive tumor cells, or average H score of 60 was previously used] because the CBCS 3 staining protocol resulted in more intense staining with increased sensitivity [140, 142].

RNA-based TP53 status was also determined for cases in CBCS 3. In addition to the PAM50 genes, the Nanostring probe set contained 52 genes for a previously validated TP53 signature [143] that was used to classify CBCS 3 tumors as TP53 mutant-like or wild-type-like. For the TP53 signature, the mutant-like versus wild-type-like class was determined based on a similarity-to-centroid approach (Pearson coefficient) for each case [143]. The TP53 signature is independent of intrinsic subtype and can be used to detect deficiencies in the TP53 pathway.

RNA-based TP53 status was available in TCGA. TP53 status, as determined by the aforementioned TP53 gene signature was determined using RNA gene expression data from microarrays in the same manner used for CBCS 3 as described above [110, 143]. Tumors were classified as TP53 mutant-like or wild-type-like.

4.2.6 Statistical analysis

Generalized linear models were used to estimate relative frequency differences (RFDs) and 95% confidence intervals (95% CIs) as the measure of association between histologic subtype and age, race, tumor characteristics, clinical subtype, intrinsic subtype,

and TP53 status [144]. In CBCS 1-3, unweighted sample size counts are presented alongside weighted percentages to account for the sampling design of CBCS. The following variables were studied in association with histologic subtype in CBCS and TCGA: age (≤50, >50), race [CBCS: self-reported black, non-black (>98% white, 2% other (referred to as white); TCGA: black and white (other races were excluded in race-specific analyses due to low sample sizes)], combined mitotic and clinical tumor grade (CBCS 1 and 3 only) (low-intermediate, high), AJCC stage of disease (I/II, III/IV), lymph node status (positive, negative), tumor size (≤2cm, >2cm), ER, PR, HER2 (negative/positive), 3-marker IHC-based clinical subtype, PAM50 intrinsic breast cancer subtype (excluding Normal-like subtype, which generally reflects insufficient tumor cellularity in the sampled biospecimens), TP53 status (IHC: negative/wild-type, positive/mutant; RNA: mutant-like, wild-type-like). Sample percentages and generalized linear regression analyses were done in SAS version 9.4 (SAS Institute, Cary, NC). Graphs were constructed using GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla, CA). P-values were produced for a two-sided test with an alpha of 0.05 for statistical significance.

4.3 Results

4.3.1 Characteristics of lobular and mixed tumors were quantitatively different from those for ductal tumors.

Lobular and mixed tumors, compared to ductal, displayed consistent differences in patient and tumor characteristics in CBCS and TCGA (Table 4.1a and 4.1b). Associations between histology and age and race were weak, with black women and women ≤50 years of age slightly less likely to be diagnosed with lobular and mixed disease, relative to ductal, in both studies. Lobular and mixed tumors displayed similar patterns of association with tumor characteristics in CBCS and TCGA. Compared to ductal, lobular tumors tended to be larger [CBCS Lobular Relative Frequency Difference (RFD) (>5 cm vs ≤2cm): 14.0%, 95% CI (10.7, 17.4)], less frequently high grade [CBCS Lobular RFD (high vs. low-intermediate):

-43.4%, 95% CI (-45.2, -41.6)], and higher stage [CBCS Lobular RFD (III/IV vs. I/II): 12.6, 95% CI (9.7, 15.4)]. In CBCS and TCGA, lobular and mixed tumors were less frequently ER-, PR-, and, HER2+. TCGA had higher proportions of larger tumors, positive lymph node status, and stage III/IV disease and lower proportions of black and younger women than CBCS, although the magnitude and direction of RFDs for the associations between patient and tumor characteristics and histologic subtypes were similar in CBCS and TCGA.

4.3.2 Ductal tumors are molecularly diverse.

IHC-based clinical subtype and RNA-based intrinsic subtype distributions by histologic subtype in CBCS3 and TCGA, are presented in Figure 4.1. Distributions of ductal, and other histologic subtype tumors for CBCS3 and TCGA were similar, but TCGA had a slightly higher percentage of lobular tumors (CBCS; 9%; TCGA:15%) and lower proportion of mixed tumors (CBCS: 16%; TCGA: 12%). Figure 4.2 displays the distributions of Luminal A and Triple Negative/Basal-like tumors by histologic subtype. Lobular tumors were predominantly Luminal A and proportions were similar by molecular subtyping method and study (CBCS IHC: 89%, RNA: 84%; TCGA IHC 86%, RNA:92%). In general, mixed tumors were more similar to lobular and were largely Luminal A subtype in CBCS and TCGA. Ductal tumors displayed more diversity in molecular subtype than lobular tumors. Proportions of ductal tumors with Luminal A subtype were similar between studies, but varied by technical method with lower percentages of ductal tumors classified as Luminal A by RNA than IHC in CBCS 3 and the TCGA (CBCS IHC: 58%, RNA: 39%; TCGA IHC 55%, RNA:42%). Ductal tumors from CBCS had higher proportions of Triple Negative/Basal-like subtype than TCGA by both IHC and RNA (CBCS IHC: 26%, RNA: 27%; TCGA IHC: 20%, RNA:23%).

4.3.3 Lobular tumors are predominantly Luminal A.

Figure 4.3 displays the relative frequency differences (RFD) and 95% confidence intervals (95% CI) for Luminal A compared to Non-Luminal A subtypes (clinical: Luminal B, TN, and HER2+; RNA, PAM50: Luminal B, Basal-like, HER2-enriched) among lobular and

mixed tumors compared to ductal tumors. Lobular tumors were more likely to be classified as any hormone receptor positive and HER2- (Luminal A clinical subtype), with similar magnitude RFDs in CBCS 1-3 and TCGA [CBCS: RFD (LumA vs. Non-LumA): 30.9%, 95% CI: (28.6, 33.2); TCGA: RFD 30.8%, 95% CI: (21.1, 40.4)]. The same association held for PAM50 subtype, the magnitude of effect was larger in TCGA than CBCS 3 [CBCS: RFD (LumA vs. Non-LumA): 44.9%, 95% CI (39.6, 50.1); TCGA: RFD: 50.5%, 95% CI (43.9, 57.1)]. Compared to ductal, mixed tumors were also more likely to be Luminal A subtype by IHC and PAM50, but the RFDs were attenuated in TCGA relative to CBCS.

4.3.4 Unique associations between patient and tumor characteristics for lobular tumors persisted among Luminal A tumors only in CBCS.

Based on the observed associations between histologic and Luminal A subtype, analyses were performed among Luminal A tumors only to determine if the observed associations between histologic subtype and tumor characteristics, race, and age persisted after accounting for intrinsic subtype (Table 4.2). When restricted to PAM50 Luminal A subtype, differences persisted for race and age with black and younger women less likely to be diagnosed with lobular or mixed as compared to ductal disease, but estimates were attenuated. Similarly, in RNA analyses restricted to Luminal A subtype, lobular tumors remained larger in size, less likely to be high combined grade, and were diagnosed at higher stages of disease than ductal tumors. Associations were similar in magnitude and direction by IHC-based Luminal A subtype in CBCS 1-3 (results not shown).

4.3.5 Lobular tumors have lower frequencies of TP53 mutation and TP53 pathway defects than ductal tumors.

Based on a TCGA report of fewer TP53 DNA mutations among lobular tumors relative to ductal tumors, we assessed the distribution of TP53 status by histologic subtype (Table 4.3). We evaluated TP53 mutant status by IHC (CBCS 3) and by RNA through use of a 52-gene signature (CBCS 3 and TCGA). Compared to ductal, lobular tumors were less likely to have TP53 mutant status [CBCS 3: IHC mutant vs wild-type RFD: -21.0%; 95% CI (-24.4 -17.6); RNA mutant-like vs wild-type-like RFD: -34.5% (-39.5, -29.4)] with larger

magnitude of association observed for lobular tumors in TCGA [RNA mutant-like vs wild-type-like RFD: -41.8% (-50.9, -32.8)].

Table 4.1a. Relative Frequency Differences (RFD) and 95% Confidence Intervals (95% CI) for the associations between age, race, and tumor characteristics by histologic subtype in CBCS 1-3

	CBCS1-3				
	Ductal	Lobular		Mixed	
	N (%*)	N (%*)	RFD (95% CI)[a]	N (%*)	RFD (95% CI) [a]
Race					
White	1,463 (76.7) (72.4)[b]	200 (83.9) (10.7)[b]	Referent	287 (83.9) (16.8) [b]	Referent
Black	1,393 (23.3) (80.7)[b]	126 (16.1) (7.5)[b]	-7.3 (-9.8, -4.8)	186 (16.1) (11.8)[b]	-7.3 (-9.4, -5.1)
Age at diagnosis					
>50	1,232 (62.7)	166 (71.2)	Referent	257 (74.2)	Referent
≤50	1,624 (37.2)	160 (28.8)	-8.4 (-11.5, -5.4)	216 (25.8)	-11.5 (-13.9, -9.0)
Tumor Size (cm)					
≤2	1,444 (58.8)	140 (44.8)	Referent	269 (64.6)	Referent
>2	1,348 (41.2)	176 (55.2)	14.0 (10.7, 17.4)	198 (35.4)	-5.8 (-8.5, -3.2)
Missing	64	10		6	
Tumor Grade^					
Low-Intermed	976 (53.7)	233 (97.7)	Referent	394 (95.4)	Referent
High	1,153 (46.3)	9 (2.9)	-43.4 (-45.2, -41.6)	29 (4.6)	-41.7 (-43.4, -40.0)
Missing	727	84		50	
Lymph Node Status					
Negative	1,717 (65.0)	187 (58.6)	Referent	270 (63.9)	Referent
Positive	1,110 (35.0)	135 (41.4)	6.5 (3.2, 9.7)	197 (36.1)	1.1 (-1.5, 3.8)
Missing	29	4		6	
AJCC Stage					
I, II	2,388 (88.1)	237 (75.6)	Referent	392 (86.8)	Referent
III, IV	407 (11.9)	81 (24.4)	12.6 (9.7, 15.4)	75 (13.2)	1.3 (-0.5, 3.2)
Missing	61	8		6	
ER+	1,641 (65.5)	284 (90.8)	Referent	419 (92.2)	Referent
ER-	1,164 (34.5)	39 (9.2)	-25.3 (-27.4, -23.2)	48 (7.9)	-26.6 (-28.3, -24.8)
Missing	51	3		6	
PR+	1,377 (54.3)	248 (76.4)	Referent	349 (76.8)	Referent
PR-	1,424 (45.7)	73 (23.6)	-22.0 (-24.9, -19.1)	116 (23.2)	-22.5 (-24.9, -20.1)
Missing	55	5		8	
HER2-	2,227 (83.3)	289 (93.1)	Referent	407 (88.6)	Referent
HER2+	482 (16.7)	20 (6.9)	-9.7 (-11.6, -7.9)	53 (11.4)	-5.2 (-7.1, -3.3)
Missing	147	17		13	

*All percentages weighted for sampling fractions. ^Grade unavailable for CBCS Phase 2
[a]Univariable model for each tumor characteristic and histologic subtype.
[b]Row percentages.

Table 4.1b. Relative Frequency Differences (RFD) and 95% Confidence Intervals (95% CI) for the associations between age, race, and tumor characteristics by histologic subtype in TCGA

	TCGA				
	Ductal	Lobular		Mixed	
	N (%*)	N (%*)	RFD (95% CI)[a]	N (%*)	RFD (95% CI)[a]
Race					
White	340 (84.6) (67.2)[b]	107 (92.2) (21.2)[b]	Referent	59 (93.7) (11.7)[b]	Referent
Black	62 (15.4) (82.7)[b]	9 (7.8) (12.0)[b]	-7.6 (-13.7, -1.7)	4 (6.4) (5.3)[b]	-9.1 (-16.1, -2.1)
Age at diagnosis					
>50	320 (66.1)	98 (77.2)	Referent	70 (81.4)	Referent
≤50	164 (33.9)	29 (22.8)	-11.1 (-19.5, -2.6)	16 (18.6)	-15.3 (-24.5, -6.1)
Tumor Size (cm)					
≤2	134 (29.2)	21 (16.7)	Referent	30 (36.1)	Referent
>2	325 (70.8)	105 (83.3)	12.5 (4.8, 20.3)	53 (63.9)	-7.0 (-18.1, 4.2)
Missing	25	1		3	
Lymph Node Status					
Negative	230 (48.4)	54 (42.9)	Referent	35 (40.7)	Referent
Positive	245 (51.6)	72 (57.1)	5.6 (-4.2, 15.3)	51 (59.3)	7.7 (-3.6, 19.0)
Missing	9	1		0	
AJCC Stage					
I, II	372 (78.3)	79 (62.7)	Referent	59 (71.2)	Referent
III, IV	103 (21.7)	47 (37.3)	15.6 (6.4, 24.8)	24 (28.9)	7.2 (-3.2, 17.7)
Missing	9	1		3	
ER+	321 (70.7)	113 (94.2)	Referent	73 (91.3)	Referent
ER-	133 (29.3)	7 (5.8)	-23.5 (-29.4, -17.5)	7 (8.8)	-20.6 (-28.0, -13.1)
Missing	30	7		6	
PR+	278 (61.4)	97 (81.5)	Referent	65 (81.3)	Referent
PR-	175 (38.6)	22 (18.5)	-20.1 (-28.4, -11.9)	15 (18.8)	-19.9 (-29.5, -10.2)
Missing	31	8		6	
HER2-	240 (74.8)	68 (89.5)	Referent	42 (70.0)	Referent
HER2+	81 (25.2)	8 (10.5)	-14.7 (-23.1, -6.3)	18 (30.0)	4.8 (-7.8, 17.3)
Missing	163	51		26	

[a]Univariable model for each tumor characteristic and histologic subtype.
[b]Row percentages.

Figure 4.1. Histologic and clinical subtype distributions in CBCS 1-3 and TCGA and histologic and PAM50 intrinsic subtype distributions in CBCS 3 and TCGA.

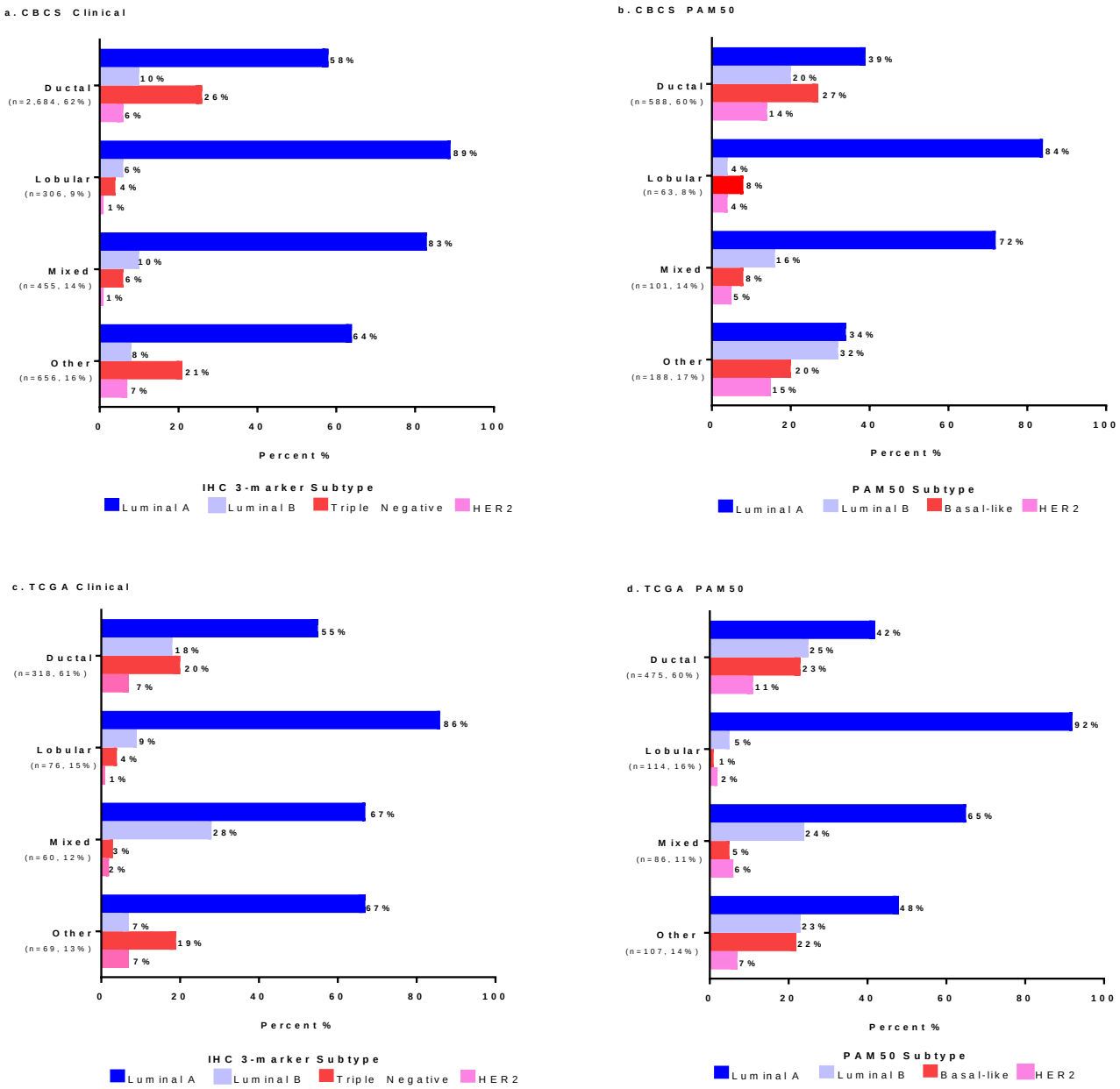

Figure 4.2. Clinical and PAM50 subtype distributions for Luminal A and Triple Negative/Basal-like subtype by histologic subtype in CBCS and TCGA.

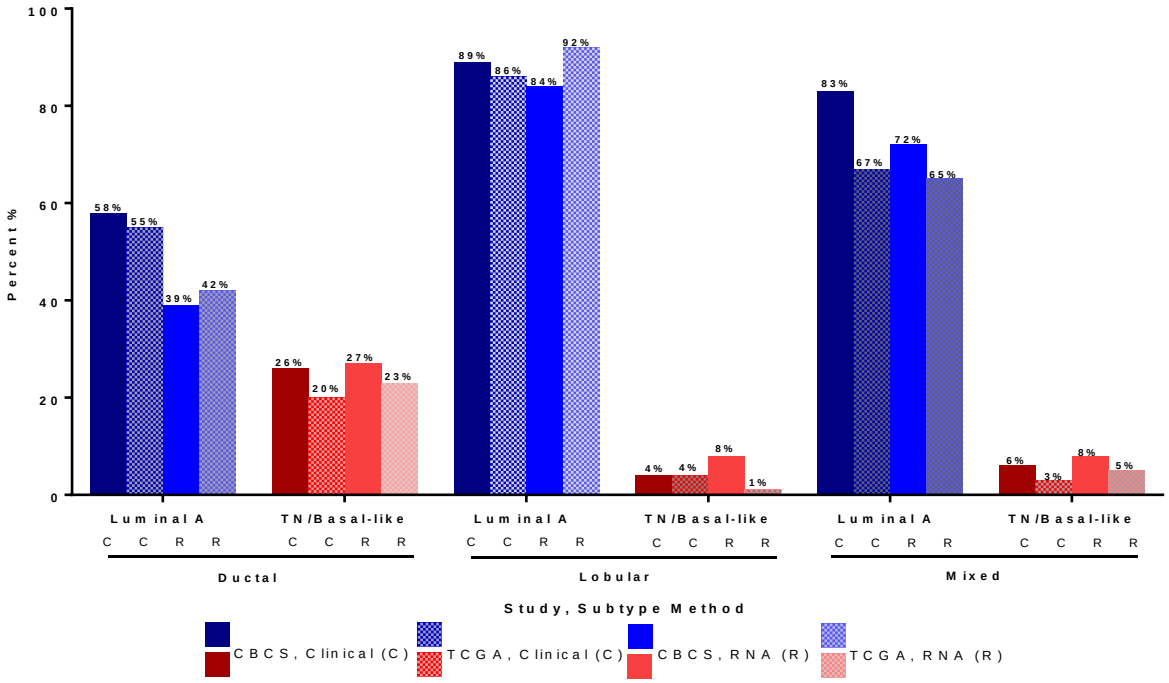

Figure 4.3. Relative Frequency Differences (RFD) and 95% Confidence Intervals (95% CI) of Luminal A subtype among lobular and mixed tumors compared to ductal tumors in CBCS and TCGA

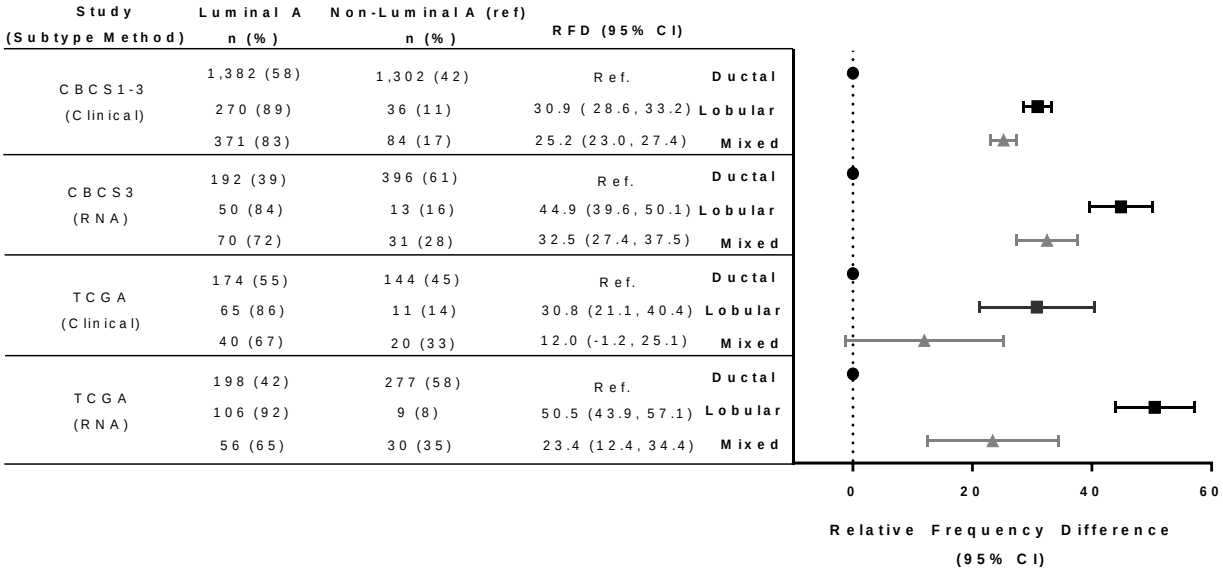

Table 4.2. Relative Frequency Differences (RFD) and 95% Confidence Intervals (95% CI) for the associations between patient and clinical characteristics by histologic subtype among Luminal A intrinsic subtype tumors, CBCS 3

	PAM50 Luminal A				
	Ductal	Lobular		Mixed	
	N (%*)	N (%*)	RFD (95% CI)	N (%*)	RFD (95% CI)
Race					
White	119 (84.0)	35 (89.0)	Referent	47 (89.0)	Referent
Black	73 (16.0)	11 (11.0)	-5.0 (-10.2, 0.0)	23 (11.0)	-5.0 (-10.0, -0.4)
Age at diagnosis					
>50	106 (72.6)	30 (78.8)	Referent	42 (79.8)	Referent
≤50	86 (27.4)	20 (21.2)	-6.2 (-12.9, 0.0)	28 (20.2)	-7.3 (-13.0, -1.5)
Tumor Size (cm)					
≤2	124 (70.7)	16 (32.7)	Referent	41 (60.7)	Referent
>2	68 (29.3)	34 (67.3)	38.0 (30.6, 45.5)	29 (39.3)	10.1 (3.4, 16.7)
Missing					
Tumor Grade					
Low-Intermed	144 (81.1)	48 (95.8)	Referent	67 (97.9)	Referent
High	48 (18.9)	2 (4.2)	-14.7 (-18.9, -10.7)	3 (2.1)	-16.8 (-20.2, -13.3)
Missing					
Lymph Node					
Negative	126 (69.5)	31 (57.5)	Referent	44 (35.5)	Referent
Positive	66 (30.5)	19 (42.5)	12.1 (4.3, 19.8)	26 (64.5)	5.0 (-1.6, 11.6)
Missing					
AJCC Stage					
I, II	168 (89.9)	35 (64.2)	Referent	65 (94.0)	Referent
III, IV	24 (10.1)	15 (35.8)	25.7 (18.6, 32.8)	5 (6.0)	-4.1 (-7.7, -0.5)
Missing					

*All percentages weighted for the CBCS sampling fractions.

Table 4.3. Relative Frequency Differences (RFD) and 95% Confidence Intervals (95% CI) for TP53 mutant status (IHC) and mutant-like status (RNA) among lobular and mixed compared to ductal invasive breast tumors in CBCS 3 and TCGA

	Ductal	Lobular		Mixed	
	N (%)	N (%)	RFD (95% CI)	N (%)	RFD (95% CI)
			CBCS3*		
TP53 IHC					
wild-type	482 (73.1)	73 (94.2)	Referent	154 (87.2)	Referent
mutant	213 (26.9)	6 (5.8)	-21.0 (-24.4, -17.6)	22 (12.8)	-14.1 (-17.3, -10.9)
missing	952	103		257	
TP53 RNA					
wild-type-like	260 (50.6)	57 (85.1)	Referent	90 (81.6)	Referent
mutant-like	346 (49.4)	10 (14.9)	-34.5 (-39.5, -29.4)	20 (18.4)	-31.0 (-35.4, -26.7)
missing	1,041	115		191	
			TCGA		
TP53 RNA					
wild-type-like	141 (35.2)	87 (77.0)	Referent	40 (64.5)	Referent
mutant-like	260 (64.8)	26 (23.0)	-41.8 (-50.9, -32.8)	22 (35.5)	-29.4 (-42.2, -16.6)
missing	83	14		24	

*Percentages weighted for sampling design

4.4 Discussion

Characteristics of lobular and mixed tumors were quantitatively different from those of ductal tumors in CBCS and TCGA. In both studies, lobular tumors were significantly more likely to be Luminal A and have lower frequencies of TP53 pathway defects than ductal tumors. As previously reported, we found lobular disease to be more common among older and white women [17, 21, 25, 26], more likely to be low-intermediate tumor grade, larger in size, diagnosed at later stage of disease [3, 5, 7–18, 21, 24–26, 37, 95]. However, we are the first to show that the associations between patient and tumor characteristics and lobular tumors persisted among Luminal A subtype only, suggesting histology contributes to these observed associations even after restricting to the dominant lobular intrinsic subtype.

By comparing CBCS and TCGA, we observed associations between lobular and ductal histology and Luminal A subtype to be quantitively similar between studies, but vary by technical method. We found that lobular tumors were 31% more likely to be Luminal A by IHC (CBCS and TCGA) and 45-51% more likely to be Luminal A by RNA (CBCS3 and TCGA, respectively). In a sample of 75 lobular tumors from the I-SPY trial, Lips *et al.* (2012) determined PAM50 subtypes and observed an RFD of 44% for Luminal A subtype among lobular compared to ductal tumors, which is similar to our estimates [16]. Other studies using IHC have reported RFDs for Luminal A subtype among lobular compared to ductal tumors ranging from 6%-39%, reflecting tremendous molecular variability in the tumors studied [3, 6, 8–10, 12–15, 111]. Proportions of Luminal A ductal tumors were overestimated by clinical data relative to RNA, suggesting RNA subtype may be important for understanding the differences in Luminal A prevalence among ductal tumors.

In our study, we observed higher proportions of TP53 mutant-like tumors (RNA) than TP53 mutations (IHC) in each histologic subtype, indicating a higher frequency of TP53 pathway defects that may be important for tumor etiology and progression in a subset of tumors. We found that compared to ductal tumors, lobular tumors are less frequently TP53 mutant by IHC, as has been previously reported [17, 25, 26], and less frequently TP53

mutant-like by RNA in CBCS 3 and TCGA. Nearly 50% of CBCS 3 ductal tumors were TP53 mutant-like by RNA, suggesting that TP53 pathway deficiencies may be integral to the development of the ductal phenotype. Relative to CBCS, TCGA had higher frequencies of mutant-like TP53 among all histologic subtypes, reflecting the higher proportion of more advanced tumors in TCGA.

Our study should be interpreted in light of some limitations. We included a mixed histology category of tumors, defined as <80% of a dominant phenotype and ≥20% of a second histologic subtype. This practice has been used in previous studies based on expert reviews, but some studies have adjudicated mixed tumors, classifying them as either lobular or ductal. For example, in TCGA, an expert panel sought consensus for lobular and ductal, and then found that mixed tumors were either lobular-like or ductal-like based on mRNA gene expression and other genomic alterations, including E-cadherin mutation status [5]. Future work should consider adjudicating the expert pathology review with molecular validation. Such approaches, if validated, could increase sample size of ductal and lobular tumors for etiologic and survival analyses. A second limitation of our work was that we were unable to study molecular characteristics of the rare histologic subtypes diagnosed in less than 2% of cases in CBCS due to power constraints.

4.5 Conclusions

To conclude, this analysis documents that while The Cancer Genome Atlas is not a population-based study of invasive breast tumors, the associations between the tumor and the molecular characteristics are similar to those from the population-based, Carolina Breast Study. Furthermore, patterns of association between tumor characteristics and histology were similar when restricting to Luminal A subtype, suggesting histology reflects some true biological differences, some of which have already been identified. Future research may leverage the molecular differences between lobular, ductal, and mixed tumors to improve histologic classification and should seek to identify how risk and survival vary for fully characterized subgroups of lobular tumors.

CHAPTER 5: REPRODUCTIVE RISK FACTOR ASSOCIATIONS WITH LOBULAR AND DUCTAL CARCINOMA IN THE CAROLINA BREAST CANCER STUDY

5.1 Introduction

The intrinsic breast cancer subtypes, including Luminal A, Luminal B, HER2-enriched, and Basal-like cancers, show distinct risk factor profiles and are hypothesized to be independent diseases within the breast. Ductal tumors, diagnosed in up to 80% of cases [18, 51, 52], are comprised of approximately 50% Luminal A tumors, and while Basal-like and HER2-enriched intrinsic subtype tumors represent a minority of ductal tumors, the vast majority of Basal-like and HER2-enriched tumors are ductal [3, 5, 6, 8–10, 12–15, 111]. Lobular and mixed ductal-lobular breast cancers, diagnosed in up to 15% of cases each [18, 51, 52], tend to be 80-90% Luminal A intrinsic subtype [3, 5, 6, 8–10, 12–15, 18, 51, 52, 111]. Thus, there are strong associations between histologic and molecular subtype.

Previous studies of etiologic heterogeneity according to histology have suggested that lobular disease is more strongly associated with a number of reproductive risk factors and hormone-modulating exposures. However, it is unclear if the observed associations depend on intrinsic breast cancer subtypes, which also have unique reproductive risk factor profiles [4, 6]. We sought to disentangle the associations between reproductive breast cancer risk factors and breast cancer subtype, considering both histology and Luminal A subtype in the Carolina Breast Cancer Study Phases 1-3 (1993-2013).

5.2 Methods

5.2.1 Study population

The present analysis includes 3,655 cases of invasive breast cancer from the Carolina Breast Cancer Study (CBCS) Phases 1-3 (1993-2013). The CBCS is a population-

based study among black and white women, initiated in 1993, that recruited participants from 24 (CBCS 1-2) to 44 (CBCS 3) of the 100 North Carolina counties [137]. CBCS oversampled women less than 50 years of age and black women [3, 82].

For CBCS 1-3, case eligibility criteria included: women with a first diagnosis of breast cancer [(invasive or *in situ* (CBCS2 only)], aged 20-74 years at diagnosis, and residence in specified counties. Cases were enrolled following rapid case ascertainment from the NC Central Cancer Registry and controls (CBCS 1-2) were identified using DMV and Medicare lists. Controls were race and age frequency matched to cases (CBCS 1-2). All participants provided informed consent for study enrollment and cases granted access to tumor tissue blocks/slides and medical records from treatment centers. Self-report, risk factor data was collected during in-person interviews by a trained study nurse. Cases eligible for this analysis had invasive tumor tissue available for centralized pathology review to determine histologic subtype as ductal (n=2,856), lobular (n=326), or mixed ductal/lobular (n=473) (henceforth referred to as mixed) breast cancer. The study maintains Institutional Review Board approval at the University of North Carolina.

5.2.2 Histologic subtype

Histologic subtype for CBCS1-3 was determined via centralized pathologist review. Tumors classified as ductal, lobular, and mixed ductal-lobular comprise 84% of all CBCS 1-3 cases with histologic subtype available. Ductal or lobular histologic subtypes tumors were defined as at least 80% representative of that histology. Mixed tumors contained ≥20% of one histologic subtype and <80% of the second histologic subtype. The following histologic subtypes were excluded: mixed ductal/non-lobular (n=285), mucinous (n=89), mixed ductal/metaplastic (n=63), metaplastic (n=44), DCIS w/focal invasion (n=44), undifferentiated high grade (n=29), tubular (n=23), micropapillary (n=21), papillary (n=19), medullary (n=18), pleomorphic lobular (n=17), anaplastic (n=14), apocrine (n=11), cribriform (n=9), neuroendocrine (n=3), others (n=15). Cases with unknown (n=99) or missing (n=376) were also excluded.

5.2.3 Immunohistochemistry (IHC)-based clinical breast cancer subtypes

For CBCS 1-2, estrogen receptor (ER) and progesterone receptor (PR) status were abstracted from medical records for 80% of cases, the remaining 20% with tumor tissue available had IHC for ER and PR performed at UNC. For tissue that was stained at UNC, a study pathologist determined ER and PR positivity defined using contemporaneous clinical cut points [3]. HER2 staining was performed at UNC for all CBCS 1-2 cases with available tissue as described previously with positivity defined as membrane/membrane plus cytoplasmic staining classified as weak or greater intensity in ≥10% of tumor cells [3].

In CBCS 3, 98% cases had ER, PR, and HER2 information in their medical records. For the remaining 2% of cases without ER, PR, and HER2 data, IHC staining was performed at UNC with positivity cut points of ≥10% for ER and PR. HER2 positivity was defined as 3+ staining intensity [negativity was defined as 0/1+ (cases with equivocal staining, 2+ staining intensity, were excluded)] as described in Allott *et al.* (2016) [139].

For CBCS 1-3, 3-marker clinical subtypes were defined as follows: Luminal A (ER+ or PR+ and HER2-), Luminal B (ER+ or PR+ and HER2+), Triple Negative (TN) (ER- and PR- and HER2-), and HER2+ (ER- and PR- and HER2+).

5.2.4 Statistical analysis

Patient characteristics

The associations between histologic subtype and race, age, menopausal status, and clinical subtype were estimated using generalized linear models that were adjusted for age, race, and study phase (1, 2, 3), where appropriate. Relative frequency differences (RFDs) and 95% confidence intervals (95% CIs) were estimated as the measure of association [144]. To account for the CBCS sampling design, weighted percentages are presented alongside unweighted sample size counts. Patient characteristics were defined as: race [self-report: black, non-black (>98% white, henceforth referred to as white)], age (years)

(<40, 40-49, 50-59, ≥60), menopausal status (pre-, post-), and clinical subtype as defined above.

Reproductive risk factor analyses

The association between each reproductive breast cancer risk factor and histologic subtype was estimated in case-control (CBCS 1-2) and case-case (CBCS 1-3) analyses. Polytomous logistic regression was used to obtain odds ratios (ORs) and 95% confidence intervals (95% CIs) as the measure of association. The following risk factors were studied in association with breast cancer subtype: parity (nulliparous, 1, 2, ≥3), years since last birth (defined as: age at study enrollment minus age at last birth; among parous women only) (0-≤10, 10-≤20, >20), age at first live birth (years) (parous women only) (<26, ≥26), lifetime lactation duration (months) (parous women only) (never, 0-≤12, >12), oral contraceptive use (never, former, current), and hormone therapy (HT) use [never, estrogen alone, combined estrogen + progesterone (E+P)]. Additional variables used in case-control analyses included: study phase (1, 2, 3), age continuous (20-74), family history of breast cancer (yes, no), alcohol intake (ever, never), smoking duration (years) (never, <10, 11-19, ≥20), oral contraceptive use (ever, never), breastfeeding (ever, never), age at menarche (years) (<13, ≥13), and the offset term to account for the sampling design of CBCS. Case-case analyses are presented for CBCS 1-3 to assess etiologic heterogeneity by histologic subtype [4, 145]. In case-case analyses, ductal served as the referent group compared to lobular and mixed. Case-case models are adjusted for age, race, and study phase.

We tested the null hypothesis that the slope of the line equals zero for age at diagnosis, parity, and lactation duration, each modeled as continuous variables. We conducted a race-stratified sensitivity analysis of the risk factor associations with histologic subtype, but we were unable to assess racial differences in oral contraceptive and HT use due to low sample size among black women for current OC use and combined E+P use. All

analyses were done in SAS version 9.4 (SAS Institute, Cary, NC). P-values were produced for a two-sided test with an alpha of 0.05 for statistical significance.

5.3 Results

5.3.1 Compared to ductal tumors, lobular tumors were less frequent among black women and younger women and were predominantly Luminal A clinical subtype.

Women from the Carolina Breast Cancer Study Phases 1-3, displayed patterns consistent with established histological associations by age, race, menopausal status, and clinical subtype (Table 5.1). Relative to ductal histologic subtype, lobular and mixed tumors were less frequent among young women and black women, and in age, race, and study phase-adjusted analysis, lobular and mixed tumors were more frequent among premenopausal women. As other studies have shown, lobular and mixed tumors were predominantly Luminal A clinical subtype (ER+ or PR+ and HER2-) (lobular 88.8%, mixed 83.1%); whereas ductal tumors are less frequently Luminal A (57.9%). After adjusting for age, race, and study phase, associations with molecular subtype were statistically significant, with lobular tumors significantly more likely to be Luminal A clinical subtype [Relative Frequency Difference (RFD) compared to ductal: 26.3%, 95% CI (24.0, 28.5)].

5.3.2 Compared to ductal tumors, lobular tumors have unique risk factor patterns.

We observed unique risk factor patterns for lobular tumors relative to ductal tumors (Table 5.2). Ductal tumors were inversely associated with parity, increasing lactation duration, and estrogen-only hormone therapy (HT) use. Among lobular tumors, parity was inversely associated with having 1 child versus being nulliparous, but the association was attenuated as parity increased. Among parous women, we observed a positive association between age ≥26 years at first birth versus <26 years and lobular disease [Lobular OR: 1.32; 95% CI (0.86-2.03)] and a null effect for ductal tumors [Ductal OR: 0.94; 95% CI (0.77-1.16)]. Lifetime lactation duration >12 months was positively associated with lobular disease [Lobular OR: 1.62; 95% CI (0.99-2.67)] and inversely associated for ductal disease [Ductal OR: 0.78, 95% CI (0.60-1.02)]. Former oral contraceptive (OC) use was associated

with lobular disease [Lobular OR: 1.43, 95% CI (0.92-2.22)] but not ductal disease [Ductal OR: 0.96, 95% CI (0.79-1.71)]. Associations with hormone therapy use were stronger for lobular compared to ductal disease, with estrogen alone and having a larger inverse association with lobular disease [OR: 0.59, 95% CI (0.33-1.06)] and combined estrogen plus progesterone (E+P) hormone therapy (HT) use having a larger positive association with lobular disease [OR: 1.74, 95% CI (0.99-3.06)].

5.3.3 Etiologic differences emerged for lobular tumors relative to ductal tumors.

To assess whether these risk factor patterns were indicative of significant etiologic differences between ductal (referent) and lobular disease, we conducted case-case analyses. Case-case analyses showed a statistically significant difference in the associations between lobular and ductal disease for age ≥26 years at first birth [OR: 1.35, 95% CI (1.03-1.78)], lifetime lactation duration >12 months [OR: 1.86, 95% CI (1.33-2.60)], and current OC use [OR: 1.86, 95% CI (1.08-3.20)] (Table 5.3). These associations did not appear to differ by race (all p-values for heterogeneity >0.50) (Table 5.4). We also observed that associations for mixed tumors were typically intermediate in magnitude between the estimates for lobular and ductal disease (Table 5.2 and 5.3).

5.3.4 Etiologic differences remained between lobular tumors relative to ductal tumors after restricting to Luminal A subtype.

To address our main research question of whether risk factor-histologic subtype associations persist after restricting to Luminal A clinical subtype, which is not evenly distributed by histologic subtype, we performed case-case analyses in CBCS 1-3 that were restricted to Luminal A tumors only (Table 5.5). After restricting to Luminal A clinical subtype, associations for lobular disease relative to ductal were similar in direction and magnitude to the overall case-case risk factor associations for lobular tumors, particularly for age at first birth ≥26 years [OR: 1.26, 95% CI: 0.82-1.93)], lactation duration >12 months [OR: 1.51, 95% CI (1.02-2.25)], and oral contraceptive use [current OR: 1.82, 95% CI (0.99-3.36); former OR: 1.48, 95% CI (1.06-2.06)].

Table 5.1. Relative Frequency Differences (RFD) and 95% Confidence Intervals (95% CI) for the associations between race, age at diagnosis, menopausal status and clinical subtype comparing lobular and mixed ductal-lobular to ductal histologic subtype breast tumors, CBCS 1-3

	Ductal	Lobular		Mixed	
Risk factor	N (%*)	N (%*)	RFD (95% CI)	N (%*)	RFD (95% CI)
Age at diagnosis[a]					
≥60	697 (37.4)	109 (46.1)	Ref	159 (48.8)	Ref
50-59	591 (28.1)	65 (27.6)	-3.5 (-4.3, -0.7)	107 (27.2)	-3.6 (-5.7, -1.5)
40-49	1,105 (24.7)	131 (23.0)	-3.9 (-4.7, -0.2)	160 (19.0)	-5.9 (-8.0, -3.7)
<40	463 (9.8)	21 (3.3)	-8.8 (-10.8, -	47 (5.0)	-8.5 (-11.2, -5.8)
Race[b]					
White	1,463 (76.7)	200 (83.9)	Ref	287 (83.9)	Ref
Black	1,393 (23.3)	126 (16.1)	-3.5 (-5.0, -2.0)	186 (16.1)	-4.4 (-6.5, -2.2)
Menopausal Status[c]					
Pre	1,400 (33.2)	141 (26.9)	2.0 (0.0, 4.0)	201 (26.5)	3.9 (0.8, 6.9)
Post	1,456 (66.8)	185 (73.1)	Ref	272 (73.5)	Ref
Clinical Subtype[c,d]					
Luminal A	1382 (57.9)	270 (88.8)	26.3 (24.0,28.5)	371 (83.1)	20.8 (18.6, 23.1)
Non-Luminal A	1302 (41.3)	36 (11.2)	Ref.	84 (16.9)	Ref.
Luminal B	282 (10.3)	16 (6.3)		43 (10.0)	
HER2+	193 (6.3)	4 (0.7)		9 (1.4)	
Triple Negative	827 (25.5)	16 (4.2)		32 (5.5)	
Missing	172	20		18	

*Percentages weighted for sampling fractions.
[a]Adjusted for race (white, black) and study phase (1, 2, 3).
[b]Adjusted for age (continuous) and study phase.
[c]Adjusted for age, race, and study phase.
[d]Luminal A (ER+ or PR+/HER2-); Non-Luminal A [Luminal B (ER+ or PR+/HER2+), Triple Negative (ER-/PR-/HER2-), HER2+ (ER-/PR-/HER2+)]

Table 5.2. Case-Control Odds Ratios (OR) and 95% Confidence Intervals (95% CI) for the associations between risk factors and ductal, lobular and mixed histologic subtype breast tumors compared to controls, CBCS 1-2

	Controls	Ductal		Lobular		Mixed	
Risk factor	N (%*)	N (%*)	OR (95% CI)	N (%*)	OR (95% CI)	N (%*)	OR (95% CI)
Parity[a]							
Nulliparous	174 (22.6)	193 (15.0)	Ref	19 (10.9)	Ref	13 (10.7)	Ref
1	281 (16.8)	219 (16.9)	0.80 (0.60-1.06)	22 (12.0)	0.77 (0.39-1.52)	19 (17.6)	0.90 (0.42-1.93)
2	495 (32.8)	372 (32.3)	0.83 (0.63-1.08)	44 (32.3)	0.87 (0.47-1.60)	37 (33.9)	1.03 (0.52-2.06)
≥3	614 (27.8)	425 (35.8)	0.85 (0.65-1.12)	59 (44.8)	1.02 (0.55-1.88)	37 (37.8)	0.96 (0.46-1.97)
p-value			0.53		0.63		0.96
Years since last live birth (parous only)[b]							
0-≤10	194 (32.0)	182 (12.1)	0.77 (0.54-1.10)	12 (5.8)	0.79 (0.33-1.93)	15 (9.4)	1.05 (0.42-2.66)
>10-≤20	365 (24.9)	271 (19.5)	0.93 (0.72-1.22)	40 (25.3)	1.57 (0.86-2.86)	18 (11.2)	0.80 (0.38-1.66)
>20	829 (43.0)	562 (68.4)	Ref.	73 (68.8)	Ref	59 (79.3)	Ref
Missing	2	1		0		1	
Age at first live birth (parous only; years)							
<26	1050	760 (73.4)	Ref.	85 (71.0)	Ref	61 (59.2)	Ref
≥26	335 (29.7	247 (26.6)	0.94 (0.77-1.16)	39 (29.0)	1.32 (0.86-2.03)	32 (30.1)	1.63 (1.01-2.62)
Missing	5	9		1		1	
Lactation duration (parous only)[c]							
Never	794 (56.1)	622 (58.7)	Ref	67 (48.5)	Ref	54 (57.3)	Ref
>0-≤12	408 (31.6)	272 (28.9)	0.90 (0.66-0.98)	30 (28.2)	0.69 (0.43-1.12)	24 (31.5)	0.77 (0.46-1.29)
>12	186 (12.2)	121 (12.4)	0.78 (0.60-1.02)	28 (23.3)	1.62 (0.99-2.67)	14 (11.2)	1.05 (0.57-1.97)
Missing	2	1		0		1	
p-value			0.02		0.24		0.79
Oral Contraceptive Use[d]							
Never	572 (23.5)	412 (39.4)	Ref	48 (28.5)	Ref	37 (44.6)	Ref
Current	76 (4.9)	77 (6.4)	1.01 (0.68-1.50)	4 (2.8)	0.86 (0.28-2.71)	4 (3.8)	0.82 (0.26-2.66)
Former	905 (58.3)	716 (59.4)	0.96 (0.79-1.17)	91 (63.6)	1.43 (0.92-2.22)	65 (61.3)	1.15 (0.69-1.89)
Missing	11	4		1		0	
Hormone Therapy Use[e]							
Never	1080	893 (64.6)	Ref	98 (53.1)	Ref	72 (55.1)	Ref
Estrogen	307 (13.8)	164 (16.8)	0.68 (0.54-0.86)	18 (19.6)	0.59 (0.33-1.06)	21 (31.6)	0.90 (0.51-1.60)
Combined	149 (8.2)	126 (18.6)	1.20 (0.89-1.60)	25 (27.3)	1.74 (0.99-3.06)	11 (13.3)	1.00 (0.48-2.06)
Missing	28	26		3		2	

*All percentages weighted for study sampling design. [a]Adjusted for race (black, white), age (continuous), study phase (1, 2, 3), family history (yes, no), alcohol intake (ever, never), smoking duration (never, <10 years, 11-19, ≥20), oral contraceptive use (ever, never) breastfeeding (ever, never), menopausal status (pre-, post-) age at menarche (<13, ≥13), CBCS offset term.[b]Among parous women only. Adjusted for race, age, study phase, menopausal status, family history, parity, alcohol intake, smoking duration, oral contraceptive use, age at menarche, CBCS offset term. [c]Among parous women only. Adjusted for race, age, study phase, menopausal status, family history, alcohol intake, smoking duration, oral contraceptive use, age at menarche, CBCS offset term. [d]Adjusted for race, age, study phase, family history, alcohol intake, smoking duration, oral contraceptive use, parity (nulliparous, 1-2, ≥3), breastfeeding, age at menarche, CBCS offset term. [e]Adjusted for race, age, study phase, family history, alcohol intake, smoking duration, oral contraceptive use, parity, breastfeeding, menopausal status, age at menarche, CBCS offset term.

Table 5.3. Case-Case Odds Ratios (OR) and 95% Confidence Intervals (95% CI) for the associations between risk factors comparing lobular and mixed to ductal histologic subtype breast tumors, CBCS 1-3

	Ductal	Lobular		Mixed	
Risk factor	N (%*)	N (%*)	OR (95% CI)	N (%*)	OR (95% CI)
Parity[a]					
Nulliparous	428 (14.4)	47 (12.6)	Ref	72 (14.0)	Ref
1	560 (19.8)	46 (11.7)	0.75 (0.49-1.15)	71 (12.6)	0.72 (0.50-1.03)
2	937 (34.7)	123 (43.4)	1.14 (0.80-1.63)	173 (39.7)	1.00 (0.74-1.35)
≥3	931 (31.1)	110 (32.4)	1.01 (0.70-1.46)	157 (33.7)	0.97 (0.71-1.33)
p-value			0.82		0.96
Years since last live birth (parous only)[a]					
0-≤10	480 (13.3)	34 (6.8)	1.16 (0.68-1.99)	69 (9.5)	1.77 (1.14-2.76)
>10-≤20	612 (18.2)	84 (22.6)	1.80 (1.23-2.64)	92 (16.5)	1.49 (1.05-2.10)
>20	1330 (68.6)	161 (70.6)	Ref	238 (74.1)	Ref
Missing	6	0		2	
Age at first live birth (parous only; years)[a]					
<26	1710 (67.4)	180 (66.2)	Ref.	259 (67.5)	Ref.
≥26	702 (32.6)	98 (33.8)	1.35 (1.03-1.78)	141 (32.5)	1.25 (0.98-1.58)
Missing	16	1		1	
Lactation duration (parous only)[a]					
Never	1399 (55.1)	136 (43.6)	Ref	206 (52.1)	Ref
>0-≤12	714 (30.9)	84 (32.6)	1.19 (0.89-1.59)	126 (32.7)	1.13 (0.88-1.45)
>12	313 (14.1)	59 (23.8)	1.86 (1.33-2.60)	68 (15.3)	1.36 (1.00-1.86)
Missing	2	0		1	
p-value					0.04
Oral Contraceptive Use[a]					
Never	756 (27.5)	82 (23.4)	Ref	117 (25.8)	Ref
Current	193 (5.3)	24 (5.0)	1.86 (1.08-3.20)	30 (5.1)	1.19 (0.74-1.92)
Former	1889 (67.2)	217 (69.6)	1.33 (0.99-1.78)	324 (69.1)	1.09 (0.86-1.40)
Missing	18	3		2	
Hormone Therapy Use[a]					
Never	2196 (68.0)	229 (60.3)	Ref	342 (59.8)	Ref
Estrogen alone	377 (17.7)	49 (21.0)	0.97 (0.69-1.38)	84 (27.8)	1.13 (0.85-1.51)
Combined E+P	229 (14.3)	41 (18.7)	1.25 (0.85-1.83)	38 (12.4)	0.88 (0.60-1.30)
Missing	54	7		3	

*All percentages weighted for study sampling design.

[a] Adjusted for race, age, study phase.

Table 5.4. Sensitivity analysis case-case Odds Ratios (OR) and 95% Confidence Intervals (95% CI) for the associations between risk factors comparing lobular and mixed ductal-lobular to ductal histologic subtype breast tumors stratified by race, CBCS 1-3

	Ductal	Lobular		Mixed	
Risk factor	N (%*)	N (%*)	OR (95% CI)	N (%*)	OR (95% CI)
			White women		
Parity[a]					
Nulliparous	244 (14.9)	30 (12.5)	Ref	40 (13.7)	Ref
1	302 (20.1)	27 (11.1)	0.72 (0.41-1.24)	42 (12.0)	0.83 (0.52-1.34)
2	556 (37.0)	88 (46.0)	1.23 (0.79-1.91)	122 (42.2)	1.26 (0.85-1.87)
≥3	361 (28.0)	55 (30.4)	1.05 (0.65-1.71)	83 (32.1)	1.26 (0.82-1.92)
p-value			0.48		0.36
Years since last live birth (parous only)[b]					
0-≤10	267 (12.9)	20 (6.2)	1.06 (0.52-2.17)	45 (9.0)	2.03 (1.13-3.67)
>10-≤20	313 (17.1)	54 (22.3)	1.93 (1.16-3.20)	61 (16.3)	1.80 (1.13-2.87)
>20	626 (70.0)	96 (71.5)	Ref	139 (74.7)	Ref
Missing	3	0		2	
Age at first live birth (parous only; years)[b]					
<26	723 (62.7)	99 (64.5)	Ref	142 (65.8)	Ref
≥26	489 (37.3)	70 (35.5)	1.21 (0.86-1.70)	104 (34.2)	1.16 (0.86-1.55)
Missing	7	1		1	
Lactation duration (months)[b]					
Never	600 (51.4)	68 (39.7)	Ref	112 (50.2)	Ref
>0-≤12	418 (33.0)	60 (34.8)	1.40 (0.96-2.03)	84 (33.7)	1.11 (0.81-1.53)
>12	201 (15.6)	52 (24.5)	2.12 (1.39-3.25)	50 (16.1)	1.37 (0.93-2.01)
Missing	0	0		1	
p-value			0.03		0.05
			Black Women		
Parity[a]					
Nulliparous	184 (12.6)	17 (12.9)	Ref.	32 (15.7)	Ref.
1	258 (18.8)	19 (14.8)	0.80 (0.40-1.59)	29 (15.8)	0.57 (0.33-0.99)
2	381 (27.3)	35 (29.6)	1.00 (0.54-1.84)	51 (26.4)	0.68 (0.42-1.10)
≥3	570 (41.3)	55	0.96 (0.54-1.70)	74 (42.2)	0.69 (0.44-1.09)
p-value			0.43		0.59
Years since last live birth (parous only)[b]					
0-≤10	213 (14.5)	14 (10.2)	1.38 (0.61-3.15)	24 (11.9)	1.53 (0.78-3.02)
>10-≤20	299 (21.6)	30 (23.8)	1.65 (0.92-2.98)	31 (17.6)	1.15 (0.68-1.97)
>20	694 (64.0)	65 (66.0)	Ref	99 (70.5)	Ref
Missing	3	0		0	
Age at first live birth (parous only; years)[b]					
<26	987 (82.3)	81 (74.6)	Ref	117 (76.6)	Ref
≥26	213 (17.7)	28 (25.2)	1.69 (1.07-2.67)	37 (23.4)	1.45 (0.97-2.17)
Missing	9	0		0	
Lactation duration (months)[b]					
Never	799 (66.9)	68 (63.6)	Ref.	94 (61.9)	Ref.
>0-≤12	296 (23.9)	24 (21.4)	0.95 (0.58-1.54)	42 (27.4)	1.18 (0.79-1.74)
>12	112 (9.2)	17 (15.0)	1.63 (0.91-2.90)	18 (10.7)	1.37 (0.79-2.38)
Missing	2	0		0	
p-value			0.44		0.42

*All percentages weighted for study sampling design.

[a] Adjusted for age (continuous) and study phase (1, 2, 3),

[b] Among parous women only. Adjusted for age and study phase.

Table 5.5. Case-Case Odds Ratios (OR) and 95% Confidence Intervals (95% CI) for the associations between risk factors comparing lobular and mixed to ductal histologic subtype breast tumors among Luminal A clinical subtype tumors only, CBCS 1-3

	Luminal A: Case-Case[a]				
	Ductal	Lobular		Mixed	
Risk factor	N (%*)	N (%*)	OR (95% CI)	N (%*)	OR (95% CI)
Parity					
Nulliparous	226 (15.3)	38 (12.7)	Ref	59 (14.4)	Ref
1	278 (20.1)	41 (12.9)	0.87 (0.54-1.40)	53 (11.3)	0.69 (0.45-1.05)
2	433 (33.5)	108 (46.1)	1.44 (0.96-2.16)	142 (41.7)	1.17 (0.82-1.66)
≥3	445 (31.0)	83 (28.4)	1.04 (0.68-1.59)	117 (32.6)	0.98 (0.68-1.59)
p-value			0.76		0.89
Years since last live birth (parous only)					
0-≤10	204 (11.6)	26 (6.5)	0.73 (0.39-1.34)	54 (9.6)	1.26 (0.75-2.12)
>10-≤20	291 (17.2)	68 (20.8)	1.26 (0.82-1.93)	70 (15.4)	1.09 (0.72-1.65)
>20	657 (71.2)	138 (72.6)	Ref	187 (74.9)	Ref
Missing	4	0		1	
Age at first live birth (parous only; years)					
<26	782 (65.3)	146 (64.6)	Ref	198 (66.4)	Ref
≥26	364 (34.7)	85 (35.4)	1.31 (0.96-1.78)	114 (33.6)	1.17 (0.88-1.55)
Missing	10	1		0	
Lactation duration (parous only)					
Never	634 (54.8)	117 (45.2)	Ref	155 (50.0)	Ref
>0-≤12	371 (31.9)	73 (34.0)	1.08 (0.78-1.50)	100 (33.1)	1.06 (0.79-1.42)
>12	151 (13.2)	42 (20.8)	1.51 (1.02-2.25)	57 (16.9)	1.51 (1.06-2.18)
Missing	0	0		0	
p-value			0.15		0.21
Oral Contraceptive Use					
Never	402 (29.8)	65 (24.7)	Ref	88 (25.5)	Ref
Current	96 (5.3)	20 (4.6)	1.82 (0.99-3.36)	25 (5.0)	1.30 (0.76-2.25)
Former	880 (65.0)	182 (70.7)	1.48 (1.06-2.06)	256 (69.6)	1.25 (0.94-1.66)
Missing	4	3		2	
Hormone Therapy Use					
Never	1031 (64.8)	188 (60.8)	Ref	268 (66.7)	Ref
Estrogen alone	186 (18.3)	41 (20.4)	1.07 (0.72-1.60)	67 (26.7)	1.25 (0.88-1.76)
Combined E+P	137 (16.9)	35 (18.8)	1.23 (0.80-1.90)	28 (11.7)	0.77 (0.48-1.21)
Missing	28	6		8	

*All percentages weighted for study sampling design.

[a]Adjusted for race, age, study phase.

5.4 Discussion

In the Carolina Breast Cancer Study, we observed differences in reproductive risk factor profiles between ductal and lobular invasive breast cancers. Lobular disease was consistently, positively associated with more than 10 and up to 20 years since last birth, older age (≥26 years) at first birth, lactation duration greater than 12 months, oral contraceptive use, and combined estrogen plus progesterone (E+P) hormone therapy (HT) use. These associations did not vary by race and were not altered by restriction to Luminal A tumors only, suggesting that risk factor associations for histologic subtypes persist even after restricting to the most common molecular subtype.

In agreement with the previous literature, we found that relative to ductal cancers, lobular cancers were less frequent among black versus white women [21, 37, 87, 146], less frequent among younger women [17, 21, 23, 25, 26, 29, 37, 87], but more common among premenopausal women after controlling for age [19, 30]. As has been shown previously, we observed that lobular tumors were predominantly Luminal A subtype while ductal tumors displayed diversity in clinical subtype and contained a majority of the TN and HER2+ tumor types [3, 5, 6, 8–10, 12–15, 111].

Associations between a number of reproductive risk factors and histology among women from the CBCS were similar to associations reported elsewhere. We and others have shown lobular tumors are more strongly associated with older age at first birth [24, 27–29, 33–36, 147–149], oral contraceptive use [33, 83, 150], and combined E+P HT use [24, 28, 49, 50]. In case-case analyses, lobular tumors were significantly associated with older age at first birth, lactation duration greater than 12 months, and current OC use suggesting that these risk factors may contribute to etiologic differences between ductal and lobular tumors. Even after restricting to the predominant lobular clinical subtype, Luminal A, the aforementioned risk factor associations persisted, which mirrors work by Kotsopoulos *et al.* (2010) who reported risk factor associations among lobular tumors persisted after restricting to ER+ and PR+ tumors [28]. Concerning lactation duration and risk of lobular

disease, our findings differed from those reported previously where slightly inverse or null associations for lactation duration and lobular disease have been reported [27, 35, 147, 150]. We found that among parous women only, lactation duration greater than 12 months was significantly associated with lobular disease relative to ductal, even after restricting to Luminal A subtype. Previous studies have included more women over the age of 50 and had lower proportions of women who reported never breastfeeding (30-40%) than was observed in our study (50%) [27, 35, 147, 150]. Generational differences in breastfeeding practices and geographic variation of breastfeeding initiation [151] may contribute to the observed differences in the literature for lactation duration and lobular carcinoma. Overall, our study supports different risk factor profiles between ductal and lobular tumors, particularly for risk factors that are thought to impact hormone levels.

The findings for invasive lobular carcinoma displaying consistent associations with a host of hormone-modulating, reproductive risk factors even after restricting to Luminal A clinical subtype beg the question: is histologic subtype etiologic in origin or is it a result of selective pressures in the breast that encourage a tumor to develop into one histologic subtype over another? Intrinsic breast cancer subtypes are hypothesized to be etiologic in nature with Luminal A and B arising from the luminal epithelial cells of the mammary gland and Basal-like and HER2-enriched arising from the basal cells of the mammary gland [2]. Conversely, histologic subtype is subjective in nature and determined by a pathologist from the visual appearance of the epithelial cells where ductal histology is characterized by tubules and solid nests of epithelial cells and lobular carcinoma is characterized by a non-cohesive phenotype with single-file strands of epithelial cells scattered throughout the stroma [17, 51]. Lobular carcinomas are characterized by down regulation of E-cadherin in >90% of cases; however, this can also be observed in a smaller percentage of ductal carcinomas [25]. Therefore, it is not clear that histologic subtype is truly etiologic in nature or if histology is a result of exposure to phenotype-modulating selective pressures present in breast tissue over the life course.

Our study should be interpreted in light of some limitations. We were unable to include rare histologic subtypes (e.g. medullary, papillary, metaplastic, etc.) in our analyses due to limited sample sizes. Because lobular tumors are predominantly Luminal A subtype, we were unable to study differences in risk factor profiles between lobular tumors that were of Triple Negative or HER2+ subtype. We, like other studies of histology, acknowledge some uncertainty around histologic classification. Interobserver reliability for histologic subtype is reported to be around 80% [152], possibly leading to some instability in associations across studies. We sought to eliminate this problem by using centralized pathology review to classify CBCS invasive breast tumors into histologic subtypes, by focusing on tumors that had a dominant ductal or lobular phenotype (at least 80% ductal or lobular), and by considering mixed ductal-lobular tumors separately. This may have impacted our power, though, and TCGA found that molecularly, mixed tumors were not a distinct disease and displayed genomic features that would classify them as ductal- or lobular-like [5]. Therefore, applying these molecular classifications to mixed ductal-lobular tumors in epidemiologic studies may be a step toward better characterizing risk factor profiles for mixed tumors.

5.5 Conclusions

To conclude, we observed differences in risk factor profiles between ductal and lobular tumors in the Carolina Breast Cancer Study that persisted after restricting to Luminal A subtype. Using both case-control and case-case analyses, we found that lobular tumors have unique risk factor profiles from ductal tumors when considering older age at first birth, increasing lactation duration, current oral contraceptive use, and combined E+P HT use. When we restricted to Luminal A subtype, we found that the observed reproductive risk factor associations by histologic subtype were not altered. Overall, our findings suggest potential etiologic or phenotype-modulating differences between ductal and lobular disease that are not driven by intrinsic subtype alone. Our findings strengthen the evidence that lobular tumors are sensitive to hormone-modulating exposures.

CHAPTER 6: BREAST CNACER-SPECIFIC SURVIVAL BY HISTOLOGIC SUBTYPE

6.1 Introduction

While consistent patterns of survival by intrinsic breast cancer subtype have shown that Luminal A tumors generally have better survival than other subtypes, patterns of survival by histologic subtype are not as clear [3, 109, 110]. Overall survival at 5 and 10 years tends to be fairly similar by histologic subtypes [15, 17, 20, 26, 37–40, 95]. However, the findings for risk of breast cancer-specific death are complex and hard to compare across studies as methods for stratification by histologic and molecular subtype vary greatly with no studies reporting survival estimates for histologic subtype stratified by intrinsic subtype [12, 18, 22, 41, 108]. In a large study pooling data from 15 clinical trials, Pestalozzi *et al.* (2008) found that lobular tumors had an early survival advantage over ductal tumors, but the trend was reversed after 10 years and similar survival patterns were observed when stratifying by ER status [23]. As lobular tumors are predominantly Luminal A intrinsic subtype [3, 5] and survival by Luminal A subtype has been shown to differ from that of other intrinsic subtypes, we sought to estimate the association between histologic subtype and breast cancer-specific death, overall and among Luminal A tumors only, using the Carolina Breast Cancer Study Phases 1-2 (1993-2001).

6.2 Methods

6.2.1 Study population

The present analysis includes 1,459 cases of invasive breast cancer from the Carolina Breast Cancer Study (CBCS) Phases 1-2 (1993-2001). The CBCS is a population-based study among black and white women, that recruited participants from 24 of the 100 North Carolina counties [137]. CBCS oversampled women <50 years of age and black women

[82]. For CBCS 1-2, eligibility criteria included: being female, having a first diagnosis of breast cancer [(invasive or *in situ* (CBCS2 only)], aged 20-74 years at diagnosis, and residence in specified counties. Cases were enrolled via rapid case ascertainment from the NC Central Cancer registry. Cases provided informed consent for study enrollment and cases granted access to tumor tissue blocks/slides and medical records from treatment centers. To be eligible for this analysis, participants must have had invasive tumor tissue available for centralized pathology review and histologic subtype classified as ductal (n=1,209), lobular (n=144), or mixed ductal/lobular (henceforth referred to as mixed) breast cancer (n=106). The study maintains approval by the University of North Carolina School of Medicine Institutional Review Board.

6.2.2 Histologic subtype

Histologic subtype for CBCS 1-2 was determined through centralized pathology review. To be classified as ductal or lobular histologic subtype tumors were ≥80% representative of that histology; whereas, mixed tumors contained <80% of one histologic subtype and ≥20% of the second histologic subtype.

6.2.3 Clinical breast cancer subtype

Estrogen receptor (ER) and progesterone receptor (PR) status was obtained from the medical records for 80% of CBCS 1-2 cases as previously described [3, 109]. The remaining cases with available tumor tissue were stained at UNC for ER and PR and all cases with tumor tissue available were stained for HER2 [3, 109]. Clinical 3-marker subtypes were defined as follows: Luminal A (ER+ or PR+ and HER2-), Luminal B (ER+ or PR+ and HER2+), Triple Negative (TN) (ER- and PR- and HER2-), and HER2+ (ER- and PR- and HER2+).

6.2.4 Survival data

CBCS survival data has been previously described [62, 153, 154]. To identify breast cancer specific deaths International Classification of Death codes, ICD-9: 174.9 and ICD-10: 50.9, were used to determine the primary cause of death in the National Death Index (NDI)

database where participants were matched based on social security number and date of birth. All cases with a primary cause of death not identified with the specified breast cancer-specific codes were classified as "other" causes of death and were censored in the breast cancer-specific survival analyses. The linkage with NDI for the CBCS 1-2 occurred on 12/31/2011 and individuals who were recorded as living were censored on that date. The median follow-up time was 13.5 years with a range of 0.2-18.7 years.

6.2.5 Statistical analysis

Associations between histology (ductal, lobular, and mixed) and race [self-report: black, non-black (>98% white, henceforth referred to as white)], age (years) (<40, 40-49, 50-59, ≥60), menopausal status (pre-, post-), tumor size (≤2cm, >2cm), combined mitotic and clinical tumor grade (CBCS 1 only) (low-intermediate, high), lymph node status (positive, negative), AJCC stage of disease (I/II, III/IV), ER, PR, HER2 status (negative/positive), and 3-marker clinical subtype (Luminal A, Luminal B, Triple Negative, HER2+) were assessed using weighted sample percentages and unweighted sample counts.

Breast cancer-specific survival proportions by histologic subtype were assessed at 5-, 10-, and 15-year time points. Kaplan-Meier survival curves were constructed and non-parametric Log-Rank tests were conducted. Log-log plots were examined to assess deviation from the proportional hazards assumption, which appeared to occur around the 10-year time point. Therefore, time-stratified Cox Proportional Hazards models were used to estimate hazards ratios (HRs) and 95% confidence intervals (95% CI) stratified on ≤10 years and >10 years. To further assess violations of the proportional hazards assumption, 1st-, 2nd-degree polynomial and log(time) interaction terms between time and histologic subtype were examined. Based on Likelihood Ratio Tests (LRT) and Akaike information criterion (AIC) an interaction term between time modeled as a quadratic term and histologic subtype was included in all Cox regression models. All analyses were done in SAS version 9.4 (SAS Institute, Cary, NC). Graphs were constructed using GraphPad Prism version 7.02

for Windows (GraphPad Software, La Jolla, CA). P-values were produced for a two-sided test with an alpha of 0.05 for statistical significance.

6.3 Results

6.3.1 Lobular tumors have unique tumor characteristics.

In the Carolina Breast Cancer Study Phase 1 and 2, lobular tumors were more common among older and white women, were more frequently >2cm in size and were often low-intermediate tumor grade. Compared to ductal, lobular tumors had higher proportions of node positive and stage III/IV disease. Lobular tumors were more frequently ER positive (77%), PR positive (72%), and HER2 negative (90%) than ductal (60% ER+, 58% PR+, 82% HER2-). Similarly, lobular tumors were predominantly Luminal A clinical subtype (79%). This information served to inform subsequent survival analyses.

6.3.2 Lobular tumors have a breast cancer-specific survival advantage over ductal tumors.

Breast cancer specific survival percentages by histologic subtype at 5, 10, and 15 years were: ductal (85%, 78%, 75%, respectively), lobular (92%, 86%, 78%, respectively), and mixed (88%, 78%, 72%, respectively). Kaplan-Meier survival curves suggested a survival advantage for lobular breast tumors over ductal and mixed tumors (Figure 6.1) (Log-Rank p=0.45). A similar, but attenuated trend was observed after restricting to Luminal A subtype (Log-Rank p=0.65) (Figure 6.1).

6.3.3 Risk of breast cancer-specific death may vary over time for lobular tumors.

Using Cox regression models, the effect of histologic subtype was time-dependent ($p<0.01$) and as such, time stratified models are presented in Table 6.2. Relative to ductal, lobular and mixed tumors were associated with a reduced risk of breast cancer-specific death from 0-10 years after diagnosis [Lobular HR: 0.70, 95% CI (0.38-1.29); Mixed HR: 0.73, 95% CI (0.42-1.28)]. Associations were similar in direction and magnitude after adjusting for clinical subtype. After restricting to Luminal A tumors only, lobular and mixed tumors had a larger inverse association with breast cancer-specific death relative to ductal

for 0-10 years after diagnosis. Lobular and mixed tumors showed an increased risk of breast cancer-specific death as compared to ductal at >10 years since diagnosis; however, sample sizes in these strata were reduced as reflected in the wide confidence intervals.

Table 6.1. Descriptive summary of ductal, lobular and mixed ductal-lobular histologic subtype tumors, CBCS 1-2

Risk factor	Ductal N (%*)	Lobular N (%*)	Mixed N (%*)
Age at diagnosis			
≥60	293 (36.5)	45 (42.2)	34 (48.8)
50-59	226 (25.7)	26 (26.8)	24 (24.9)
40-49	473 (26.3)	66 (28.7)	32 (17.5)
<40	217 (11.5)	7 (2.2)	16 (8.7)
Race			
White	665 (77.8)	93 (84.7)	74 (87.1)
Black	544 (22.2)	51 (15.3)	32 (12.9)
Menopausal Status			
Pre	610 (35.9)	66 (30.2)	47 (27.9)
Post	599 (64.1)	78 (69.8)	59 (72.1)
Tumor Size (cm)			
≤2	601 (58.4)	62 (49.8)	65 (71.5)
>2	552 (41.6)	73 (50.2)	37 (28.5)
Missing	56	9	4
Tumor Grade^			
Low-Intermed	263 (55.6)	61 (95.2)	57 (91.5)
High	270 (44.4)	3 (4.8)	11 (8.5)
Missing	676	80	38
Lymph Node Status			
Negative	745 (67.6)	80 (57.6)	54 (59.2)
Positive	438 (32.4)	61 (43.4)	48 (40.8)
Missing	26	3	
AJCC Stage			
I, II	1,017 (91.0)	105 (83.7)	89 (90.7)
III, IV	133 (9.0)	31 (16.3)	11 (9.3)
Missing	59	8	6
ER+	610 (60.1)	110 (77.2)	73 (77.9)
ER-	551 (39.9)	32 (22.8)	27 (22.1)
Missing	48	2	6
PR+	590 (57.8)	108 (72.2)	61 (65.6)
PR-	566 (42.2)	32 (27.8)	38 (34.4)
Missing	53	4	7
HER2-	874 (82.1)	118 (89.7)	80 (75.3)
HER2+	202 (17.9)	11 (10.3)	17 (24.7)
Missing	133	15	9
Clinical Subtype			
Luminal A	519 (55.7)	103 (78.6)	58 (60.4)
Non-Luminal A	535 (44.3)	23 (21.4)	34 (39.6)
Luminal B	114 (11.2)	9 (9.7)	14 (22.8)
HER2+	81 (6.4)	2 (0.9)	2 (2.4)
Triple Negative	340 (26.6)	12 (10.8)	18 (14.4)
Missing	155	18	14
Vital Status			
Alive	734 (61.6)	88 (55.7)	63 (61.8)
Deceased	475 (38.4)	56 (44.3)	43 (38.2)
Missing	0	0	0
Cause of Death			
Breast Cancer	289 (47.7)	30 (51.4)	29 (56.2)
Other	186 (52.3)	26 (48.6)	14 (43.8)
Missing	0	0	0

*All percentages weighted for study sampling design.
^CBCS 1 only

Figure 6.1 a. Breast cancer-specific survival by histologic subtype. b. Breast cancer-specific survival by histologic subtype among Luminal A clinical subtype tumors only, CBCS 1-2.

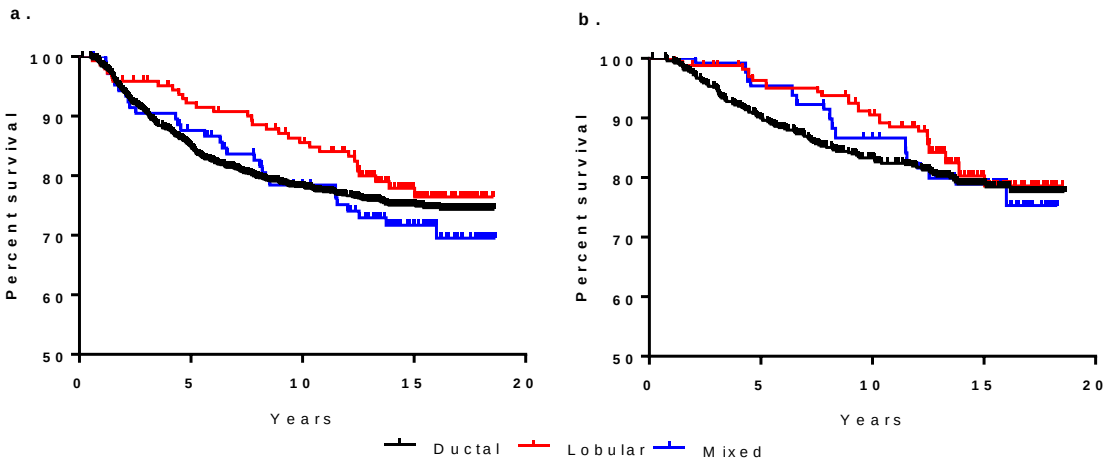

Table 6.2. Time stratified analyses at 0-10 years and >10 years for Hazard Ratios (HR) and 95% Confidence Intervals (95% CI) for associations between histologic subtype and breast cancer-specific death among all clinical subtypes, adjusted for clinical subtype, and among Luminal A tumors only, CBCS 1-2

	Overall		Adjusted for clinical subtype		Among Luminal A tumors only	
	Events/Censored [n (%*)]	HR (95% CI)[a]	Events/Censored [n (%*)]	HR (95% CI)[b]	Events/Censored [n (%*)]	HR (95% CI)[c]
			0-≤10 Years			
Ductal	256 (56.5) / 113 (43.5)	Ref.	228 (58.9) / 94 (41.1)	Ref.	83 (42.7) / 64 (57.3)	Ref.
Lobular	20 (66.0) / 10 (34.0)	0.70 (0.38-1.29)	14 (60.5) / 10 (39.5)	0.67 (0.31-1.35)	11 (61.1) / 8 (38.9)	0.50 (0.18-1.42)
Mixed	22 (56.3) / 10 (43.7)	0.73 (0.42-1.28)	20 (58.5) / 7 (41.5)	0.77 (0.42-1.41)	7 (48.4) / 3 (51.6)	0.45 (0.15-1.30)
			>10 Years			
Ductal	33 (3.5) / 807 (96.5)	Ref.	29 (3.5) / 703 (96.5)	Ref.	19 (4.2) / 353 (95.8)	Ref.
Lobular	10 (10.2) / 104 (89.8)	1.18 (0.12-11.91)	10 (11.3) / 92 (88.7)	1.39 (0.13-15.00)	9 (11.8) / 75 (88.2)	1.74 (0.11-27.61)
Mixed	7 (7.8) / 67 (92.3)	1.09 (0.08-14.63)	7 (8.5) / 59 (91.5)	1.33 (0.09-18.99)	4 (5.9) / 44 (94.1)	1.23 (0.05-28.9)

*All percentages weighted for study sampling design.
[a]Model adjusted for age, race.
[b]Model adjusted for age, race, and clinical subtype.
[a]Among Luminal A tumors only. Model adjusted for age, race.

6.4 Discussion

In the Carolina Breast Cancer Study phases 1-2, we observed that overall survival proportions at 5, 10, and 15 years were similar by histologic subtype. Visually, Kaplan-Meier curves suggested an early survival advantage for lobular tumors relative to mixed and ductal tumors, but mixed tumors displayed survival disadvantage relative to ductal at more than 10 years. Due to stratification at 10 years, sample sizes were reduced for Cox models, but the hazards ratios for lobular and mixed tumors versus ductal indicated an inverse association with breast cancer-specific death at 0-10 years, even after restricting to Luminal A tumors. Whereas, the hazards ratios for lobular carcinoma relative to ductal carcinoma were positively association with breast cancer-specific death at 10+ years, even among Luminal A tumors only.

Previous reports in the literature on breast cancer-specific survival by histologic subtype are inconsistent. Our finding of a reduced risk of death for lobular compared to ductal tumors was similar in magnitude to the association reported by Campbell *et al*. (2015) [108]. In analyses stratified by ER and PR status and age in SEER data, Li (2010) reported a slightly reduced risk of death for lobular compared to ductal tumors that did not remain when restricted to ER+/PR+ tumors [22]. Conversely, in all strata of ER/PR status in SEER data, Dunnwald *et al.* (2007) observed a reduced risk of breast cancer-specific death among lobular tumors, which mirrors our findings of a reduced risk of death among lobular tumors when restricted to Luminal A subtype. One study has reported an increased risk of breast cancer-specific death among lobular relative to ductal disease [12]. Our findings mirror those reported by Pestalozzi *et al.* (2008), who showed a reversal of the risk of overall death at >10 years, for lobular as compared to ductal tumors [23], which warrants further investigation in epidemiologic studies with large sample sizes and survival data of 15 or more years.

A limitation of this analysis is that the CBCS has extensive follow-up data on phases 1-2, but phase 3 is still accumulating survival data as this phase finished enrollment in

2013, which has limited our power to study these associations. Similarly, Nanostring analyses are not available for CBCS 1-2, which would allow for PAM50 intrinsic subtype to be accounted for in survival analyses. Finally, we were unable to resolve the mixed tumors into either ductal-like or lobular-like, but using molecular data may allow for these categorizations in the future, which would increase the sample sizes of lobular tumors for survival analyses.

6.5 Conclusions

Overall, while we had limited sample sizes in our time-stratified survival analyses, our results mirror those reported previously and lend strength to the hypothesis that lobular cancers have better survival in the early years following diagnosis, but a cross-over may occur resulting in lobular cancers having worse long-term survivorship. These findings could be due to later stage at diagnosis and differences in metastatic spread between ductal and lobular tumors [23, 25, 26, 95, 155], but mechanistic data is lacking. We found that the associations by histologic subtype persisted when restricting to Luminal A subtype, suggesting both histology and molecular subtype may play a role in survival.

CHAPTER 7: DISCUSSION

7.1 Main Findings

The main research aims of this dissertation were first, to estimate the association between molecular and clinical characteristics and histologic subtype, overall and among Luminal A tumors, and to compare these findings to those observed in The Cancer Genome Atlas (TCGA). Second, to estimate the association between breast cancer risk factors (race, age at diagnosis, menopausal status, recency of last birth, age at first birth, parity, lactation duration, and exogenous hormone use) and histologic subtype, overall and among Luminal A tumors. Third, to estimate the association between histologic subtype and breast cancer-specific survival, overall and among Luminal A tumors.

First, we found that the tumor and molecular characteristics of lobular and mixed tumors were quantitatively different from those for ductal tumors in CBCS and TCGA. In both studies, lobular tumors are significantly more likely to be Luminal A. While TCGA had higher proportions of more aggressive tumors, the magnitude of the relative frequency differences was similar between CBCS and TCGA for lobular as compared to ductal tumors. As previously reported, we observed unique associations between patient and tumor characteristics for lobular tumors, but we were able to take our analysis a step further by restricting to Luminal A subtype, where we found that the associations between tumor characteristics and histology persisted. Based on TCGA and others reporting lower frequencies if TP53 mutations in lobular tumors, we investigated the associations between histologic subtype and TP53 mutation as measured by IHC and TP53 mutant-like status by RNA, which represents pathway defects in TP53 independent of intrinsic breast cancer subtype and has not been evaluated in ductal, lobular, and mixed ductal-lobular tumors. We

observed that lobular tumors had lower frequencies of TP53 pathway defects than ductal or mixed tumors. In whole, the molecular findings concerning PAM50 subtype and TP53 status for CBCS, a population-based study, mirror those from TCGA for lobular tumors suggesting that lobular tumors may be molecularly distinct from ductal tumors, necessitating the identification of risk factors that may contribute to etiologic differences by histologic subtype.

Next, in the CBCS, we evaluated several reproductive risk factors in association with ductal, lobular, and mixed ductal-lobular breast tumors. Overall, we observed differences in reproductive risk factor profiles between ductal and lobular invasive breast cancers and intermediate risk factor profiles for mixed breast tumors. Lobular disease was positively associated with more than 10 and up to 20 years since last birth, older age (≥26 years) at first birth, lactation duration greater than 12 months, oral contraceptive use, and combined estrogen plus progesterone (E+P) hormone therapy (HT) use. These associations did not vary by race and were nearly unchanged in analyses restricted to Luminal A tumors. Our findings strengthen the evidence that lobular tumors are sensitive to hormone-modulating exposures and the case-case analysis suggested potential etiologic or phenotype-modulating differences between ductal and lobular disease that are not driven by Luminal A intrinsic subtype alone.

Finally, we characterized survival differences by histologic subtype of invasive breast cancer in CBCS 1-2. Based on previous reports, there is no clear pattern of survival by histologic subtype. Differences between studies are due to various methods of stratification by age groups, ER status, and survival time. In CBCS 1-2, we observed that 5-, 10-, and 15-year survival proportions were similar between ductal, lobular, and mixed histologic subtype tumors. Kaplan-Meier curves suggested an early survival advantage for lobular tumors relative to mixed and ductal tumors, but mixed tumors displayed a survival disadvantage relative to ductal at greater than 10 years after diagnosis. Therefore, we used time-stratified Cox regression models, which negatively impacted our power in each strata.

We observed an inverse association with breast cancer-specific death at 0-10 years for lobular and mixed tumors compared to ductal tumors, and the associations were similar in magnitude after restricting to Luminal A tumors. Conversely, we observed an increased risk of breast cancer-specific death for lobular compared to ductal tumors at 10 or more years, that persisted even among Luminal A tumors only. However, our findings should be interpreted with caution due to the low sample sizes available for survival analysis.

7.2 Is histologic subtype etiologic or does it arise from phenotype-modulating exposures acting on the breast tissue over the life course?

A recent in-depth molecular analysis by TCGA identified genomic differences between ductal and lobular tumors, even after restricting to Luminal A intrinsic subtype [5]. While TCGA and others [114, 156] have shown molecular differences between ductal and lobular breast tumors, loss of E-cadherin among invasive lobular tumors is the most consistent difference and is hypothesized to contribute to the non-cohesive phenotype observed under the microscope for lobular tumors. In a small study of lobular carcinoma *in situ* (LCIS) lesions, E-cadherin loss is suggested may be an early event in lobular tumorigenesis as 12 of the 13 LCIS samples studied harbored E-cadherin DNA mutations and were E-cadherin negative by immunohistochemistry similar to invasive lobular tumors [157]. However, it is not clear that all LCIS will proceed to invasive lobular disease. In an analysis of SEER data, Li *et al.* (2006) showed that among women diagnosed with invasive breast cancer following an initial diagnosis and treatment for LCIS and who did not undergo a mastectomy, 51% developed an invasive ductal carcinoma and 49% developed an invasive lobular carcinoma [158]. Further work remains to be done to examine the continuum of invasive lobular breast carcinogenesis and the underlying biologic mechanisms of development.

It could be argued that histologic subtypes of invasive breast cancer result from phenotype-modulating exposures that impact the tumor microenvironment and encourage the development of one histologic subtype over another. In risk factor analyses, we identified significant differences between ductal and lobular tumors for older age at first

birth, increased lactation duration, and current oral contraceptive use. These finding suggest that perhaps these risk factors are phenotype-modulating exposures that have encouraged the development of lobular, rather than ductal, carcinoma. It may follow biologically that over a time a ductal tumor could progress into a lobular tumor through loss of normal E-cadherin expression. One way to investigate this hypothesis in epidemiologic data could be to study the molecular profiles of mixed ductal-lobular tumors to see if they could be further parsed into ductal and lobular histologic subtypes. Mixed tumors are usually smaller in size like ductal tumors, which may suggest they are in a histologic transitional period at the time of diagnosis, shifting from ductal to lobular histology. Therefore, if we could molecularly classify mixed tumors as ductal-like or lobular-like, as performed in TCGA, then we may be better able to study the characteristics and risk factors of true ductal tumors and true lobular tumors. In turn, this may help to identify modifiable risk factors that reduce the burden of breast cancer, including reducing the risk of mixed tumors, which have been hard to identify since their risk factor profiles are generally intermediate to those observed for ductal and lobular tumors; however, this remains an important public health concern as up to 15% of breast cancer cases, approximately 34,000 of the estimated 230,000 diagnoses of invasive breast cancers each year, are classified as mixed ductal/lobular breast tumors.

7.3 Limitations

While this work had a number of methodologic strengths and was the first epidemiologic study that incorporated intrinsic breast cancer subtype into the study of histologic subtypes, there are some limitations to this dissertation. Even though we were able to compare our findings in CBCS for PAM50 intrinsic subtype and a TP53 gene signature and how these molecular characteristics varied by histologic subtype in TCGA, we were unable to extend beyond these molecular characteristics and explore other TCGA-identified genomic associations, namely E-cadherin mutation. CBCS currently has Nanostring data on 194 genes and the addition of 200 more genes, including E-cadherin, to the Nanostring panel is underway for participants from CBCS 1-3. Therefore, it may be

possible to look at a number of genes and how their expression varies by histologic subtype in the CBCS in the future.

Another limitation faced in comparing our findings in the CBCS to those from other studies is the lack of uniformity in histologic classification between studies and by pathologists, which impacts histologic breast cancer subtype research as a whole. This continues to pose challenges for studying histologic subtypes as illustrated here in the differences in proportion and magnitude of association for Luminal A subtype in mixed as compared to ductal tumors in CBCS and TCGA. CBCS used centralized pathology review and TCGA used an expert pathologist committee plus the pathology report to reach a histologic subtype consensus. While the TCGA classification system may appear to be a purer classification approach, it is not representative of the manner in which histology is determined in the clinic at the time of diagnosis or how it is determined in other epidemiologic studies. Therefore, the development of a universal classification scheme for mixed tumors will serve to clarify findings for this histologic subtype in molecular and epidemiologic research and aid in the identification of risk factors for mixed disease.

Due to low proportions of lobular and mixed tumors that were Luminal B, Triple Negative, and HER2+, we were unable to look at differences in risk factors for these histo-molecular subtypes. While there are distinct risk factor differences by intrinsic subtype, estimating these associations among each histologic subtype is important for identifying etiologic differences by histo-molecular subtypes and in developing risk reduction strategies for all histo-molecular subtypes of invasive breast cancer.

7.4 Significance

Our work attempts to answer an important biologic question in the study of histologic breast cancer heterogeneity-do the observed associations for histology persist when restricted to Luminal A subtype? Based on the collective findings of this work, it could be argued that histology defines an etiologic subtype as the observed associations for tumor characteristics, reproductive breast cancer risk factors, and survival all persisted when the

respective analyses were restricted to Luminal A subtype tumors only. In the first aim, we found that even after restricting to Luminal A subtype, lobular tumors were more likely to be larger in size, lower grade, and diagnosed at later stages of disease than ductal. In the second aim using case-case analyses, after restricting to Luminal A subtype, lobular tumors were associated with older age at first birth, lactation duration greater than 12 months, and oral contraceptive use. Finally, although power was lacking in the third aim, we estimated the association between histologic subtype and breast cancer-specific survival. We observed that after restricting to Luminal A subtype, lobular tumors had a reduced risk of death from 0-10 years and an increased risk of death at greater than 10 years. As a whole, these finding suggest that although histologic subtype of breast cancer is determined subjectively by a pathologist, there may be a number of molecular and biologic factors that contribute to the observed histologic phenotype and the identification of risk reduction strategies remains important.

By comparing and contrasting the molecular characteristics for histologic subtypes in CBCS and TCGA, we found that even though tumors included in TCGA had more aggressive characteristics, the differences between molecular, patient and tumor characteristics for ductal and lobular histologic subtypes were relatively robust. These comparisons are important for two reasons. First, TCGA is the 'gold-standard' for the molecular characterization of tumors and TCGA seeks to identify etiologic events in tumorigenesis that may inform the development of therapeutic targets; however, the more aggressive tumor types represented in TCGA may not be representative of the true population distributions of tumor characteristics. Second, understanding how the patients in TCGA compare to participants in population-based epidemiologic studies can help to better inform the application of the results from the TCGA. Our use of the same molecular signatures in CBCS also illustrates that the methods used in TCGA can be expanded into population-based studies with molecular tumor data for validation.

Finally, this dissertation was also able to provide population-based estimates of the association between histologic subtype and race. There are few epidemiologic studies that have a large enough proportion of black women to investigate these associations. We were able to confirm that black women in the CBCS are indeed less likely than white women to be diagnosed with lobular or mixed ductal-lobular breast cancer compared to ductal breast cancer. Given differences in reproductive patterns between black and white women, we then performed race-stratified analyses and found similar patterns of association between reproductive risk factors among black and white women for lobular relative to ductal disease suggesting that etiologic differences by histologic subtype do not vary by race.

7.5 Future Directions

There are a number of future directions that may be informative to the epidemiologic study of the histologic subtypes of invasive breast cancer. First, there are potential approaches that could be used to create better resolution of the mixed tumor category. As TCGA found mixed tumors to be lobular-like or ductal-like, similar classifications could be applied in epidemiologic studies using the proportion of each tumor that is lobular or ductal in conjunction with E-cadherin status. By resolving mixed tumors into ductal-like or lobular like, these efforts could increase sample sizes among lobular tumors where reduced sample size precludes studies from looking at lobular tumors with respect to risk and survival.

Identifying molecular signatures to distinguish between lobular and ductal tumors, independent of intrinsic subtype would greatly advance this field of study. As we move toward the inclusion of higher level genomic data in epidemiologic studies these research goals are more attainable. The molecular resolution of histologic subtype through histology-specific gene expression signatures would work toward eliminating the interobserver variation in histologic classification of breast tumors by pathologists. In turn, this may help to identify more appropriate treatment options and help to identify modifiable risk reduction strategies.

Creating consortia of epidemiologic studies of breast cancer could increase the sample sizes to include the rarer histologic subtypes, such as metaplastic, papillary, and tubular in risk factor and survival analyses, which we were unable to assess here. These subtypes, while rare, still require the identification of risk reduction strategies to reduce the breast cancer burden as a whole. Future work in epidemiologic studies with long term follow-up are needed to estimate these associations with appropriate power.

7.6 Conclusions

In conclusion, we estimated differences in association between molecular characteristics, risk factors, and breast cancer-specific survival and ductal, lobular, and mixed ductal-lobular histologic subtypes of invasive breast cancer. We observed that histology and intrinsic breast cancer subtype likely both contribute to the observed associations with tumor characteristics. We also observed that histology is associated with reproductive risk factors and breast cancer-specific survival, even after restricting analyses to Luminal A subtype only. Additionally, this dissertation presented a comparison of molecular findings for histologic subtype between the CBCS and TCGA and found the associations between tumor characteristics and histologic subtypes to be similar. This underscores the ability and importance of validating molecular findings in epidemiologic studies and should serve to encourage the use of molecular tumor data in epidemiologic research to better understand tumor biology and how it relates to risk and survival.

REFERENCES

1. (2012) Comprehensive molecular portraits of human breast tumors The Cancer Genome Atlas Network. Nature 490:61–70. doi: 10.1038/nature11412.Comprehensive

2. Perou CM, Sørlie T, Eisen MB, et al (2000) Molecular portraits of human breast tumours. Nature 406:747–52. doi: 10.1038/35021093

3. Carey LA, Perou CM, Livasy CA, et al (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 295:2492–502. doi: 10.1001/jama.295.21.2492

4. Millikan RC, Newman B, Tse C-K, et al (2008) Epidemiology of basal-like breast cancer. Breast Cancer Res Treat 109:123–39. doi: 10.1007/s10549-007-9632-6

5. Ciriello G, Gatza ML, Beck AH, et al (2015) Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell 163:506–519. doi: 10.1016/j.cell.2015.09.033

6. Sisti JS, Collins LC, Beck AH, et al (2016) Reproductive risk factors in relation to molecular subtypes of breast cancer: Results from the nurses' health studies. Int J Cancer 138:2346–2356. doi: 10.1002/ijc.29968

7. Iorfida M, Maiorano E, Orvieto E, et al (2012) Invasive lobular breast cancer: Subtypes and outcome. Breast Cancer Res Treat 133:713–723. doi: 10.1007/s10549-012-2002-z

8. Azim HA, Malek RA, Azim HA (2014) Pathological features and prognosis of lobular carcinoma in Egyptian breast cancer patients. Womens Health (Lond Engl) 10:511–8. doi: 10.2217/whe.14.48

9. Braunstein LZ, Brock JE, Chen Y-H, et al (2015) Invasive lobular carcinoma of the breast: local recurrence after breast-conserving therapy by subtype approximation and surgical margin. Breast Cancer Res Treat 149:555–64. doi: 10.1007/s10549-015-3273-y

10. Caldarella A, Buzzoni C, Crocetti E, et al (2013) Invasive breast cancer: A significant correlation between histological types and molecular subgroups. J Cancer Res Clin Oncol 139:617–623. doi: 10.1007/s00432-012-1365-1

11. Cha YJ, Jung WH, Cho NH, Koo JS (2015) Expression of sarcosine metabolism-related proteins in invasive lobular carcinoma: comparison to invasive ductal carcinoma. Yonsei Med J 56:598–607. doi: 10.3349/ymj.2015.56.3.598

12. García-Fernández A, Lain JM, Chabrera C, et al (2015) Comparative Long-term Study of a Large Series of Patients with Invasive Ductal Carcinoma and Invasive Lobular Carcinoma. Loco-Regional Recurrence, Metastasis, and Survival. Breast J 21:533–537. doi: 10.1111/tbj.12455

13. Lim ST, Yu JH, Park HK, et al (2014) A comparison of the clinical outcomes of patients with invasive lobular carcinoma and invasive ductal carcinoma of the breast according to molecular subtype in a Korean population. World J Surg Oncol 12:56. doi: 10.1186/1477-7819-12-56

14. Engstrøm MJ, Opdahl S, Vatten LJ, et al (2015) Invasive lobular breast cancer: the prognostic impact of histopathological grade, E-cadherin and molecular subtypes. Histopathology 66:409–419. doi: 10.1111/his.12572

15. Jung S-Y, Jeong J, Shin S-H, et al (2010) The invasive lobular carcinoma as a prototype luminal A breast cancer: a retrospective cohort study. BMC Cancer 10:664. doi: 10.1186/1471-2407-10-664

16. Lips EH, Mukhtar R a., Yau C, et al (2012) Lobular histology and response to neoadjuvant chemotherapy in invasive breast cancer. Breast Cancer Res Treat 136:35–43. doi: 10.1007/s10549-012-2233-z

17. Arpino G, Bardou VJ, Clark GM, Elledge RM (2004) Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. Breast Cancer Res 6:R149-56. doi: 10.1186/bcr767

18. Arps DP, Healy P, Zhao L, et al (2013) Invasive ductal carcinoma with lobular features: A comparison study to invasive ductal and invasive lobular carcinomas of the breast. Breast Cancer Res Treat 138:719–726. doi: 10.1007/s10549-013-2493-2

19. (2012) Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. Lancet Oncol 13:1141–1151. doi: 10.1016/S1470-2045(12)70425-4

20. Cristofanilli M, Gonzalez-Angulo A, Sneige N, et al (2005) Invasive Lobular Carcinoma Classic Type: Response to Primary Chemotherapy and Survival Outcomes. J Clin Oncol 23:41–48. doi: 10.1200/JCO.2005.03.111

21. Li CI, Uribe DJ, Daling JR (2005) Clinical characteristics of different histologic types of breast cancer. Br J Cancer 93:1046–1052. doi: 10.1038/sj.bjc.6602787

22. Li CI (2010) Risk of mortality by histologic type of breast cancer in the United States. Horm Cancer 1:156–65. doi: 10.1007/s12672-010-0016-8

23. Pestalozzi BC, Zahrieh D, Mallon E, et al (2008) Distinct clinical and prognostic features of infiltrating lobular carcinoma of the breast: Combined results of 15 International Breast Cancer Study Group clinical trials. J Clin Oncol 26:3006–3014. doi: 10.1200/JCO.2007.14.9336

24. Phipps AI, Li CI, Kerlikowske K, et al (2010) Risk factors for ductal, lobular, and mixed ductal-lobular breast cancer in a screening population. Cancer Epidemiol Biomarkers Prev 19:1643–54. doi: 10.1158/1055-9965.EPI-10-0188

25. Rakha EA, Gill MS, El-Sayed ME, et al (2009) The biological and clinical characteristics of breast carcinoma with mixed ductal and lobular morphology. Breast Cancer Res Treat 114:243–250. doi: 10.1007/s10549-008-0007-4

26. Zengel B, Yararbas U, Duran A, et al (2015) Comparison of the clinicopathological features of invasive ductal, invasive lobular, and mixed (invasive ductal + invasive lobular) carcinoma of the breast. Breast Cancer 22:374–381. doi: 10.1007/s12282-013-0489-8

27. Li CI (2006) Relationship between Established Breast Cancer Risk Factors and Risk of Seven Different Histologic Types of Invasive Breast Cancer. Cancer Epidemiol Biomarkers Prev 15:946–954. doi: 10.1158/1055-9965.EPI-05-0881

28. Kotsopoulos J, Chen WY, Gates M a, et al (2010) Risk factors for ductal and lobular breast cancer: results from the nurses' health study. Breast Cancer Res 12:R106. doi: 10.1186/bcr2790

29. Reeves GK, Pirie K, Green J, et al (2009) Reproductive factors and specific histological types of breast cancer: prospective study and meta-analysis. Br J Cancer 100:538–544. doi: 10.1038/sj.bjc.6604853

30. Reeves GK, Beral V, Green J, et al (2006) Hormonal therapy for menopause and breast-cancer risk by histological type: a cohort study and meta-analysis. Lancet Oncol 7:910–918. doi: 10.1016/S1470-2045(06)70911-1

31. Kelsey JL, Gammon MD, John EM (1993) Reproductive factors and breast cancer. EpidemiolRev 15:36–47.

32. Falk RT, Maas P, Schairer C, et al (2014) Alcohol and risk of breast cancer in postmenopausal women: an analysis of etiological heterogeneity by multiple tumor characteristics. Am J Epidemiol 180:705–17. doi: 10.1093/aje/kwu189

33. Nyante SJ, Gammon MD, Malone KE, et al (2008) The association between oral contraceptive use and lobular and ductal breast cancer in young women. Int J Cancer 122:936–941. doi: 10.1002/ijc.23163

34. Xu W-H, Zhang X-L, Gao Y-T, et al (2007) Joint effect of cigarette smoking and alcohol consumption on mortality. Prev Med An Int J Devoted to Pract Theory 45:313–319. doi: 10.1016/j.ypmed.2007.05.015

35. Beaber EF, Holt VL, Ph D, et al (2009) Reproductive factors, age at maximum height, and risk of three histologic types of breast cancer. 17:3427–3434. doi: 10.1158/1055-9965.EPI-08-0641.Reproductive

36. Newcomb P, Trentham-Dietz A, Hampton J, et al (2013) Late age at first full term birth is strongly associated with lobular breast cancer. Cancer 18:1199–1216. doi: 10.1016/j.micinf.2011.07.011.Innate

37. Bharat A, Gao F, Margenthaler JA (2009) Tumor characteristics and patient outcomes are similar between invasive lobular and mixed invasive ductal/lobular breast cancers but differ from pure invasive ductal breast cancers. Am J Surg 198:516–9. doi: 10.1016/j.amjsurg.2009.06.005

38. Korhonen T, Huhtala H, Holli K (2004) A comparison of the biological and clinical features of invasive lobular and ductal carcinomas of the breast. Breast Cancer Res Treat 85:23–29. doi: 10.1023/B:BREA.0000021038.97593.8b

39. Warren A, Newcomb P, Hampton, JM Titus-Ernstoff L, et al (2013) Reproductive factors and histologic subtype in relation to mortality after a breast cancer diagnosis. Breast Cancer Res Treat 18:1199–1216. doi: 10.1016/j.micinf.2011.07.011.Innate

40. Ellis IO, Galea M, O'Rouke S, et al (1992) Pathological prognostic factors in breast cancer. III. Vascular invasion: relationship with recurrence and survival in a large study with long-term follow-up. Histopathology 24:41–7. doi: 10.1111/j.1365-2559.1994.tb01269.x

41. Dunnwald LK, Rossing MA, Li CI (2007) Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. Breast Cancer Res 9:R6. doi: 10.1186/bcr1639

42. Hunt KK, Robertson JR BK (2014) The Breast. In: Brunicardi F, Andersen DK, Billiar TR, Dunn DL, Hunter JG, Matthews JB, Pollock RE. eds. Schwartz's Principles of Surgery, 10e. http://accessmedicine.mhmedical.com/content.aspx?bookid=980&Sectionid=59610859. Accessed 19 Jan 2016

43. Milanese TR, Hartmann LC, Sellers TA, et al (2006) Age-Related Lobular Involution and Risk of Breast Cancer. JNCI J Natl Cancer Inst 98:1600–1607. doi: 10.1093/jnci/djj439

44. Radisky DC, Hartmann LC (2009) Mammary involution and breast cancer risk: transgenic models and clinical studies. J Mammary Gland Biol Neoplasia 14:181–91. doi: 10.1007/s10911-009-9123-y

45. Hutson SW, Cowen PN, Bird CC (1985) Morphometric studies of age related changes in normal human breast and their significance for evolution of mammary cancer. d:281–287.

46. Wellings SR, Jensen HM, Marcum RG (1975) An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. J Natl Cancer Inst 55:231–273.

47. Ghosh K, Vachon CM, Pankratz VS, et al (2010) Independent association of lobular involution and mammographic breast density with breast cancer risk. J Natl Cancer Inst 102:1716–23. doi: 10.1093/jnci/djq414

48. Ferretti G, Felici A, Cognetti F (2007) Re: Age-related lobular involution and risk of breast cancer. J Natl Cancer Inst 99:571–2; author reply 572. doi: 10.1093/jnci/djk116

49. Li CI, Daling JR, Haugen KL, et al (2013) Use of menopausal hormone therapy and risk of ductal and lobular breast cancer among women 55–74 years of age. 18:1199–1216. doi: 10.1016/j.micinf.2011.07.011.Innate

50. Calle EE, Feigelson HS, Hildebrand JS, et al (2009) Postmenopausal hormone use and breast cancer associations differ by hormone regimen and histologic subtype. Cancer 115:936–45. doi: 10.1002/cncr.24101

51. Rosen PP (2009) Rosen's Breast Pathology, Third. Lippincott Williams & Wilkins, Phildelphia, PA

52. Makki J (2015) Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. Clin Med Insights Pathol 8:23–31. doi: 10.4137/CPath.S31563.TYPE

53. Corben AD (2013) Pathology of Invasive Breast Disease. Surg Clin North Am 93:363–392. doi: 10.1016/j.suc.2013.01.003

54. Tavassoli F a, Schnitt SJ, Hoefler H, et al (2003) World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Breast and Female Genital Organs. 63–74.

55. O'Brien KM, Cole SR, Millikan RC, et al (2013) RACE AND SUBTYPE DIFFERENCES IN THE REPLICATION OF PREVIOUSLY IDENTIFIED BREAST CANCER SUSCEPTIBILITY LOCI : A BAYESIAN APPROACH Katie M . O ' Brien A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulf. Univ. North Carolina Diss.

56. Sørlie T, Perou CM, Tibshirani R, et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869–74. doi: 10.1073/pnas.191367098

57. Hammond MEH, Hayes DF, Dowsett M, et al (2010) American society of clinical oncology/college of American pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). Arch Pathol Lab Med. doi: 10.1043/1543-2165-134.7.e48

58. Rosenthal SI, Depowski PL, Sheehan CE, Ross JS (2002) Comparison of HER-2/neu oncogene amplification detected by fluorescence in situ hybridization in lobular and ductal breast cancer. Appl Immunohistochem Mol Morphol 10:40–46. doi: 10.1097/00022744-200203000-00007

59. Cuzick J, Dowsett M, Pineda S, et al (2011) Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer. J Clin Oncol 29:4273–4278. doi: 10.1200/JCO.2010.31.2835

60. Dowsett M, Nielsen TO, A'Hern R, et al (2011) Assessment of Ki67 in Breast Cancer: Recommendations from the international Ki67 in breast cancer working Group. J Natl Cancer Inst 103:1656–1664. doi: 10.1093/jnci/djr393

61. Goldhirsch A, Winer EP, Coates AS, et al (2013) Personalizing the treatment of women with early breast cancer: Highlights of the st gallen international expert consensus on the primary therapy of early breast Cancer 2013. Ann Oncol 24:2206–2223. doi: 10.1093/annonc/mdt303

62. O'Brien KM, Cole SR, Tse C-K, et al (2010) Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clin Cancer Res 16:6100–10. doi: 10.1158/1078-0432.CCR-10-1533

63. Kwan ML, Kushi LH, Weltzien E, et al (2009) Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors. Breast cancer Res 11:R31. doi: 10.1186/bcr2261

64. Li C, Littman A, White E (2007) Relationship between age maximum height is attained, age at menarche, and age at first full-term birth and breast cancer risk. Cancer Epidemiol Biomarkers Prev 16:2144–2149. doi: 10.1158/1055-9965.EPI-07-0242

65. Brinton LA, Schairer C, Hoover RN, Fraumeni JF (1988) Menstrual factors and risk of breast cancer. Cancer Invest 6:245–254. doi: 10.3109/07357908809080645

66. Hsieh CC, Trichopoulos D, Katsouyanni K, Yuasa S (1990) 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 46:796–800. doi: 10.1002/ijc.2910460508

67. Key TJ, Verkasalo PK, Banks E (2001) Epidemiology of breast cancer. Lancet Oncol 2:133–140. doi: 10.1016/S1470-2045(00)00254-0

68. Stalsberg H, Thomas D (1993) Age distribution of histologic types of breast carcinoma. Int J Cancer 54:1–7.

69. Nyante SJ, Dallal CM, Gierach GL, et al (2013) Risk factors for specific histopathological types of postmenopausal breast cancer in the NIH-AARP Diet and Health Study. Am J Epidemiol 178:359–71. doi: 10.1093/aje/kws471

70. Hendrix SL, Langer RD, Stefanick ML, et al (2003) Influence of Estrogen Plus Progestin on Breast Cancer and Mammography in Healthy Postmenopausal Women The Women ' s Health Initiative Randomized Trial. 289:3243–3253.

71. Million Women Study Collaborators (2003) Breast cancer and hormone replacement therapy in the Million Women Study. Lancet 362:419–427.

72. Di Meglio A, Freedman RA, Lin NU, et al (2016) Time trends in incidence rates and survival of newly diagnosed stage IV breast cancer by tumor histology: a population-based analysis. Breast Cancer Res Treat 157:587–596. doi: 10.1007/s10549-016-3845-5

73. Phipps AI, Ichikawa L, Bowles EJA, et al (2011) NIH Public Access. 67:60–66. doi: 10.1016/j.maturitas.2010.04.015.Defining

74. Dossus L, Benusiglio PR (2015) Lobular breast cancer: incidence and genetic and non-genetic risk factors. Breast Cancer Res 17:546. doi: 10.1186/s13058-015-0546-7

75. Suhrke P, Zahl P-H (2015) Breast cancer incidence and menopausal hormone therapy in Norway from 2004 to 2009: a register-based cohort study. Cancer Med 4:1303–1308. doi: 10.1002/cam4.474

76. Newcomer LM, Newcomb PA, Potter JD, et al (2003) Postmenopausal hormone therapy and risk of breast cancer by histologic type (United States). Cancer Causes Control 14:225–233.

77. Medina D (2005) Mammary developmental fate and breast cancer risk. Endocr Relat Cancer 12:483–495. doi: 10.1677/erc.1.00804

78. Victora CG, Bahl R, Barros AJD, et al (2016) Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. Lancet 387:475–490. doi: 10.1016/S0140-6736(15)01024-7

79. Collaborative Group on Hormonal Factors in Breast Cancer (2002) Breast cancer and breastfeeding: collaborative reanalysis of individual data form 47 epidemiological studies in 30 countries, including 50 302 women with berast cancer and 96 973 women without disease. Br J Cancer 87:1234–1245. doi: DOI 10.1038/sj.bjc.6600596

80. Hulka BS, Moorman PG (2001) Breast cancer: Hormones and other risk factors. Maturitas 38:103–116. doi: 10.1016/S0378-5122(00)00196-1

81. Terry MB, Gammon MD, Schoenberg JB, et al (2002) Oral Contraceptive Use and Cyclin D1 Overexpression in Breast Cancer among Young Women Oral Contraceptive Use and Cyclin D1 Overexpression in Breast Cancer among Young Women. 1100–1103.

82. Moorman PG, Millikan RC, Newman B (2001) Oral contraceptives and breast cancer among African-american women and white women. J Natl Med Assoc 93:329–34.

83. Newcomer LM, Newcomb P a, Trentham-Dietz A, et al (2003) Oral contraceptive use and risk of breast cancer by histologic type. Int J Cancer 106:961–4. doi: 10.1002/ijc.11307

84. Li CI, Anderson BO, Daling JR, Moe RE (2003) Trends in Incidence Rates of Invasive Lobular and Ductal Breast Carcinoma. JAMA 289:1421–1424.

85. Ademuyiwa FO, Gao F, Hao L, et al (2014) US breast cancer mortality trends in young women according to race. Cancer 1–8. doi: 10.1002/cncr.29178

86. Eheman CR, Shaw KM, Ryerson AB, et al (2009) The Changing Incidence of In situ and Invasive Ductal and Lobular Breast Carcinomas: United States, 1999-2004. Cancer Epidemiol Biomarkers Prev 18:1763–1769. doi: 10.1158/1055-9965.EPI-08-1082

87. Chlebowski RT, Chen Z, Anderson GL, et al (2005) Ethnicity and breast cancer: factors influencing differences in incidence and outcome. J Natl Cancer Inst 97:439–48. doi: 10.1093/jnci/dji064

88. Porter PL, Lund MJ, Lin MG, et al (2004) Racial differences in the expression of cell cycle-regulatory proteins in breast carcinoma. Cancer 100:2533–2542. doi: 10.1002/cncr.20279

89. Metzger Filho O, Giobbie-Hurder A, Mallon E, et al (2015) Relative Effectiveness of Letrozole Compared With Tamoxifen for Patients With Lobular Carcinoma in the BIG 1-98 Trial. J Clin Oncol 33:2772–2779. doi: 10.1200/JCO.2015.60.8133

90. Badowska-Kozakiewicz AM, Patera J, Sobol M, Przybylski J (2015) The role of oestrogen and progesterone receptors in breast cancer – immunohistochemical evaluation of oestrogen and progesterone receptor expression in invasive breast cancer in women. Współczesna Onkol 3:220–225. doi: 10.5114/wo.2015.51826

91. Biglia N, Maggiorotto F, Liberale V, et al (2013) Clinical-pathologic features, long term-outcome and surgical treatment in a large series of patients with invasive lobular carcinoma (ILC) and invasive ductal carcinoma (IDC). Eur J Surg Oncol 39:455–460. doi: 10.1016/j.ejso.2013.02.007

92. Delpech Y, Coutant C, Hsu L, et al (2013) Clinical benefit from neoadjuvant chemotherapy in oestrogen receptor-positive invasive ductal and lobular carcinomas. Br J Cancer 108:285–91. doi: 10.1038/bjc.2012.557

93. Kwast ABG, Groothuis-Oudshoorn KCGM, Grandjean I, et al (2012) Histological type is not an independent prognostic factor for the risk pattern of breast cancer recurrences. Breast Cancer Res Treat 135:271–80. doi: 10.1007/s10549-012-2160-z

94. Rakha E a., El-Sayed ME, Menon S, et al (2008) Histologic grading is an independent prognostic factor in invasive lobular carcinoma of the breast. Breast Cancer Res Treat 111:121–127. doi: 10.1007/s10549-007-9768-4

95. Suryadevara A, Paruchuri LP, Banisaeed N, et al (2010) The clinical behavior of mixed ductal/lobular carcinoma of the breast: a clinicopathologic analysis. World J Surg Oncol 8:51. doi: 10.1186/1477-7819-8-51

96. Conners AL, Jones KN, Hruska CB, et al (2015) Direct-Conversion Molecular Breast Imaging of Invasive Breast Cancer: Imaging Features, Extent of Invasive Disease, and Comparison Between Invasive Ductal and Lobular Histology. Am J Roentgenol 205:W374–W381. doi: 10.2214/AJR.14.13502

97. Vlug E, Ercan C, van der Wall E, et al (2014) Lobular breast cancer: pathology, biology, and options for clinical intervention. Arch Immunol Ther Exp (Warsz) 62:7–21. doi: 10.1007/s00005-013-0251-0

98. Weiss A, Chang DC (2016) A Comparison of Breast-Specific Gamma Imaging of Invasive Lobular Carcinomas and Ductal Carcinomas. 2015–2017. doi: 10.1001/jamasurg.2015.0962.Author

99. Langlands F, White J, Kearins O, et al (2015) Contralateral breast cancer: incidence according to ductal or lobular phenotype of the primary. Clin Radiol 71:159–163. doi: 10.1016/j.crad.2015.10.030

100. Kanumuri P, Hayse B, Killelea BK, et al (2015) Characteristics of Multifocal and Multicentric Breast Cancers. Ann Surg Oncol 22:2475–2482. doi: 10.1245/s10434-015-4430-6

101. Arps DP, Jorns JM, Zhao L, et al (2014) Re-excision rates of invasive ductal carcinoma with lobular features compared with invasive ductal carcinomas and invasive lobular carcinomas of the breast. Ann Surg Oncol 21:4152–8. doi: 10.1245/s10434-014-3871-7

102. Cao A-Y, Huang L, Wu J, et al (2012) Tumor characteristics and the clinical outcome of invasive lobular carcinoma compared to infiltrating ductal carcinoma in a Chinese population. World J Surg Oncol 10:152. doi: 10.1186/1477-7819-10-152

103. Ng CH, Pathy NB, Taib NA, et al (2012) The estrogen receptor negative-progesterone receptor positive breast carcinoma is a biological entity and not a technical artifact. Asian Pac J Cancer Prev 13:1111–3. doi: 10.7314/APJCP.2012.13.4.1111

104. Rosenberg LU, Einarsdottir K, Friman EI, et al (2006) Risk Factors for Hormone Receptor-Defined Breast Cancer in Postmenopausal Women. Cancer Epidemiol Biomarkers Prev 15:2482–2488. doi: 10.1158/1055-9965.EPI-06-0489

105. Ross JS, Wang K, Sheehan CE, et al (2013) Relapsed Classic E-Cadherin (CDH1)-Mutated Invasive Lobular Breast Cancer Shows a High Frequency of HER2 (ERBB2) Gene Mutations. Clin Cancer Res 19:2668–2676. doi: 10.1158/1078-0432.CCR-13-0295

106. Korhonen T, Kuukasjärvi T, Huhtala H, et al (2013) The impact of lobular and ductal breast cancer histology on the metastatic behavior and long term survival of breast cancer patients. Breast 22:1119–24. doi: 10.1016/j.breast.2013.06.001

107. Boyle DP, Mcart DG, Irwin G, et al (2014) The prognostic significance of the aberrant extremes of p53 immunophenotypes in breast cancer. Histopathology 340–352. doi: 10.1111/his.12398

108. Campbell I, Scott N, Seneviratne S, et al (2015) Breast Cancer Characteristics and Survival Differences between Maori, Pacific and other New Zealand Women Included in the Quality Audit Program of Breast Surgeons of Australia and New Zealand. Asian Pac J Cancer Prev 16:2465–2472.

109. O'Brien KM, Cole SR, Tse C-K, et al (2010) Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clin Cancer Res 16:6100–10. doi: 10.1158/1078-0432.CCR-10-1533

110. Parker JS, Mullins M, Cheang MCU, et al (2009) Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. J Clin Oncol 27:1160–1167. doi: 10.1200/JCO.2008.18.1370

111. Cha YJ, Kim YH, Cho NH, Koo JS (2014) Expression of autophagy related proteins in invasive lobular carcinoma: comparison to invasive ductal carcinoma. Int J Clin Exp Pathol 7:3389–98.

112. Brinck U, Jacobs S, Neuss M, et al (2004) Diffuse Growth Pattern Affects E-Cadherin Expression in Invasive Breast Cancer. 2242:2237–2242.

113. Cleton-Jansen A-M (2002) E-cadherin and loss of heterozygosity at chromosome 16 in breast carcinogenesis: different genetic pathways in ductal and lobular breast cancer? Breast Cancer Res 4:5–8. doi: 10.1186/bcr416

114. Korkola JEJ, DeVries S, Fridlyand J, et al (2003) Differentiation of lobular versus ductal breast carcinomas by expression microarray analysis. Cancer Res 63:7167–7175.

115. Grabowska MM (2012) Soluble E-cadherin: more than a symptom of disease. Front Biosci 17:1948. doi: 10.2741/4031

116. de Deus Moura R, Wludarski SCL, Carvalho FM, Bacchi CE (2013) Immunohistochemistry applied to the differential diagnosis between ductal and lobular carcinoma of the breast. Appl Immunohistochem Mol Morphol 21:1–12. doi: 10.1097/PAI.0b013e318255bafa

117. (2014) The Cancer Genome Atlas Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature 507:315–22. doi: 10.1038/nature12965

118. Fei F, Zhang D, Yang Z, et al (2015) The number of polyploid giant cancer cells and epithelial-mesenchymal transition-related proteins are associated with invasion and metastasis in human breast cancer. J Exp Clin Cancer Res 34:158. doi: 10.1186/s13046-015-0277-8

119. Lewis-Tuffin LJ, Rodriguez F, Giannini C, et al (2010) Misregulated E-cadherin expression associated with an aggressive brain tumor phenotype. PLoS One 5:e13665. doi: 10.1371/journal.pone.0013665

120. Reddy P, Liu L, Ren C, et al (2005) Formation of E-Cadherin-Mediated Cell-Cell Adhesion Activates Akt and Mitogen Activated Protein Kinase via Phosphatidylinositol 3 Kinase and Ligand-Independent Activation of Epidermal Growth Factor Receptor in Ovarian Cancer Cells. Mol Endocrinol 19:2564–2578. doi: 10.1210/me.2004-0342

121. Wen L, Sun L, Xi Y, et al (2015) Expression of calcium sensing receptor and E-cadherin correlated with survival of lung adenocarcinoma. Thorac Cancer 6:754–760. doi: 10.1111/1759-7714.12255

122. Hu Q-P, Kuang J-Y, Yang Q-K, et al (2015) Beyond a tumor suppressor: Soluble E-cadherin promotes the progression of cancer. Int J Cancer 0:n/a-n/a. doi: 10.1002/ijc.29982

123. Derksen PWB, Liu X, Saridin F, et al (2006) Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. Cancer Cell 10:437–449. doi: 10.1016/j.ccr.2006.09.013

124. Gilmore AP (2005) Anoikis. Cell Death Differ 12:1473–1477. doi: 10.1038/sj.cdd.4401723

125. Hanahan D, Weinberg RA, Francisco S (2000) The Hallmarks of Cancer Review University of California at San Francisco. 100:57–70.

126. Silwal-Pandit L, Vollan HKM, Chin SF, et al (2014) TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance. Clin Cancer Res 20:3569–3580. doi: 10.1158/1078-0432.CCR-13-2943

127. Fadare O, Wang S a, Hileeto D (2008) The expression of cytokeratin 5/6 in invasive lobular carcinoma of the breast: evidence of a basal-like subset? Hum Pathol 39:331–6. doi: 10.1016/j.humpath.2007.07.014

128. Harbhajanka A, Lamzabi I, Singh RI, et al (2014) Correlation of clinicopathologic parameters and immunohistochemical features of triple-negative invasive lobular carcinoma. Appl Immunohistochem Mol Morphol 22:e18-26. doi: 10.1097/PAI.0b013e3182a5a02d

129. van't Veer LJ, Dai H, Vijver MJ Van De, et al (2002) Gene expression pro ® ling predicts clinical outcome of breast cancer. doi: 10.1038/415530a

130. West M, Blanchette C, Dressman H, et al (2001) Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci U S A 98:11462–11467. doi: 10.1073/pnas.201162998

131. Huang B, Omoto Y, Iwase H, et al (2014) Differential expression of estrogen receptor α, β1, and β2 in lobular and ductal breast cancer. Proc Natl Acad Sci 111:1933–1938. doi: 10.1073/pnas.1323719111

132. Sikora MJ, Cooper KL, Bahreini A, et al (2013) Invasive lobular carcinoma cell lines are characterized by unique estrogen-mediated gene expression patterns and altered tamoxifen response. 18:1199–1216. doi: 10.1016/j.micinf.2011.07.011.Innate

133. Huang B, Warner M, Gustafsson J-Å (2014) Estrogen receptors in breast carcinogenesis and endocrine therapy. Mol Cell Endocrinol 418:240–244. doi: 10.1016/j.mce.2014.11.015

134. Scholzen T, Gerdes J (2000) The Ki-67 protein: From the known and the unknown. J Cell Physiol 182:311–322. doi: 10.1002/(SICI)1097-4652(200003)182:3<311::AID-JCP1>3.0.CO;2-9

135. Engels CC, Fontein DBY, Kuppen PJK, et al (2014) Immunological subtypes in breast cancer are prognostic for invasive ductal but not for invasive lobular breast carcinoma. Br J Cancer 2:1–7. doi: 10.1038/bjc.2014.338

136. WHI Writing Group (2002) Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women. Jama 288:321–333. doi: 10.1001/jama.288.3.321

137. Newman B, Moorman PG, Millikan R, et al (1995) The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. Breast Cancer Res Treat 35:51–60.

138. Hall IJ, Moorman PG, Millikan RC, Newman B (2005) Comparative analysis of breast cancer risk factors among African-American women and white women. Am J Epidemiol 161:40–51. doi: 10.1093/aje/kwh331

139. Allott EH, Cohen SM, Geradts J, et al (2016) Performance of three-biomarker immunohistochemistry for intrinsic breast cancer subtyping in the AMBER consortium. Cancer Epidemiol Biomarkers Prev 25:470–478. doi: 10.1158/1055-9965.EPI-15-0874

140. Taylor NJ, Nikolaishvili-feinberg N, Midkiff BR, et al (2015) Rational Manual and Automated Scoring Thresholds for the Immunohistochemical Detection of TP53 Missense Mutations in Human Breast Carcinomas. 0:1–7.

141. Conway K, Edmiston SN, Cui L, et al (2002) Prevalence and spectrum of p53 mutations associated with smoking in breast cancer. Cancer Res 62:1987–1995.

142. Furberg H, Millikan R. C., Geradts J., et al (2002) Environmental Factors in Relation to Breast Cancer Characterized by p53 Protein Expression. Cancer Epidemiol Biomarkers Prev 11:829–835.

143. Troester M a, Herschkowitz JI, Oh DS, et al (2006) Gene expression patterns associated with p53 status in breast cancer. BMC Cancer 6:276. doi: 10.1186/1471-2407-6-276

144. (2005) Easy SAS Calculations for Risk or Prevalence Ratios and Differences. Am J Epidemiol 162:199–200. doi: 10.1093/aje/kwi188

145. Begg CB, Zhang ZF (1994) Statistical analysis of molecular epidemiology studies employing case- series. Cancer Epidemiol Biomarkers Prev 3:173–175.

146. Li CI, Malone KE, Daling JR (2002) Differences in Breast Cancer Hormone Receptor Status and Histology by Race and Ethnicity among Women 50 Years of Age and Older. Cancer Epidemiol Biomarkers Prev 11:601–607.

147. Ursin G, Bernstein L, Lord SJ, et al (2005) Reproductive factors and subtypes of breast cancer defined by hormone receptor and histology. Br J Cancer 93:364–71. doi: 10.1038/sj.bjc.6602712

148. Albrektsen G, Heuch I, Thoresen S (2010) Histological type and grade of breast cancer tumors by parity, age at birth, and time since birth: a register-based study in Norway. BMC Cancer 10:226. doi: 10.1186/1471-2407-10-226

149. Lööf-Johanson M, Brudin L, Sundquist M, et al (2011) Breastfeeding and prognostic markers in breast cancer. Breast 20:170–5. doi: 10.1016/j.breast.2010.08.007

150. Li CI, Malone KE, Porter PL, et al (2003) Reproductive and anthropometric factors in relation to the risk of lobular and ductal breast carcinoma among women 65-79 years of age. Int J Cancer 107:647–651. doi: 10.1002/ijc.11465

151. Kogan MD, Singh GK, Dee DL, et al (2008) Multivariate analysis of state variation in breastfeeding rates in the United States. Am J Public Health 98:1872–1880. doi: 10.2105/AJPH.2007.127118

152. Longacre TA, Ennis M, Quenneville LA, et al (2006) Interobserver agreement and reproducibility in classification of invasive breast carcinoma: an NCI breast cancer family registry study. Mod Pathol 19:195–207. doi: 10.1038/modpathol.3800496

153. Sun X, Nichols HB, Tse CK, et al (2016) Association of parity and time since last birth with breast cancer prognosis by intrinsic subtype. Cancer Epidemiol Biomarkers Prev 25:60–67. doi: 10.1158/1055-9965.EPI-15-0864

154. Sun X, Nichols HB, Robinson W, et al (2015) Post-diagnosis adiposity and survival among breast cancer patients: influence of breast cancer subtype. Cancer Causes Control 26:1803–11. doi: 10.1007/s10552-015-0673-6

155. Sastre-Garau X, Jouve M, Asselain B, et al (1996) Infiltrating lobular carcinoma of the breast: Clinicopathologic analysis of 975 cases with reference to data on conservative therapy and metastatic patterns. Cancer 77:113–120. doi: 10.1002/(SICI)1097-0142(19960101)77:1<113::AID-CNCR19>3.0.CO;2-8

156. Zhao H, Langerød A, Ji Y, et al (2004) Different Gene Expression Patterns in Invasive Lobular and Ductal Carcinoma of the Breast. Mol Biol Cell 15:2523–2536. doi: 10.1091/mbc.E03

157. Mastracci TL, Tjan S, Bane AL, et al (2005) E-cadherin alterations in atypical lobular hyperplasia and lobular carcinoma in situ of the breast. Mod Pathol 18:741–751. doi: 10.1038/modpathol.3800362

158. Li CI, Malone KE, Saltzman BS, Daling JR (2006) Risk of invasive breast carcinoma among women diagnosed with ductal carcinoma in situ and lobular carcinoma in situ, 1988-2001. Cancer 106:2104–12. doi: 10.1002/cncr.21864