From DEPARTMENT OF MEDICAL EPIDEMIOLOGY AND BIOSTATISTICS
Karolinska Institutet, Stockholm, Sweden

MAMMOGRAPHIC DENSITY AND BREAST CANCER PHENOTYPES

Louise Eriksson

Karolinska Institutet

Stockholm 2012

Published by Karolinska Institutet. Printed by Larserics Digital Print AB

ISBN 978-91-7457-939-0

To my family

ABSTRACT

Mammographic density is one of the strongest risk factors for breast cancer and has been thoroughly studied as such. Extensive mammographic density also decreases screening sensitivity, thereby increasing the risk of interval cancers. Whether density acts as fertile ground for all types of breast cancer, or whether it influences tumor growth in a specific direction, was not known when we embarked upon the studies of this thesis. We therefore aimed to investigate the association between density, tumor characteristics, molecular subtypes, recurrence, and survival, focusing on interval cancers in the last study.

For studies I, III, and IV, we used the cases included in a population-based case-control study, in which cases were all Swedish women, aged 50-74, with incident breast cancer, diagnosed 1993-1995 (n=3345). We only included postmenopausal women with no prior history of cancer other than non-melanoma skin cancer and cervical cancer in situ (n=2720). Of these women, 1774 women had eligible mammograms.

For study II, in which we investigated the relationship between density and molecular subtypes, the study population was based on all women with breast cancer operated at a large university hospital in Stockholm 1994-1996 (n=524). Women with available gene expression profiling and mammograms were included in the study (n=110).

Pre-diagnostic/diagnostic density of the unaffected breast was assessed using a semi-automated, computer-assisted thresholding technique, Cumulus. Density was either measured as the dense area in cm^2 (absolute density=AD) or percentage density (PD) (the absolute dense area/the total breast area).

We did not find an association between density and tumor characteristics (lymph node metastasis, hormone-receptor status, grade, and histopathological classification) except for tumor size. However, this association seemed at least in part to be due to masking delaying diagnosis. In accordance with the lack of association between PD and most tumor characteristics, we did not find an association between density and molecular subtypes, nor between density, distant recurrence, and survival. We did, however, see a relatively strong association between PD and both local and locoregional recurrence, independent of established risk factors.

In the last study, we investigated the differences in survival between interval cancers and screening-detected cancers, taking mammographic density into account. We could show that interval cancers in both dense and non-dense breasts were associated with poorer prognosis compared to screening-detected cancers. However, the poorer prognosis seen in interval cancers in dense breasts seemed mainly attributable to delayed detection, whereas the group of interval cancers in non-dense breasts primarily seemed composed of truly aggressive tumors which we believe need further study.

LIST OF PUBLICATIONS

I. Eriksson L, Czene K, Rosenberg L, Humphreys K, Hall P: **The influence of mammographic density on breast tumor characteristics**. *Breast Cancer Res Treat* 2012, **134**(2):859-866.

II. Eriksson L, Hall P, Czene K, dos Santos Silva I, McCormack V, Bergh J, Bjohle J, Ploner A: **Mammographic density and molecular subtypes of breast cancer**. *Br J Cancer* 2012, **107**(1):18-23.

III. Eriksson L, Czene K, Rosenberg L, Humphreys K, Hall P.
Mammographic density influences the risk of local and locoregional recurrence of breast cancer
Submitted

IV. Eriksson L, Czene K, Rosenberg L, Törnberg S, Humphreys K, Hall P.
Influence of mammographic density on survival in interval cancers
Manuscript

CONTENTS

LIST OF ABBREVIATIONS

AD	Absolute density
BMI	Body mass index
CC	Cranio-caudal (view)
CI	Confidence interval
e.g.	Exempli gratia
ER	Estrogen receptor
GWAS	Genome-wide association study
Gy	Gray
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
HRT	Hormone replacement therapy
IC	Interval cancer
i.e.	Id est
IGF	Insulin-like growth factor
MD	Mammographic density
MLO	Medio-lateral oblique (view)
N.B.	Nota bene
OR	Odds ratio
PD	Percentage density
PR	Progesterone receptor
RR	Relative risk
RRR	Relative risk ratio
SD	Standard deviation
SNP	Single nucleotide polymorphism
Sv	Sievert
TDLU	Terminal duct lobular units

1 INTRODUCTION

Breast cancer is the most common cancer in women in Sweden and worldwide. Survival has improved greatly over the past decades, due to, among other things, mammography screening and adjuvant therapy. However, ~1500 women still die of breast cancer each year in Sweden. Although 5-year breast cancer survival is high, which for most diseases reflects the curability of disease, a woman is unfortunately never cured of breast cancer. She may therefore die from the disease even 20-30 years after diagnosis.

Over the past decade, there has been growing evidence that breast cancer is not one disease, rather it is composed of several different subtypes. These are heterogeneous with respect to etiology, tumor characteristics, response to treatment, and prognosis. Certain tumors could be described as indolent whereas others are highly aggressive. Unfortunately, the currently clinically available prognostic factors are unable to correctly stratify tumor subtypes; some women will therefore die despite adjuvant therapy, and some women will receive adjuvant therapy who do not need it.

Mammographic density is one of the strongest risk factors of breast cancer but its role in breast cancer development is poorly understood. The studies that form the basis of this thesis aimed to investigate the role of mammographic density in carcinogenesis and tumor progression. Is mammographic density a seedbed for all types of breast cancer, or is it associated with a certain subtype or behavior? We hereby hoped to shed light on breast cancer biology and to add valuable information to established prognosticators in order to improve breast cancer care.

2 BACKGROUND

2.1 THE NORMAL BREAST

2.1.1 Breast anatomy and physiology

The breast lies on top of the pectoralis major muscle. It is composed of parenchyma, stroma, and fatty tissue as well as skin and subcutaneous tissue. Bands of connective tissue called Cooper's ligaments are attached to the fascia of the skin and the pectoralis major muscle, holding the breast in place. As women age, these ligaments relax, leading to ptosis of the breast. Dimpling of the skin, a sign of breast cancer, is caused by tumor invasion of the ligaments of Cooper.

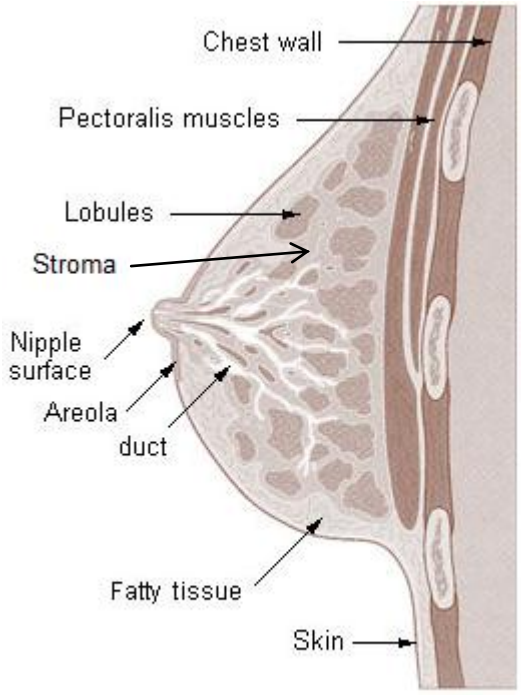


Figure 1. The anatomy of the breast (1)

The breast parenchyma consists of 15-20 lobes each drained by one major lactiferous duct. Every lobe of the breast contains a system of ever-branching ducts ending blindly in a network of terminal ductules. A cluster of terminal ductules and their duct of origin compose a lobule which is surrounded by a specialized stroma. The terminal ductal lobular units (TDLU:s) include the terminal ductules, and their specific extra- and intralobular terminal duct (Figure 2). The TDLU:s are the functional units of the breast, producing and secreting milk. It is within these that most breast cancers occur; lobular cancers originating from the terminal ductules and ductal cancers from the ducts. The ducts are lined by an inner layer of luminal epithelial cells consisting of secretory cells, hormone receptor-expressing cells (2), and progenitor cells (3, 4), and an outer layer of basal cells consisting of myoepithelial cells (2), which have the ability to contract when stimulated by oxytocin, and mammary stem cells (5, 6). Milk is transported from the TDLU:s through interlobular ducts which merge and drain into the major lactiferous ducts. The major lactiferous ducts dilate into sinuses beneath the areola in which milk

can accumulate and these are in turn connected to 6-10 openings in the nipple through which the milk exits.

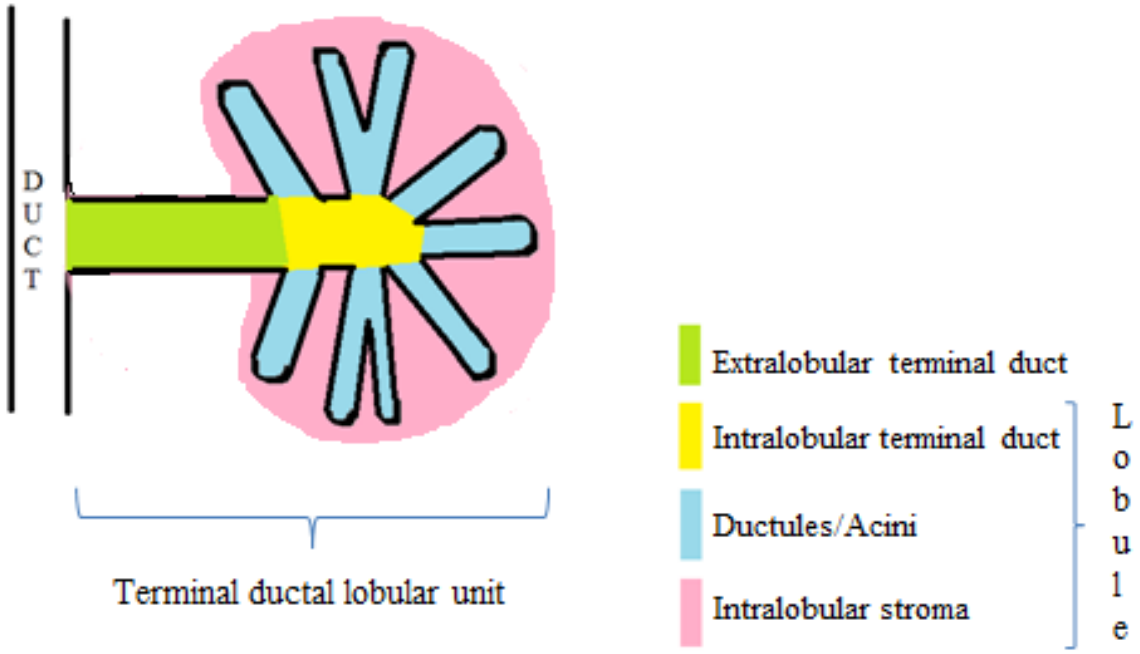


Figure 2. Schematic picture of the terminal ductal lobular unit

Although the breast epithelium is the functional part of the breast, it is only a minor component of the breast; the majority of the breast is composed of stroma and fatty tissue. During menopause the amount of epithelium diminishes as it involutes and is replaced by fatty tissue (see 2.1.2 Breast development).

For clinical purposes, the breast is further divided into the area behind the areola and the quadrants which compose the rest of the breast. The upper, outer quadrant contains the most breast tissue, and it is also here most breast cancers occur.

The breast parenchyma and nipple-areola complex are supplied with blood through several arteries, of which the internal mammary artery accounts for more than half of the blood supply to the entire breast parenchyma. The circulus venosus which is situated around the nipple, carries blood from the breast towards the axilla into the internal thoracic and axillary vein. The breast is thus richly vascularized and hematogenic, metastatic spread is not uncommon as the disease progresses. The most common sites of metastasis are bones, liver, and lungs (7).

Most of the lymphatic drainage of the breast goes to the axillary lymph nodes (75%), which is why it is the most common initial location of metastasis. The number of lymph nodes varies between individuals and also depends on what is defined as the axillary lymph nodes since this depends on method of dissection, but the span of 12-40 axillary lymph nodes is commonly reported (7). Other regional lymph nodes are situated within the breast (the internal mammary chain), parasternally, and above and below the clavicula.

2.1.2 Breast development

Breast development starts during embryonic life but is only fully developed and differentiated by the end of the first full-term pregnancy. During the fifth week of embryonic life, the 'milk streak' forms, which is a thickening of the ectoderm from the axilla to the groin (8). During weeks 6-7 the mammary crest forms which is a thickening in the thoracic region, and the rest of the milk streak involutes (8). This is also when the nipple primary bud takes form. The primary bud penetrates the mesenchyme and eventually leads to the development of 15-20 secondary buds which become the lactiferous ducts and their branches (9).

During puberty, lobule formation occurs (9) and the deposition of fatty tissue increases breast size (10). Still 80% of the breast consists of stromal tissue (8) which includes fibroblasts, endothelial cells, immune cells, and extracellular matrix. During pregnancy, the distal parts of the ducts proliferate. The epithelial cells in the acini both increase in number and size, and differentiate into secretory cells that can synthesize and secrete milk (9). The breast has only now become fully differentiated.

The mammary gland changes little morphologically during lactation. However, cessation of lactation causes post-lactational involution to occur. During this process, epithelial and stromal tissues go into apoptosis and regress, and stromal remodeling ensues (11). The lobules go into a resting non-secretory state, but still appear more complex/developed than the lobules of nulliparous women (11).

Involution also occurs as an aging-process (12) and is especially marked during menopause (10) in response to the ceasing ovarian hormonal secretion (11). Involution results in a decrease of both number of acini per lobule and size of acini. Furthermore, the intralobular stroma is replaced by denser collagen to finally be replaced by fatty tissue (12).

2.2 BREAST CANCER

2.2.1 Descriptive epidemiology

2.2.1.1 Incidence

Breast cancer is the most common cancer in women worldwide. 1.38 million new breast cancer cases were diagnosed 2008 (23% of all cancers in women) (13). However, incidence rates vary from 19.3 per 100,000 women in Eastern Africa to 89.9 per 100,000 women in Western Europe (13). The differences in incidence are due to variations in environmental factors rather than genetic factors (14, 15), which is well illustrated by migrant studies which show an increase in risk for migrants moving from a low-risk to a high-risk country (16) and that the risk increases from generation to generation (17). Thus, parallel with the progress of the developing countries and the adoption of a more westernized lifestyle, incidence of breast cancer is also increasing in these countries (18).

Breast cancer is also the most common cancer in women in Sweden; approximately 8000 women developing breast cancer every year (19). The number of effected women has been steadily increasing by 1.3% since 20 years, but the last 10 years, the increase has been weaker (0.9%) (20). This may partially be due to the drastic drop in hormone-replacement therapy prescription which decreased after the Women's Health Initiative (WHI) hormone trial showed adverse effects of hormone-replacement therapy (21). Lambe et al. (22) studied the incidence of breast cancer in postmenopausal women in

Sweden between 1997 and 2007 and the possible correlation with hormone-replacement therapy prescription. They found that breast cancer incidence had significantly decreased from 2003 in the age group 50-59 and concluded that this decrease was likely due to a decline in hormone-replacement therapy use(22). A decline in breast cancer incidence has been noted in several Western countries (23-31) and saturation in mammography screening has been proposed as another possible explanation to the decline in breast cancer incidence (32).

2.2.1.2 Mortality and survival

Mortality is usually measured as a rate and refers to the rate of individuals dying from a particular disease compared to the general population. Mortality is influenced by the incidence and survival of disease and is a particularly important measure of the burden of disease (33). Conversely, survival rates measure the rate of dying within a cohort of individuals with the disease. Survival is not an adequate measure for e.g. evaluating screening programs since these programs aim to diagnose disease earlier, which will increase survival "artificially" (see *2.2.5.2* Mammography screening). Thus, mortality is the preferred estimate for evaluating secondary preventive measures. However, to the individual patient, survival is of utmost importance since it estimates the risk of that individual dying compared to other individuals with the disease.

Since breast cancer survival is much higher in developed regions where the incidence is highest, the range of mortality rates is smaller (~6-19 per 100,000) than the range of incidence rates (13). It is the most common cause of death among women worldwide (269,000 deaths in developing regions and 189,000 deaths in developed regions) (13). In Sweden, breast cancer was also the most common cause of death from cancer in women until 2005, but is now the second most common cause of death from cancer after lung cancer (34). In 2011, 1401 Swedish women died from breast cancer (35).

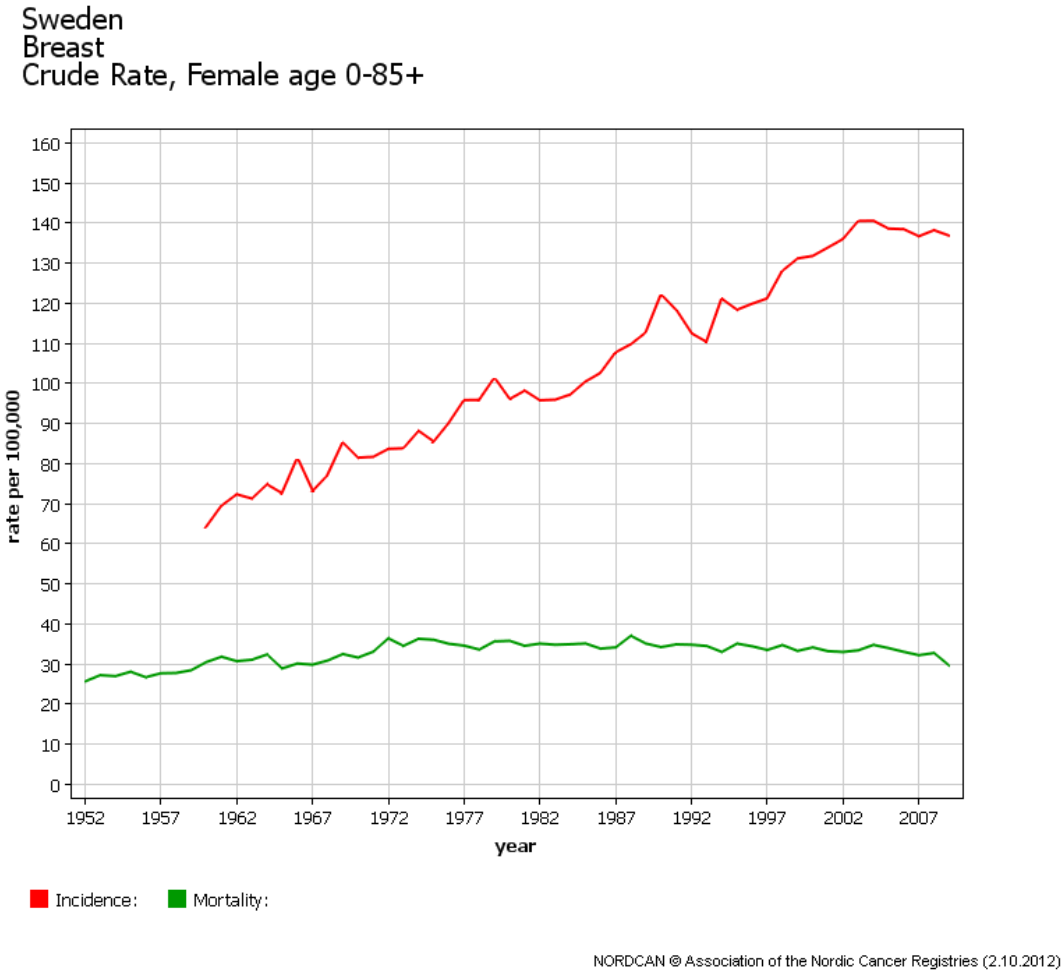


Figure 3. Trends in female breast cancer incidence and mortality in Sweden. Data from Nordcan (Association of the Nordic Cancer Registries).

2.2.1.3 Prevalence

Since breast cancer is relatively common and survival is high (the 5-year breast cancer specific survival is 86%), the prevalence of disease is high. In 2009, approximately 89000 Swedish women were living with breast cancer (36).

2.2.2 Breast cancer biology

2.2.2.1 Carcinogenesis

2.2.2.1.1 The somatic mutation theory

Cancer evolves through a multi-step process which is initialized by the genetic mutation of a single cell, either caused by environmental factors, germ-line mutations, or spontaneously. The mutated cell gives rise to a clonal expansion of genetically damaged cells, in which further mutations in oncogenes and tumor suppressor genes lead to uncontrollable growth and deranged cellular architecture and orientation, resulting in a cancer cell. During this clonal evolution, the ensuing cancer cells also acquire invasive and metastatic capability. Genetic alterations may continue to occur and accumulate in the cancerous cells, which may affect proliferation rate, the degree of invasiveness, and metastatic potential.

In 2000, Hanahan and Weinberg proposed that there were six essential changes in cell function that lead to cancer development: self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, escape from programmed cell-death (apoptosis), unrestricted replicative potential, sustained angiogenesis, and tissue invasion and metastasis (37). However, as pointed out by Yuri Lazebnik, only the last criterion of tissue invasion and metastasis differentiates malignant tumors from benign tumors (38). This is also an important distinction between in situ cancers (which do not penetrate the myoepithelial basement membrane) and invasive cancers.

Based on the advances in cancer research the past decade, Hanahan and Weinberg updated their proposed hallmarks of cancer in 2011 to include two emerging hallmarks: Evasion of immune destruction and reprogramming of energy metabolism to support unceasing cell growth and proliferation (39). Furthermore, they highlighted the role of genomic instability and tumor-promoting inflammation in cancer progression. Genomic instability increases the amount of mutated cells, leading to a high degree of diversity between cancer clones, which is beneficial for tumor survival. This may, however, also occur through epigenetic alterations such as DNA methylation and histone modifications (39).

2.2.2.1.2 The tissue organization field theory

A paradigm shifting hypothesis pertaining to the development of somatic cancer is the 'tissue organization field theory' proposed by Ana Soto and Carlos Sonnenschein in 1999 (40). It suggests two fundamental principles. First, in contrast to the prevailing paradigm of the default state of cells being quiescence, it proposes that the normal state of cells is proliferation. The second premise is that carcinogenesis is a collapse of tissue organization rather than an acquisition of crucial genetic mutations. Although proliferation is presented as the default state of cells, cell-adhesion-dependent tissue architecture is believed to restrict proliferation (41). A disturbance of stromal-epithelial interactions may cause a dysfunction of cell adhesion and/or a loss of tissue organization, which would be able to stimulate normal cells into transforming into malignant counterparts even in the absence of genetic mutations (41). Moinfar et al. (42) observed that genetic alterations and loss of heterozygosity was common in the stromal DNA adjacent to the primary breast tumor. Even more intriguing, they found that genetic alterations in the stroma seemed to precede genetic changes in the epithelium in some cases.

In line with the above, tissue reorganization affects tissue stiffness, which is now known to be an important modulator of cell proliferation, survival, migration, and differentiation (43). In recent years, tissue stiffness has also been shown to be associated with both breast tumorigenesis and progression (44). In the same study, it was shown that tissue stiffness was associated with collagen cross-linking, and the authors remarked that this may explain why mammographic density increases breast cancer risk (44).

2.2.2.1.3 Cancer is a tissue disease

Until recently, tumors were viewed as a homogeneous group of cancer cells. However, it is now known that tumors are heterogeneous, including cancer stem cells, cancer-associated fibroblasts (CAFs), immune inflammatory cells, pericytes, and endothelial cells apart from the cancer cells themselves (39). These cells, including the surrounding extracellular matrix, are referred to as the tumor microenvironment. Interaction

between the tumor microenvironment and the tumor cells has been shown to be essential to tumor progression (39). CAFs, for example, effect cell proliferation, angiogenesis, invasion, and metastasis (39).

The stromal cells of the microenvironment are recruited from adjacent normal tissue (either as preexisting stromal cells or as stem/progenitor cells,) as well as from the bone marrow (39). When the mesenchymal stem and progenitor cells reach the tumor, they may either remain undifferentiated, become partially differentiated or fully differentiate into the different stromal cell types (39).

2.2.2.2 Metastasis

Metastasis occurs when tumor cells from the primary tumor migrate to distant sites in the body and form a secondary mass at that place. For this to happen, a tumor cell must detach from adjacent cells and invade the surrounding stroma. It then has to be able to enter lymphatic channels or blood vessels (intravasation), survive the mechanical forces of circulation, avoid circulating immune cells, and exit the vessels through a process called extravasation (45). Once at the new site, the tumor cell must adapt to the new environment. It also needs to acquire the ability to proliferate without the stimulation it formerly received from the microenvironment of the primary tumor, and in spite of inhibitory signals from the new microenvironment (46). Finally, it needs to be able to develop a vascular system that can support proliferation (47).

As previously stated, metastasis can either spread via the lymphatic channels or blood vessels. In breast cancer, the former gives rise to lymph node metastasis most often occurring in the ipsilateral axilla, whereas hematogenic/systemic spread usually gives rise to metastasis in the bones, lungs, liver, or brain. Breast cancer that is spread to any of the latter sites is referred to as distant metastasis. It is not uncommon to present with axillary lymph node metastasis at diagnosis, and this is still a curable disease since the spread is well-confined. The opposite is true for distant metastasis, which is uncommon at first presentation, can affect several organs at once and several parts of the same organ, and for which only palliative therapy exists (7).

2.2.2.3 Recurrence

Recurrence of disease refers to relapse at any site, i.e. locally (in the residual breast, chest wall, or scar tissue), regionally (in the regional lymph nodes), or in distant organs (distant metastasis). Local and regional recurrence can also be combined and is then referred to as locoregional recurrence. Unlike distant metastasis, which is closely related to breast cancer death, locoregional recurrence is only moderately associated with breast cancer survival; for every four locoregional reccurences that are avoided, one breast cancer-specific death over the next 15 years is evaded (48).

2.2.3 Risk factors

2.2.3.1 Sex

Breast cancer is the most common cancer in women, while it is a very rare disease in men. In 2010, almost 8000 women were diagnosed with breast cancer in Sweden and only 33 men (49). Sex is therefore a very strong risk factor of breast cancer.

2.2.3.2 Age

Breast cancer incidence increases with age and is relatively rare in women less than 40 years of age, after which the incidence increases greatly (7, 33) (Figure 4). Conversely, the rate of the age-specific incidence rises steeply until around 50 years of age (i.e. around menopause), after which the increase is less pronounced (33, 50), suggesting that hormones are important to breast cancer development (33).

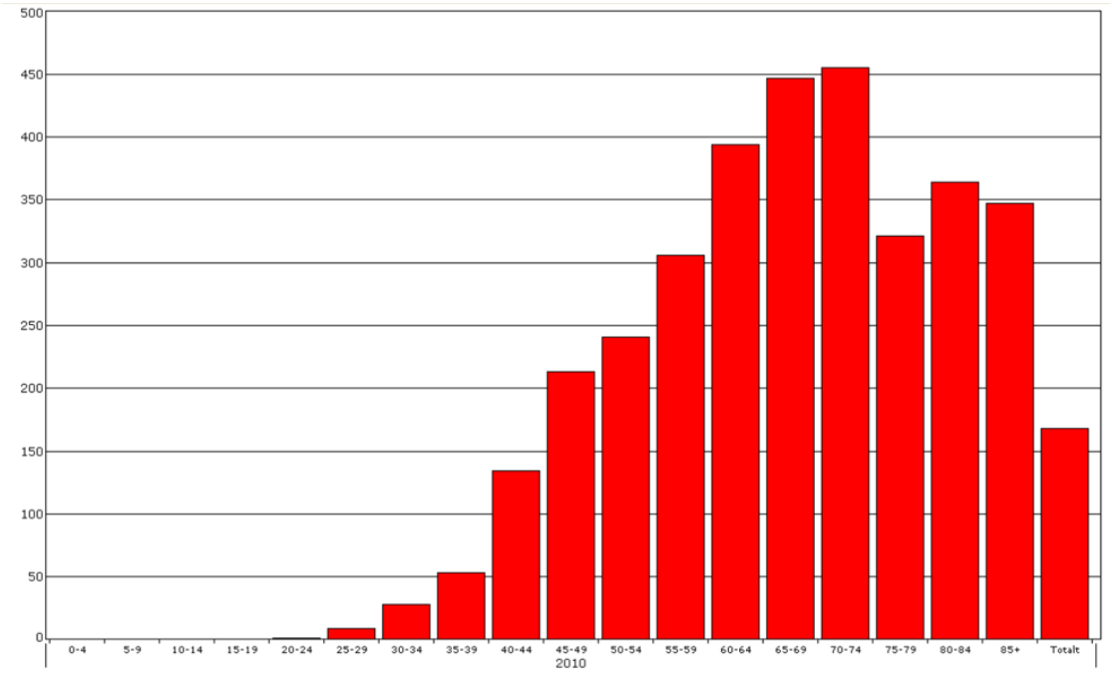


Figure 4. Breast cancer incidence in women per 100,000 individuals. Sweden, 2010. Data from the Swedish Cancer Registry.

2.2.3.3 Mammographic density

Mammographic density is one of the strongest risk factors of breast cancer. When measured quantitatively, a percentage density (PD) ≥75% has a 5-fold increased risk of breast cancer compared to women with PD<5% (51) (see 2.3 "Mammographic Density" below).

2.2.3.4 Heritability

A family history of breast cancer in a first-degree relative increases risk of breast cancer by approximately 2-fold (52), and a family history of more than one first-degree relative confers an even greater risk. Approximately 16% of breast cancers in women with a positive family history are thought to arise due to mutations in the high penetrance susceptibility genes *BRCA1* and *BRCA2* (53). High-penetrance susceptibility genes are characterized by the fact that carriers have a high likelihood of developing disease. The lifetime risk of developing breast cancer is 45-65% for carriers of *BRCA1* and *BRCA2* mutations (54), and is even higher for early-onset breast cancers (55). Furthermore, *BRCA1* and *BRCA2* both increase risk of ovarian cancer. There are four other known high penetrance genes associated with breast cancer,*TP53*, *PTEN*, *STK11*, and *CDH1*, but mutations in these genes are very rare and together all six genes

are only believed to account for 20% of the familial risk of breast cancer (55). Four, intermediate penetrance genes have also been identified: *CHEK2*, *ATM*, *BRIP1* and *PALB2*. These increase breast cancer risk by 2-4 fold but are also believed to be uncommon (55). Thus, the majority of breast cancers, even cases with family history, are thought to be polygenic, i.e. attributable to a combination of genetic variants each conferring a low risk. In order to find these low penetrance alleles, genome-wide association studies (GWAS) are carried out to compare single nucleotide polymorphisms (SNPs) (common inter-individual or inter-allelic base pair differences) in breast cancer cases and controls. However, since effects of the individual SNPs on breast cancer risk are small, large study populations are needed. GWAS are conducted without *à priori* knowledge and are thus hypothesis-generating. Indeed, GWAS has managed to shed new light on breast cancer biology; new susceptibility genes have been identified that are not, as the previously identified susceptibility genes (e.g. BRCA1/2, BRIP1, CHEK2), associated with DNA repair and sex hormone synthesis, but with cell proliferation and cell signaling, e.g. FGFR2 and MAP3K (56).

2.2.3.5 Endogenous hormones

Breast cancer is influenced by many hormonally related factors (see below) and it has therefore long been assumed that high levels of endogenous sex hormones are partially at fault. This was also corroborated when the Endogenous Hormones and Breast Cancer Collaborative Group showed that postmenopausal women who had increased levels of endogenous sex hormones, also were at higher risk of breast cancer and there was a dose-response relationship (57). Eliassen et al. observed a relationship between higher levels of free estradiol during the follicular phase and breast cancer risk in premenopausal women (58). Two studies that directly assessed endogenous progesterone and its relationship to breast cancer risk showed no associations (58, 59). Missmer et al. also tested for an interaction between estrogen and progesterone but did not find any evidence of this (59), which is surprising since especially combinations of HRT (estrogens in combination with progestagens) increase breast cancer risk (see below). The authors speculated that this could be due to differences between endogenous progesterone's and synthetic progestin's metabolic effects on breast tissue.

2.2.3.6 Reproductive factors

Age at menarche influences breast cancer risk, where older age decreases risk (60). The opposite is true for age at menopause, where older age at menopause increases risk of breast cancer (61). Hence, the longer a woman has menstrual cycles, the higher her risk of breast cancer.

An early age at first full-term pregnancy and number of full-term pregnancies both have protective effects on breast cancer risk, independent of each other (62). The reason for this is thought to be that mammary gland cells are undifferentiated until first pregnancy, and that each pregnancy decreases the number of undifferentiated cells. However, the effect of pregnancy is dual; a full-term pregnancy increases breast cancer risk immediately after birth, but the risk then gradually diminishes and, in the long-term, imparts a protective effect (63). Breast-feeding also has a protective effect on breast cancer risk independent of age at first birth and parity (63, 64). The Collaborative Group on Hormonal Factors in Breast Cancer estimated that each birth

conferred a decrease in relative risk (RR) of breast cancer of 7.0% and every 12 months of breast feeding conferred a decrease in RR of 4.3% (64).

2.2.3.7 Anthropometric factors

As with pregnancy, BMI also has a dual effect on breast cancer. Childhood and adolescent obesity is protective of breast cancer, possibly due to obesity being associated with fewer menstrual cycles (65). Conversely, in postmenopausal women, obesity increases risk of breast cancer. This may be due to obese postmenopausal women having higher levels of bioavailable estradiol (65), since the main source of estrogens in postmenopausal women is the conversion of androgens to estrogens in the adipose tissue (66). Another anthropometric measure associated with breast cancer risk is height, which is weakly and positively correlated with risk (67); a 10-cm increment in height increasing RR by 17%. The cause of this relationship is, however, unclear.

2.2.3.8 Exogenous hormones

The use of combined oral contraceptives is related to an increase in breast cancer risk of about 25% but is only associated with current or recent use (63). 10 years after cessation there seems to be no increased risk (63).

Hormone-replacement therapy (HRT), given as a combination of estrogen and progesterone or estrogen alone, increases breast cancer risk, where combination treatment increases risk the most (68). Risk also increases with duration, e.g. in women on combination therapy, breast cancer risk increases by 7.6%/year (69). However, the increased risk gradually disappears within 2 years of cessation (70). The effect of HRT on breast cancer risk is larger in lean women (68) and women with high mammographic density (71).

2.2.3.9 The intrauterine environment

In 1990, Dimitrios Trichopoulos hypothesized that breast cancer may originate in utero (72). He suggested that the high estrogen levels created a “fertile soil” that was more susceptible to cancer initiation. Many studies have evaluated this hypothesis, mainly using birth weight as a proxy of intra-uterine exposure to estrogens and other growth factors such as insulin and the insulin-like growth factors (IGF) I and II which also may influence breast cancer risk (73). Three meta-analyses have confirmed that there is a positive correlation between birth weight and breast cancer risk (73-75), especially in pre-menopausal breast cancer (73, 74). The intrauterine environment thus seems to influence mammary carcinogenesis. However, whether this is due to increased growth factors and the mechanism by which they would act, remains to be revealed.

2.2.3.10 Previous benign breast disease

An important distinction of benign breast disease when discussing its relation to breast cancer, is whether it is a non-proliferative or proliferative lesion, and whether the latter has presented with or without atypia. Non-proliferative lesions do not or only slightly increase breast cancer risk (76), whereas proliferative lesions are strong risk factors for breast cancer. Proliferative lesions without atypia are associated with a doubling of breast cancer risk, and proliferative lesions with atypia increase risk by at least 4-fold (63).

2.2.3.11 Lifestyle factors

Alcohol intake is associated with a moderate increase in breast cancer risk (77), while most studies have not found an association between smoking and breast cancer (78). Moderate physical activity seems to have a protective effect on breast cancer risk; a few hours of vigorous exercise per week may decrease risk by about 30% (63). However, studies on the relationship have been somewhat inconsistent and the effect of physical activity on breast cancer risk has varied between studies (63).

Although several factors associated with different aspects of our diets have been hypothesized to be associated with breast cancer risk, associations are modest at best and results inconsistent (33). However, for vitamin A, carotenoids, and folate a modest protective effect on breast cancer risk has consistently been shown (33).

2.2.3.12 Ionizing radiation

Studies of the effects of ionizing radiation on breast cancer risk have primarily been carried out on Japanese atomic bomb survivors and patients exposed to ionizing radiation in the diagnostic or therapeutic, clinical setting. There was a large variation in the doses the different populations received to the breast; from 0.02 Sievert (Sv) to >20 Sv (79). According to these studies there seems to be a dose-response relationship between ionizing radiation and breast cancer risk (79). Furthermore, risk seems to attenuate with age (and breast maturation) so that girls who are exposed before 20 years of age have the highest risk of breast cancer, whereas the risk in postmenopausal women is negligible (79). There may also be an increased breast cancer risk associated with ionizing radiation in pregnant women (79).

Table 1. Risk factors of breast cancer

Risk factors	**Strength of association***	**Clarifications**
Sex	↑↑↑	Females
Age	↑↑↑	Risk increases with age
Mammographic density	↑↑↑	PD≥75% compared to PD<5%
Geographical region	↑↑↑	Developed countries vs. developing countries
Previous benign breast disease	↑↑↑	Atypical hyperplasia
Family history	↑↑	First-degree relative, also increases with number of effected relatives
BMI postmenopausal	↑↑	High BMI increases risk
HRT	↑↑	Especially combined HRT
Age at menopause	↑↑	Older age
Age at menarche	↑↑	Younger age
Age at first birth	↑↑	>30 compared to <20
Ionizing radiation	↑↑	Before age 20
Alcohol intake	↑	Risk increases with increased intake
Oral contraceptives	↑	Current use
Height	↑	Risk increases with increased length
Birth weight	↑	Risk increases with increased birth weight
Previous benign breast disease	↔	Fibroadenomas
Breast feeding	↓	Longer periods of breast feeding decreases risk
Parity	↓	Long-term protective effect of parity
BMI premenopausal	↓	High BMI is protective
Physical activity	↓	Physical activity is protective

*↑↑↑ Strong risk factor; ↑↑ Moderate risk factor; ↑ Weak risk factor; ↔ no association; ↓ protective factor.

2.2.4 Risk prediction models

The Gail model was originally developed by Gail et al. in 1989 (80) and is now a publicly available tool for prediction of breast cancer risk (http://www.cancer.gov/bcrisktool) . The model uses the variables age, age at menarche, age at first birth, family history of breast cancer in a first-degree relative, number of previous breast biopsies, atypical hyperplasia, previous breast cancer

(invasive and in situ), and ethnicity to calculate risk (atypical hyperplasia and ethnicity have been added later on (81)). One caution with the Gail model is that it has not updated the age-specific incidence rates on which it is based, and this may be one of the reasons that it underestimates breast cancer risk on the individual level (82).

Since the Gail model mainly includes non-genetic risk factors, the Claus (83) and Ford (84) models were developed which both focus on family history of breast cancer. The latter also includes family history of ovarian cancer to identify individuals with BRCA mutations. The Tyrer-Cuzick model tries to incorporate both information on endogenous estrogen exposure, history of benign breast disease, and a comprehensive family history assessment (85). According to an evaluation of the four different models (86), they all performed similarly pertaining to individual risk assessment, but the Tyrer-Cuzick model was somewhat better at predicting overall number of cases.

A limitation of all models is that none of them have been validated in the general population (87). Furthermore, none of the models include mammographic density or genetic alterations. There have been attempts to incorporate mammographic density into risk prediction models, but this has only lead to moderate improvements in prediction (87). However, this may be explained by the correlation between mammographic density and the other breast cancer risk factors such as parity, age at first birth, and benign breast disease. Tice et al. (88) namely showed that a model including only age, ethnicity, and mammographic density performed as well as the Gail model. Further, Barlow et al. (89) showed that both age and mammographic density were highly predictive when studied independently, but that adding mammographic density to a model including traditional risk factors did not increase predictive power.

2.2.5 Diagnostics

2.2.5.1 Triple diagnostics

Breast cancer is diagnosed through triple-diagnostics, referring to the triad mammography, clinical assessment, and cytology. Nowadays, ultrasound is also often added as a complement to mammography, and courser needle biopsies may be taken in addition to cytology (7). About half of all breast cancers diagnosed in Sweden today are diagnosed within the screening program (7). These are usually asymptomatic cancers (7). The rest of the cancers are clinically detected as symptomatic cancers and most often present as a lump in the breast, but other symptoms such as a lump in the axilla, breast discomfort, redness of the skin, or nipple secretion can also occur. The majority of women diagnosed with symptomatic cancers are women who are not invited to screening, but also consist of women who were invited but chose not to attend screening (7). Furthermore, breast cancer may be detected in the interval between two screening examinations as a symptomatic cancer, either because a cancer was missed at the previous screening or because it is highly proliferative (see 2.2.5.2.2. Interval cancers).

High mammographic density lowers mammographic sensitivity (90). Consequently, other breast imaging modalities are used in populations where high mammographic density is prevalent. Ultrasound is for instance preferred in young (<30 years of age), pregnant, and lactating women (7), all of whom have high mammographic density. In women with a strong family history or carriers of e.g. BRCA1 or BRCA2 mutations,

who are recommended yearly follow-up from age 25-35 years of age (91-93), MRI is preferred also due to the younger age group and higher mammographic density. However, MRI is also favored in these women since BRCA2 carriers have an increased risk of lobular cancer (94) which are difficult to detect on a mammography (95), and BRCA-associated tumors may be more likely to have a benign appearance on mammography (96, 97). MRI is not only used in individuals with family history, but may also be used to further investigate unclear radiologic findings. The drawback of MRI is that, although sensitivity is high, specificity is lower than that of mammography (98), increasing the risk of false positive findings.

Galactography/ductography is also an alternative radiologic examination of the breast if a woman suffers from serous or bloody discharge from the nipple. It is carried out by injecting contrast material into the secreting duct in order to visualize intraductal proliferative disease, e.g. papillomas, and also allows for cytological examination of the discharge. Sensitivity is, however, low (7).

2.2.5.2 Mammography screening

2.2.5.2.1 History and guidelines

Mammography screening was implemented in Sweden in order to decrease breast cancer mortality. Screening programs were introduced in 1986 and were implemented in all of Sweden by 1997. The initial recommendations that came in 1986 included women from 40-74 years of age. However, two amendments were made in 1987 and 1988, respectively, allowing county councils that were short of staff to only include women between 50-69 years of age (99). Different county councils therefore included different age groups. In 1995-1996, for example, 11 out of the 25 participating county councils invited women 40-74 years old, six invited women 50-69 years old, and the eight remaining invited women between these two age spans (99).

The interval between screening examinations is between 1,5-2 years in Sweden. It is based on the time difference between the point at which a tumor first becomes detectable on a mammography (at 3-4 mm) and the point at which it becomes symptomatic. The length of this pre-symptomatic phase at which the tumor is detectable at screening has been estimated to be approximately 1,5-3 years in most cases (100).

2.2.5.2.2 Interval cancers

Interval cancers are symptomatic cancers diagnosed in the interval between two screening examinations, where the prior screening mammography was negative (or inconclusive with a negative follow-up). Since mammography does not have 100% sensitivity, some cancers will not be detected at mammography screening either due to observer error, unspecific findings, or masking, and therefore diagnosed as interval cancers. These are viewed as “false” interval cancers. Highly proliferative tumors may, however, also arise in the interval between two screening examinations, and these are referred to as “true” interval cancers. The latter are thought to be more aggressive and have a poorer prognosis than screening-detected cancers (101).

The number of interval cancers is used as a measure of the quality of the screening program. For the quality of the screening program to be acceptable, the percentage of interval cancers should not exceed 30% of the estimated breast cancer incidence

without screening the first year after a negative screen, and not exceed 40% the second year (100).

2.2.5.2.3 Views

The mammogram views routinely used in screening programs are the medio-lateral oblique (MLO) and the cranio-caudal (CC) views. The CC view is taken straight from above and the MLO view is taken at an angle of 30-60 degrees medially of the CC view. The MLO view constitutes the basis of the Swedish mammography screening program, but CC views are often added to increase detection rate. At the first mammography a woman ever does (the baseline mammography), both the CC and the MLO views are taken. When performing a diagnostic mammogram, both the MLO and CC views are used and a lateral view (90 degrees lateral of the CC view) is usually also added to further increase the detection rate.

2.2.5.2.4 Biases in estimating the effect of screening on mortality

The aim of mammography screening is to reduce breast cancer mortality. However, comparisons between screening- and non-screening-detected cancers to evaluate the efficiency of mammography screening suffer from several biases e.g. selection bias, lead-time bias (see figure 5), and length bias (see figure 6). Selection bias refers to women attending screening being different from non-attenders; they are more often of higher socio-economic status and health-conscious which increases their life-expectancy. Lead-time bias refers to tumors diagnosed at screening being detected earlier during their natural history than tumors detected otherwise; screening will then advance detection of disease and result in a survival advantage independent of whether it actually prolongs survival or not. Tumors detected by mammography screening are generally therefore of lower stage than symptomatic cancers, a phenomenon called stage shift (102). The length bias refers to differences in tumors' proliferation rates; indolent tumors spend a longer time in an asymptomatic phase and are thus more likely to be detected at screening, whereas fast-growing tumors have a shorter asymptomatic phase and are thus more likely to be detected between screening rounds.

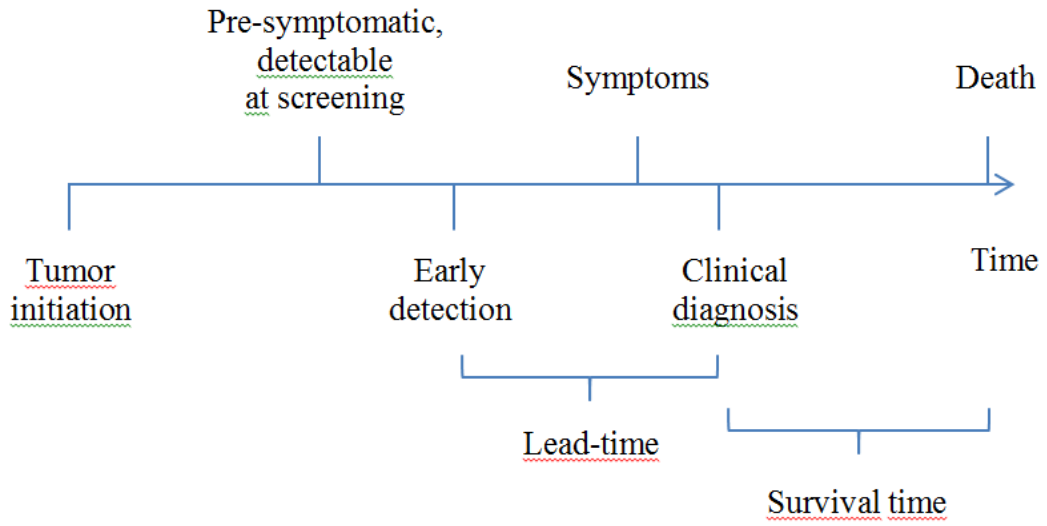


Figure 5. Lead-time bias refers to the artificial survival advantage a pre-symptomatic tumor has compared to a clinically detected tumor, due to the former being diagnosed earlier during its natural history.

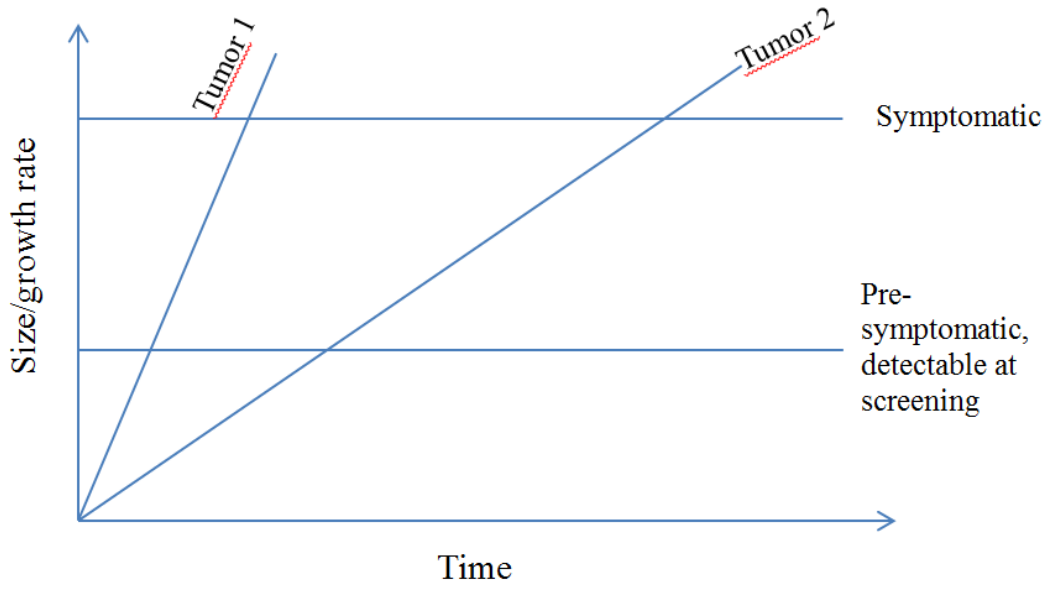


Figure 6. Length bias. Tumor 1 is a more highly proliferative and aggressive tumor with a much shorter pre-symptomatic phase than tumor 2 which is a more indolent tumor. Tumor 2 will thus more likely be diagnosed at screening than tumor 1 which will more likely be diagnosed between screening rounds. This will affect survival comparisons and is referred to as the length bias.

2.2.5.2.5 The screening debate

Several Swedish randomized controlled trials have been carried out to assess the effect of mammography screening on mortality. They have shown reductions in mortality of approximately 25% for the general population and even more for the women who participated in screening (103-105). However, a Cochrane review by Gotzsche and Nielsen of all eligible and completed randomized trials comparing breast cancer mortality in screened vs. non-screened women, has questioned the results of the trials (106). Among other things, the trials were accused of having had a suboptimal randomization process and biased assessment of breast cancer mortality (this critique was later refuted by Nyström et al. (104)). Yet, in spite of these objections, it was still calculated that mammography screening reduces breast cancer mortality by 15-20% (106). However, it was also estimated that screening led to an over-detection of breast cancer cases by ~30%, and the authors concluded that it is "not clear whether screening does more harm than good."(106)

The issue of over-detection is important, but a review in 2007 of the studies assessing over-detection (107) concluded that all the estimates were biased and deemed the studies inadequate. They specifically scrutinized the estimates presented by Gotzsche (108) (among others) and found that lead-time had not been taken into consideration. Lead-time influences the comparison of incidence in screening and control groups since it increases the incidence rate of breast cancer in the screening group at both the prevalent (the first screen a woman ever does) and subsequent, incident screens. After screening stops, the incidence rate drops in the previously screened group. This is due to the fact that cancers that would have emerged as symptomatic cancers in this period have already been diagnosed at screening. After some years (no exact time point), the cumulative incidence will even out between both groups, which is why the comparison

of the cumulative incidence should not be made until several years after screening has ended. Gotzsche based his comparisons on the cumulative incidence directly at the end of screening, and did not statistically adjust for lead-time. This biased all his estimates upward, including the estimate of 30% which is referred to in the Cochrane review.

An interesting study from Norway, contributing to the debate about screening, was published in 2010 (109). In this study it was shown that there was an effect of mammography screening on breast cancer mortality, but that a large part of this reduction was attributable to improvements in the healthcare associated with screening, e.g. the implementation of interdisciplinary breast cancer teams.

In summary, the studies above all show that mammography screening decreases breast cancer mortality. However, more and non-biased studies are needed to evaluate the extent of over-detection to be able to adequately assess the pros and cons of screening. Furthermore, I personally believe that much would be gained if we could improve screening sensitivity for women with highest density, since these are women at high risk of breast cancer and where mammography has lowest sensitivity.

2.2.6 Breast cancer classification

Cancers of the breast are almost always adenocarcinomas which originate from the epithelial cells. Even so, breast cancer is heterogeneous and can be classified according to invasiveness, tumor characteristics, and gene expression patterns.

2.2.6.1 Invasiveness

Breast cancer is divided into non-invasive and invasive breast cancer. Non-invasive cancer is further subdivided into lobular cancer in situ (LCIS) and ductal cancer in situ (DCIS). Non-invasive breast cancer has become common since the introduction of mammography screening and now accounts for 15-20 % of all breast cancers diagnosed in Sweden (7). Throughout the studies of this thesis, the term breast cancer refers to invasive breast cancer only.

2.2.6.2 Histopathological classification

The most common type of invasive cancer histopathologically is ductal cancer, which accounts for 40-75% of all invasive breast cancers (7). The second most common type of cancer is lobular cancer (5-15%) followed by tubular cancer (2-7%), medullary cancer (1-7%), invasive cribriform cancer (0.8-3.5%), mucinous cancer (2%), invasive papillary cancer (<1-2%), and invasive micropapillary cancer (<2%) (7). There are other types as well but these are even less common.

2.2.6.3 Stage

Stage is the most important prognosticator to date and is assessed using the TNM-classification, which incorporates the variables tumor size (T), lymph node involvement (N), and presence of distant metastasis (M). Lymph node involvement is the single most powerful predictor of prognosis in patients who do not have distant metastasis (110), the number of metastatic lymph nodes also strongly, negatively influencing survival (111).

2.2.6.4 Histologic grade

Tumor grade was introduced by Greenough in 1925 (112), modified by Bloom and Richardson in 1957 (113) and later on in 1991 by Elston and Ellis (114). The latter classification is the grading system currently used in Sweden and is commonly referred to as the Nottingham histologic grade (NHG), in which a score from 1 to 3 is given according to the tubular formation, size and appearance of nuclei (nuclear atypia), and number of mitotic cells. The more similar to normal breast tissue the tumor tissue appears and the lower the proliferative activity, the lower the grade, and vice versa. Grade is a prognostic factor independent of tumor size and lymph node involvement but, unlike these factors, is only moderately reproducible (115).

2.2.6.5 Proliferation rate

Cell division is carefully regulated in normal cells but not in cancer cells where one of the most important features is uncontrollable cell growth. All cells go through a cell cycle consisting of five phases; G0, G1, S, G2, and M. G0 is a resting phase. During G1, cells increase in size. This is also a checkpoint that the cell is ready to enter S-phase, the synthesis phase, during which DNA replication takes place. During G2 the cell continues to grow. This is also a checkpoint to see that the cell is ready to enter M-phase. If ready, the cell enters M-phase, when mitosis (cell division) takes place, producing two identical daughter cells.

During the time period of studies I, III, and IV, proliferation rate was measured using flow cytometry as the percentage of cells in S-phase. Nowadays, assessment of Ki67 and cyclin A, proteins produced during the proliferative part of the cell cycle (i.e. not during G0), is carried out instead, using immunohistochemistry. A tumor's proliferation rate influences prognosis, but whether it predicts response to chemotherapy is uncertain (116).

2.2.6.6 Hormone receptors

Breast cancers can be categorized based on estrogen receptor (ER) and progesterone receptor (PR) status. This is assessed using immunohistochemistry and predicts whether the primary tumor/metastasis will respond to endocrine therapy. A tumor is usually considered positive if >10% of the nuclei are stained, but tumors with only 1-10% stained nuclei may also benefit from endocrine treatment (117). ER-status is the most important factor influencing response to endocrine therapy, but PR-status may give additional information to ER-positive tumors (118, 119). Apart from predicting response to endocrine therapy, hormone-receptor status also influences prognosis; hormone receptor-positive cancers conferring a survival advantage compared to hormone receptor-negative cancers (120).

2.2.6.7 HER2 (or HER2/neu or ERBB2)

Human epidermal growth factor receptor 2 (HER2) (also known as HER2-neu and ERBB2) is amplified in 10-30% of all breast cancer cases (7). HER2 mediates growth, differentiation, and survival of cells. Women whose tumors overexpress HER2 thus have a more aggressive disease than women whose tumors do not overexpress HER2, both with respect to disease-free and overall survival (121). Assessment of amplification of the *HER2* gene has come into use in later years and was not carried out during the time of the studies included in this thesis. HER2-status is assessed using immunohistochemistry followed by in situ hybridization in cases where amplification is suspected with immunohistochemistry or where this is inconclusive. Amplification of

the *HER2* gene allows for treatment with a specific antibody against the HER2-receptor (trastuzumab) and was initially introduced in Sweden in 2001 as a treatment of metastasized breast cancer but is now also used approved for treatment of non-metastasized breast cancer. Before the availability of trastuzumab, women with HER2-positive cancers had a particularly poor prognosis, but now these women, if they are treated with trastuzumab, may even have a better prognosis than women with HER2-negative disease (122).

2.2.6.8 Triple-negative disease

Tumors that are ER-, PR-, and HER2-negative are referred to as triple-negative tumors and account for 15-20% of all newly diagnosed breast cancers (123). Triple-negative cancers are associated with younger age, African American ethnicity, and BRCA1-mutations (124). Since they lack ER- and PR-expression as well as HER-2 overexpression/gene amplification, there is no targeted therapy for these cancers, but they respond to chemotherapy (124). Despite this, triple-negative breast cancers still have a higher probability of relapse and poorer overall survival in the first years after breast cancer diagnosis compared to other types of breast cancer (123). Most triple-negative cancers are high-grade tumors and of ductal type. However, both low-grade tumors and tumors of other histopathological classification than ductal may be triple-negative. Most triple-negative breast cancers fall into a specific molecular subgroup of breast cancers, the basal group (see below), but the remainder fall into different molecular subtypes (see below). In other words, the group of triple-negative tumors is diverse (124).

2.2.6.9 Molecular subtypes

Gene expression profiling has highlighted the concept of breast cancer as a collection of different diseases albeit affecting the same organ. Individual subclasses have been identified through gene expression profiling, referred to as the Sorlie–Perou subtypes (125, 126). The heterogeneity between the subtypes is reflected in differences in etiology, clinical presentation, histopathological features, outcomes, and response to systemic therapies (127).

The five original subtypes were the luminal A, luminal B, basal-like, ERBB2, and the normal breast-like subtypes (125, 126). The claudin-low and molecular apocrine subtypes were identified later on (128, 129). Luminal A tumors are mostly ER-positive, have a low proliferation rate, and are of low grade. Luminal B tumors are also mostly ER-positive but may express low levels of hormone receptors, and are usually of high grade and have a higher proliferation rate. The basal-like subtype is often characterized by triple-negative tumors (ER-, PR-, and HER2-negative), high levels of expression of proliferation-related genes, and expression of genes associated with basal and myoepithelial cells. The claudin-low subtype is also often triple-negative. However, the claudin-low tumors are characterized by down-regulation of genes involved in cell adhesion, are more enriched in epithelial-to-mesenchymal transition features, and express stem cell-associated genes (127). The apocrine subtype is usually ER-negative but expresses androgen receptors and androgen receptor-associated genes, and has histological features suggestive of apocrine differentiation (129, 130). The ERBB2 subtype shows amplification and high expression of the *ERBB2* gene (also known as HER2 or HER2-neu). The normal breast-like subtype shows expression of many genes expressed by adipose tissue and other non-epithelial cell types, strong expression of basal epithelial genes, and low expression of luminal epithelial genes. It is, however,

unclear whether the latter subtype is a distinct group or represents poorly sampled tissue (131).

The different subtypes can be used as both prognostic and therapeutic predictive factors. The luminal A subtype, for example, has a better prognosis than the luminal B and basal-like subtypes (127). Further, the basal-like subtype responds well to chemotherapy whereas the luminal A subtype does not (127). However, the additive prognostic and predictive information offered is limited compared to ER-status, PR-status, HER2-status, and Ki67 (127). The clear advantage of molecular subtyping lies in reproducibility and quantitative assessment (127).

2.2.7 Treatment

2.2.7.1 Surgery

Surgery is the primary treatment of breast cancer and many women are cured after surgery or after surgery and radiotherapy combined (7). Depending on the size of the tumor and its relation to the total breast size, tumor stage, number of tumor foci and distance between them, and whether the patient can tolerate radiotherapy or not, either partial mastectomy (breast-conserving surgery) or total mastectomy is performed. In addition to the surgery of the breast, axillary surgery is also carried out. The primary aim of the axillary surgery is to appropriately determine tumor stage in order to plan the adjuvant therapy and also assess prognosis. The therapeutic effect of lymph node dissection is uncertain; although axillary lymph node dissection protects against axillary recurrences, it is unclear whether this effects survival (132-134).

During the time period of the studies within this thesis, i.e. the 1990's, axillary lymph node dissection was the only surgical procedure in routine use. However, axillary lymph node dissection is associated with complications such as seroma, infections, and lymphedema of the arm. Furthermore, mammography screening caused a shift in the stage at which breast cancers were detected, and more and more patients were diagnosed without spread to axillary lymph nodes. This made many question the role of axillary lymph node dissection and the sentinel node biopsy was therefore developed during the end of the 1990's for staging purposes to reduce morbidity. With this method, the first lymph node/s in the axilla draining the breast is/are identified using radioactive tracer and blue dye. This/these lymph node/s are then surgically removed and analyzed immediately (with the patient still on the operating table). If any spread is seen, an axillary lymph node dissection is performed, if not, no further surgery of the axilla is carried out. Sentinel node biopsies are currently only performed in women with unifocal tumors and where preoperative axillary lymph node status is negative.

2.2.7.2 Radiotherapy

Postoperative radiotherapy is routinely given to the residual breast tissue of women treated with breast-conserving surgery. The purpose of the treatment is to decrease risk of local recurrence but postoperative radiotherapy has also been shown to improve breast cancer-specific survival (135). However, radiotherapy also increases risk of cardiovascular events and decreases overall survival due to this (135). Other less severe side-effects include redness of the skin, swelling of the breast, pneumonitis (136), lymphedema of the arm (if the axilla is radiated), and possibly also negative effects on lung function (7). It is therefore of utter importance to balance risk and benefit. International (ASCO and Eusoma) and Swedish guidelines (SweBCG) thus recommend that women at high risk of locoregional recurrence (where risk of recurrence within 10

years is greater than 20%) should receive post-operative locoregional. According to the Swedish guidelines this means that women treated with total mastectomy should receive radiotherapy towards the thoracic wall if e.g. the tumor is >5 cm, or receive radiotherapy towards both the thoracic wall, axilla, and supraclavicular lymph nodes if e.g. there are ≥ 4 metastatic lymph nodes (7). Radiotherapy is usually given in fractions of 2 Gy every day, 5 days per week until the patient has received a total dose of 50 Gy (7).

2.2.7.3 Systemic therapy

The aim of systemic therapy is to annihilate micrometastatic disease. Factors that predict response to systemic therapy are ER-status, PR-status, and HER-2 amplification. Proliferation rate, assessed as e.g. Ki67 may add information as to who will benefit from chemotherapy. Despite these factors, the response to therapy is not certain; individuals will both be under- and over-treated.

2.2.7.3.1 Endocrine therapy

ER-status and PR-status are predictive of a tumor's response to endocrine therapy. A woman with ER-negative cancer does not benefit from endocrine therapy (7). The oldest type of endocrine therapy is the selective estrogen receptor modulator (SERM), tamoxifen. It has been in use since the 1970's and is still an important part of the arsenal of endocrine treatment. Tamoxifen reduces risk of recurrence, contralateral breast cancer, and breast cancer mortality (137). It is prescribed as a pill to be taken daily for usually 5 years.

During the time at which the study subjects within this thesis were treated, tamoxifen was the most commonly prescribed type of endocrine therapy. Other forms of endocrine therapy were megestron (Megace®), a substance similar to progesterone, and goserelin (Zoladex®), an analogue to GnRH which results in medical castration. During the 1990's, aromatase-inhibitors were introduced as an alternative treatment of hormone-positive postmenopausal breast cancer, but it was not until the late 1990's that randomized controlled trials were conducted to compare the effect of aromatase inhibitors to tamoxifen. A meta-analyses has shown that the former decreases risk of recurrence compared to latter (138). However, no statistically significant improvement in survival has been shown (137).

2.2.7.3.2 Chemotherapy

There is no specific predictive factor of response to chemotherapy but according to the St Gallen treatment recommendations (117) low expression of the estrogen receptor, HER2 amplification, and increased proliferation rates predict response to chemotherapy in general. All trials that have evaluated treatment with trastuzumab have also given HER2-positive patients chemotherapy, why there is no support of the use of trastuzumab without chemotherapy (117). Furthermore, triple-negative patients should (almost always) receive chemotherapy. Whether to give chemotherapy to women with ER-positive, HER2-negative disease is more problematic; for these patients relative indications for chemotherapy include grade 3 tumors, high proliferation rate, ≥4 positive lymph nodes, tumor size of >5 cm, extensive peritumoral vascular invasion, and patient preference (117).

Chemotherapy is usually given after surgery, but women with locally advanced/primarily inoperable tumors (T3-T4), fixed axillary metastatic lymph nodes, or parasternal or supraclavicular metastasis are recommended preoperative chemotherapy (neo-adjuvant treatment) (7). Neoadjuvant chemotherapy has the potential to decrease tumor size, increasing the possibility of performing breast-conserving surgery. Although there is no difference in survival depending on whether chemotherapy is given before or after surgery, there is an increased risk of locoregional recurrence if given before (139).

During the mid 1990's (and already in the 1980's) a combination of cyclophosphamide, methotrexate, and fluorouracil (CMF) and/or anthracyclins were standardly given, especially to premenopausal women with cancers of more advanced stages. Currently in Sweden, the decision to use chemotherapy is based on lymph node status. An anthracycline-based combination of drugs or CMF is given to lymph node negative patients requiring chemotherapy, and lymph node positive patients are recommended a more aggressive treatment with anthracycline and taxane-based polychemotherapy (7).

2.2.7.3.3 Antibodies

Trastuzumab is a monoclonal antibody that binds to the HER2-receptor and impairs HER2 signalling, thereby reducing proliferation. It is given to patients with tumors showing amplification of the *HER2* gene and was initially introduced in Sweden in 2001 as a treatment of metastasized breast cancer but is now also approved for treatment non-metastasized breast cancer (7). Women with HER2-amplified tumors should always initially be treated with chemotherapy followed by trastuzumab (7).

2.2.8 Prognosis

Breast cancer survival has improved drastically over the years due to many factors such as adjuvant therapy, mammography screening, and an improved organization. The prognosis is generally good and many women are cured by surgery alone. Yet approximately 1500 women in Sweden die from breast cancer every year. Unlike most other cancers where the 5-year survival often reflects the cure rate, breast cancer has no cure rate since women with breast cancer continue to die many years after diagnosis. This is naturally also reflected in the development of recurrence; most recurrences occur within five years, but a substantial proportion relapses first after five years and many years later (140).

The currently used prognostic factors in Sweden are age, stage (tumor size, lymph node involvement and distant metastasis), histologic grade, ER-status, PR-status, HER-2-status, and Ki67 (7). However, studies have also showed that mode of detection influences survival – screening-detected cancers have been found to have a better prognosis beyond the impact of stage shift compared to cancers detected otherwise (141, 142).

Other patient characteristics than age may also influence breast cancer prognosis. High BMI, for example, has been shown to confer a poorer prognosis (143) while physical activity after diagnosis seems to improve prognosis (144). Ethnicity also affects breast cancer outcome: African American women have a poorer survival than non-African American women (145). The former are at higher risk of developing triple-negative tumors (146) and thereby have a worse prognosis, but Sparano et al. (145) also found a

poorer prognosis for African American women with ER-positive cancers and that this was independent of e.g. disparities in healthcare. Lastly, time since childbirth is inversely associated with survival (147, 148).

HRT has been shown to influence mortality, however, results are contradictory; the Women's Health Initiative (a randomized controlled trial) showed an increased mortality for HRT users, whereas observational studies have shown a decreased mortality for HRT users (68, 149). This may be due to the fact that HRT users more frequently attend mammography than non-HRT users. This will advance the date of diagnosis for HRT users, improving survival in observational studies. In the WHI, however, mammography attendance was balanced in the comparison arms and could therefore not influence results (149).

2.3 MAMMOGRAPHIC DENSITY

2.3.1 Relationship with breast cancer

It was John Wolfe who in 1976 initially proposed that the mammographic appearance of the breast was related to breast cancer risk (150, 151). He created the following categorization according to risk: N1 (predominantly fatty breasts); P1 (ductal prominence in <25% of the breast); P2 (ductal prominence in ≥25% of the breast); and DY (extensive dysplasia). Wolfe showed that the women categorized as DY had a 37-fold increased incidence of breast cancer compared to the women in the N1 group. There was, however, much critique, since Wolfe had performed a cohort study with a relatively short follow-up period. Hence, it was believed that the association was due to women with dense breasts having masked tumors at study entrance which were then diagnosed during the study period. These objections were adequate, but both cohort studies with as much as 10 years of follow-up (in which the effect of masking will diminish), and case-control studies (in which breast cancers in dense breasts instead will be underdiagnosed due to masking) have consistently showed a strong relationship between mammographic density and breast cancer (51), although not as high as Wolfe's original estimates.

To achieve a more objective method of assessing density, a semi-automated software, Cumulus (152), was developed, which calculates the total breast area, dense and non-dense areas in mm^2, and percentage density (PD) by dividing the dense area with the total breast area. PD is the most commonly used measurement of density and a meta-analyses has shown that women with a PD≥75% have a 5-fold increased risk of breast cancer than women with PD<5% (51). This makes it the strongest risk factor for breast cancer after gender, age and BRCA mutations. However, there is an important distinction between high-penetrance genes and mammographic density, because whereas the former are fairly uncommon among women with breast cancer (approximately 5% of all cases), mammographic density of >50% may account for almost 1/3 of all breast cancer cases (153, 154).

2.3.2 Mammographic density assessment

Mammographic density (MD) reflects the different tissues of the breast. Fibroglandular tissue consists of stromal and epithelial cells and is radiodense, appearing white on a mammogram, whereas fatty tissue is radiolucent and appears black.

There are two different ways of assessing density; either qualitatively or quantitatively. Wolfe's parenchymal patterns (151), described previously, and Tabár patterns (155) are examples of qualitative assessment. They take into account certain features of the breast on a mammogram, such as the appearance of the ducts and texture of the tissue, as well as the quantity of density. The American College of Radiology created a qualitative, categorical assessment for clinical purposes, BI-RADS (Breast Imaging-Reporting and Data System) which has largely replaced both the Wolfe and Tabár patterns. BI-RADS is categorized accordingly: (1) almost entirely fatty, (2) scattered fibroglandular densities, (3) heterogeneously dense, and (4) extremely dense. However reliability between readers is modest (kappa statistic=0.56) (156).

The software Cumulus is an example of a computer-assisted technique allowing for a continuous, quantitative classification of density. In short, two thresholds are set; one threshold demarcates the breast from the background (i.e. pixels with a gray value equal to or greater than this threshold comprise the total area of the breast) and the other threshold separates non-dense from dense tissue (i.e. pixels with a grey value equal to or greater than the selected threshold comprise the dense area of the breast) (Figure 7). The total breast area, absolute dense area, absolute non-dense area, and percentage density (PD) is then calculated by the software. Since Cumulus is a technique used for quantitative measurement, the qualitative aspect is lost. Nonetheless, PD is still more strongly associated with breast cancer risk than e.g. Wolfe patterns (157).

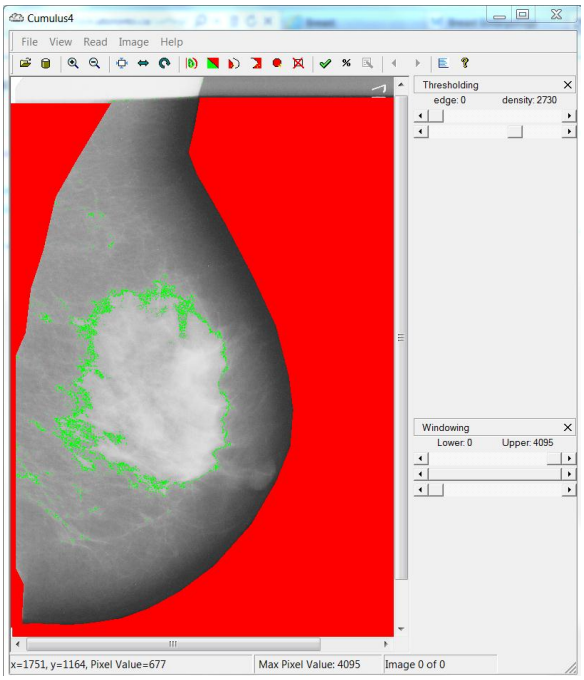


Figure 7. Assessing mammographic density using Cumulus. The pectoralis muscle and thoracic wall is demarcated as is the red background. The dense area of the breast is encircled/marked. All other tissue in the breast that is not encircled is considered non-dense.

There are important limitations with Cumulus. First, it requires digitized, analogue films, and is not yet compatible with digital images. Since many mammography units

are converting to digital images, this will become an increasing problem for future studies. The digitization of images also adds to the work load associated with density measurement, which in itself is a laborious task. Secondly, although subjectivity is decreased compared to the qualitative techniques, it is not eliminated. For both these reasons, fully automated methods are being developed (158, 159). Lastly, density, which in reality is a volume, is assessed from 2D films and is estimated as an area. New methods are being developed that measure volumetric density. Surprisingly, though, the volumetric methods do not give better risk estimations than the area-based measures (87).

Our group has developed an automated thresholding method using an established Java-based image analyses program, ImageJ, for obtaining quantitative measurements of mammographic density. We found that it compares well with Cumulus in predicting breast cancer risk(159). We also found evidence that ImageJ captured additional features of mammograms which were significant and independent markers of breast cancer risk (159). We believe that the latter may be a measurement of tissue texture. This study was, however, performed on digitized images and the MLO view, and must be validated on digital images as well as other views.

2.3.3 Percent versus absolute density

Most studies assessing density in quantitative terms use the relative measure of density, percentage density (PD). However, this does not convey any information about the absolute amounts of dense and non-dense tissues. A woman with x amount of dense tissue in a small breast will have a higher PD than a woman with the same amount of dense tissue in a larger breast (Figure 8). Moreover, two women with the same PD may have large differences in the absolute amounts of each respective tissue (Figure 8). On the other hand, PD contains information on the non-dense area which the measure of absolute density lacks unless one adjusts for the non-dense area. The non-dense area may be important, especially in postmenopausal women, since the fatty tissue is an important source of local estrogens (160), which influences the dense tissue. Furthermore, PD is more intuitive than absolute density and, used as a categorical variable, is fairly simple to apply clinically.

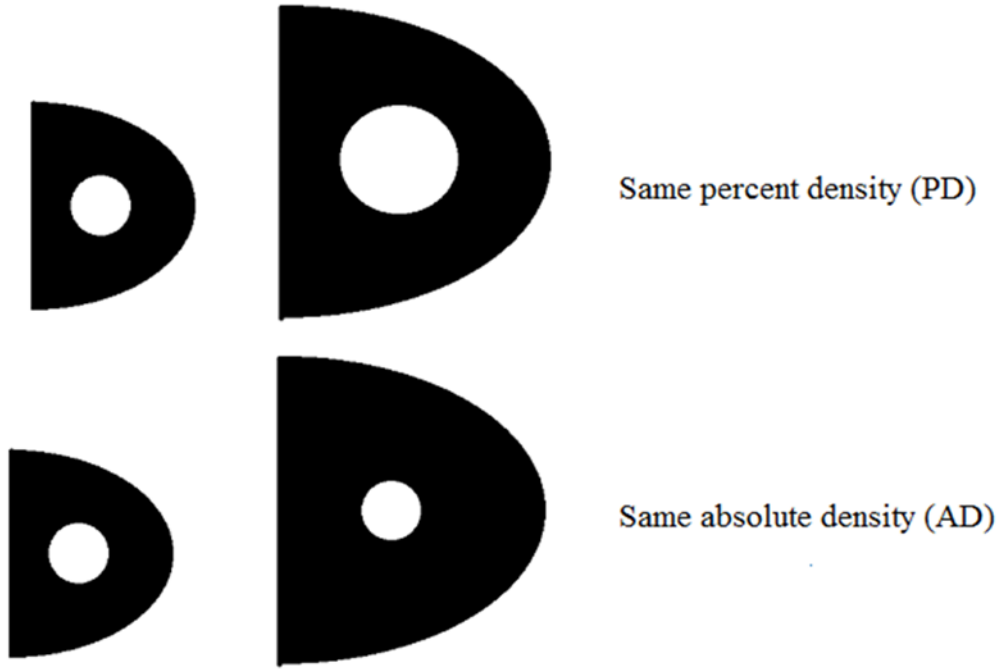


Figure 8. Percent and absolute density

However, since the amount of dense tissue reflects the amount of target tissue for carcinogenesis, it has been proposed that inferences on the etiology of mammographic density should be made using the absolute measure of density (161). Conversely, PD is satisfactory when studying breast cancer risk since both absolute density and percent density have been shown to be equally predictive of breast cancer risk (162, 163). There is no consensus on which parameter to use when studying mammographic density and breast cancer phenotypes/prognosis.

2.3.4 Histology

Through histological studies increased MD has been shown to be associated with an increased amount of stromal cells (164, 165), an increased amount of collagen (164-166), stromal fibrosis (166, 167), and an increased number of epithelial cells (164, 165, 168). Interactions between stroma and epithelium are important both in breast organogenesis and carcinogenesis and the stromal matrix effects proliferation, differentiation, migration, and apoptosis of epithelial cells (37, 39, 169).

Mammographic density has further been hypothesized to correlate with lobular involution. Only one group has investigated this relationship (170, 171). They found that there was an inverse correlation between mammographic density (assessed using the Wolfe patterns) and lobular involution (170), although breasts showing complete involution were identified both in the group with mammographically fatty breasts (N1) and in the group with densest breasts (DY). In a related study of the same cohort of women, they further showed that mammographic density, assessed using the Wolfe patterns, and lobular involution were independent risk factors of breast cancer (171); the women with no lobular involution and densest breasts having highest risks.

2.3.5 Factors influencing mammographic density

2.3.5.1 Non-genetic factors

Mammographic density is a dynamic and hormonally responsive trait. It decreases with age and during menopause which has been ascribed the process of involution (see above). The few existing longitudinal studies have shown a decrease of PD per year of ~1% (172, 173) and an especially large decrease is seen during menopause when PD may decrease by ~3% (173). Parity and increased number of live births decreases density, whereas later age at first birth is associated with higher mammographic density (174).

Hormone replacement therapy (HRT), especially combined estrogen-progestin treatment, increases MD (175). The increase in MD occurs within the first year of use (176) and the effect may disappear within three weeks after cessation (177). Lastly, the selective estrogen receptor modulator (SERM), tamoxifen, reduces MD (178). Not only has tamoxifen been seen to decrease mammographic density, but women whose density decreased also had a reduction in breast cancer risk (178). Whether this is true for other SERMs, aromatase inhibitors, and chemotherapy, remains to be confirmed (87). Although there are many examples of mammographic density as a hormonally responsive trait, most studies of mammographic density and blood hormone levels of estrogens, progesterone, SHBG, testosterone, and androstenedione have found no associations (179). Furthermore, Tamimi et al. found that mammographic density and circulating sex hormones were independent risk factors of breast cancer in postmenopausal women (180).

BMI is highly, inversely correlated with PD (181, 182). This is due to BMI being positively associated with the area of the non-dense tissue, because of an increased amount of fatty tissue in the breast. Postmenopausal women with more non-dense tissue will have an increased local production of estrogens compared to women with less non-dense tissue, which could influence the dense area. In spite of this, BMI seems to be only weakly (183), or possibly not at all associated with the absolute dense area of the breast (161).

Thus, most factors that influence mammographic density also influence breast cancer risk in the same direction, the two exceptions being age and BMI. Boyd et al. (184) related mammographic density to the Pike model (50) to try to explain the counterintuitive association between mammographic density and age. Pike et al. modelled breast cancer incidence as a function of breast tissue aging instead of chronological age (50). They proposed that breast tissue aging starts at menarche at a certain constant rate until first full-term pregnancy. It then slows and continues at that rate until menopause, when it slows even more, and after which it is constant. Pike et al. (50) further showed that it was the cumulative exposure to breast tissue aging that described the age-incidence curve for breast cancer. Rosner and Colditz extended the Pike model to include more breast cancer risk factors (185-187) and proposed that it was not breast tissue aging per se that was the underlying cause of breast cancer incidence, rather breast tissue aging was a proxy for cell proliferation and accumulation of genetic damage due to cumulative hormonal exposure. Boyd et al. applied this to mammographic density and hypothesized that cumulative exposure to mammographic density represents the epithelial and stromal cells' cumulative exposure to hormonal

and growth factor stimuli, and that the cumulative exposure to density increases with age (157).

2.3.5.2 Genetic factors

Mammographic density is highly heritable at a given age (188); lifestyle factors only explaining 20-30% of the variation in PD between women (189), and >60% of the variation attributable to genetic factors (188). As with other quantitative traits, there is evidence that the mode of inheritance is polygenic (190), i.e. that several genes are involved in the heritability of the disease.

Since mammographic density has been considered an intermediate in breast cancer development (184), investigations of associations between genetic variations and mammographic density are important to attempt to shed light on breast cancer biology and also to possibly find ways of disease prevention. To date, the following breast cancer susceptibility genes, *ZNF365* (191), *LSP1* (192, 193), *RAD51L1* (192), *ESR1* (191), and a locus on the long arm of chromosome 8 (193) have been found to correlate with mammographic density.

3 AIMS OF THIS THESIS

The overall aim of this thesis has been to investigate whether mammographic density, one of the strongest risk factors of breast cancer, also influences the breast cancer phenotype and breast cancer progression. We attempted to answer this question through the four studies carried out in this thesis. They had the following individual objectives:

Study I: To investigate if mammographic density is associated with breast tumor characteristics.

Study II: To investigate if mammographic density is associated with molecular subtypes of breast cancer.

Study III: To investigate if mammographic density influences breast cancer recurrence and survival.

Study IV: To investigate if mammographic density influences survival in interval cancers.

4 MATERIALS AND METHODS

4.1 REGISTRIES

The personal identity number was introduced in 1947. It is a unique 10-digit identifier consisting of six digits denoting the birth date (year, month, and day), two digits originally specifying the place of birth, but this has been changed so that these two digits no longer have a relation to place of birth, one digit to identify the sex of the carrier, and, lastly, a control digit. The control digit is calculated given the first nine digits so that the entrance of an incorrect number into e.g. a medical record data base will give an error message. The personal identity number is registered in all public registers, allowing for easy and unequivocal withdrawal and cross-linkage of information.

4.2 THE SWEDISH POPULATION REGISTER

The Swedish Population Register includes information on population statistics such as name, personal identity number, place of birth, civil status, address, and immigration/emigration data. Population registration is an old tradition originally maintained by the church; the oldest preserved records dating back to the early 17th century.

4.3 THE SWEDISH CANCER REGISTRY

The Swedish Cancer Registry was established in 1958 and holds information on all primary incident cancers diagnosed in Sweden (49). Reporting is mandatory for physicians as for pathologists/cytologists separately, which has led to a double notification system, and a high degree of completeness. The system also ensures that cancers diagnosed at e.g. autopsy are included and are classified according to the International Classification of Disease (ICD). 99% of the registered cancers have been morphologically verified, and in women, completeness may be as high as 99% (194).

4.4 STOCKHOLM-GOTLAND BREAST CANCER REGISTRY

All new primary breast cancers diagnosed in the Stockholm-Gotland health care region since 1976 have been reported to the Stockholm-Gotland Breast Cancer registry and it has a close to complete coverage. An important role of the registry is to evaluate the quality of breast cancer care and ensure that patients receive the same care independent of place of diagnosis.

4.5 THE CAUSE OF DEATH REGISTRY

The cause of death register covers all residents in Sweden. It was computerized in 1952 and is considered reliable since 1961. The reporting is mandatory and there are essentially no missing deaths. It has also been shown to correctly classify 98% of breast cancer deaths (195). Information in the register includes date of death, underlying cause of death, and up to 10 contributing causes, classified according to ICD.

4.6 STUDY POPULATIONS

4.6.1 Paper I, III, and IV

This study is an extension of a large case-control study, CAHRES, among all Swedish residents born in Sweden and aged 50 to 74 years at the time of enrollment, 1 October, 1993 - 31 March, 1995. Women with incident primary invasive breast cancer were identified via the six Swedish Regional Cancer Registries. The study identified 3,979 women, of whom 84% (n=3345) participated. However, of the cases included, 19 were diagnosed outside of the study period, one case had a diagnosis other than breast cancer, and 58 cases had non-invasive breast cancer, rendering them ineligible.

For the studies within this thesis, the inclusion criteria were further refined to only include postmenopausal women who had no prior diagnosis of cancer other than non-melanoma skin cancer. Menopause was defined as the age at the last menstrual period or the age at bilateral oophorectomy if at least one year prior to date of study entrance. 198 premenopausal women and 202 women with unknown menopausal status who were younger than 55 for non-smokers or 54 for smokers (the 90th percentile of age at natural menopause of study subjects) were thus excluded from the study as were 147 women with previous cancer. The study base thus consisted of 2720 breast cancer cases.

We used the national registration number to retrieve the correct patient records, and collected, among other things, information on tumor characteristics and reason for diagnostic mammography (see 4.7.1.3 Data collection of clinical variables).

Using the Swedish national registration numbers, we obtained addresses for participants from 1975 to 1995 through the civil registry. During 2007 and 2008, we visited all mammography screening units and radiology departments conducting screening mammography throughout Sweden. A second retrieval attempt was carried out in 2010-2011. For the eligible participants in these studies, we managed to collect mammograms for 2046 women (75%). For all three studies, we used the mammogram closest to diagnosis, excluding post-diagnostic mammograms. 107 women who only had post-diagnostic mammograms available were thus excluded. The median difference from date of mammography to study entrance was 50 days.

Since studies have shown that MD may differ histologically in pre- and postmenopausal women (164) and also may be affected differentially by hormones (196, 197), we excluded women who lacked postmenopausal mammograms (n=79).

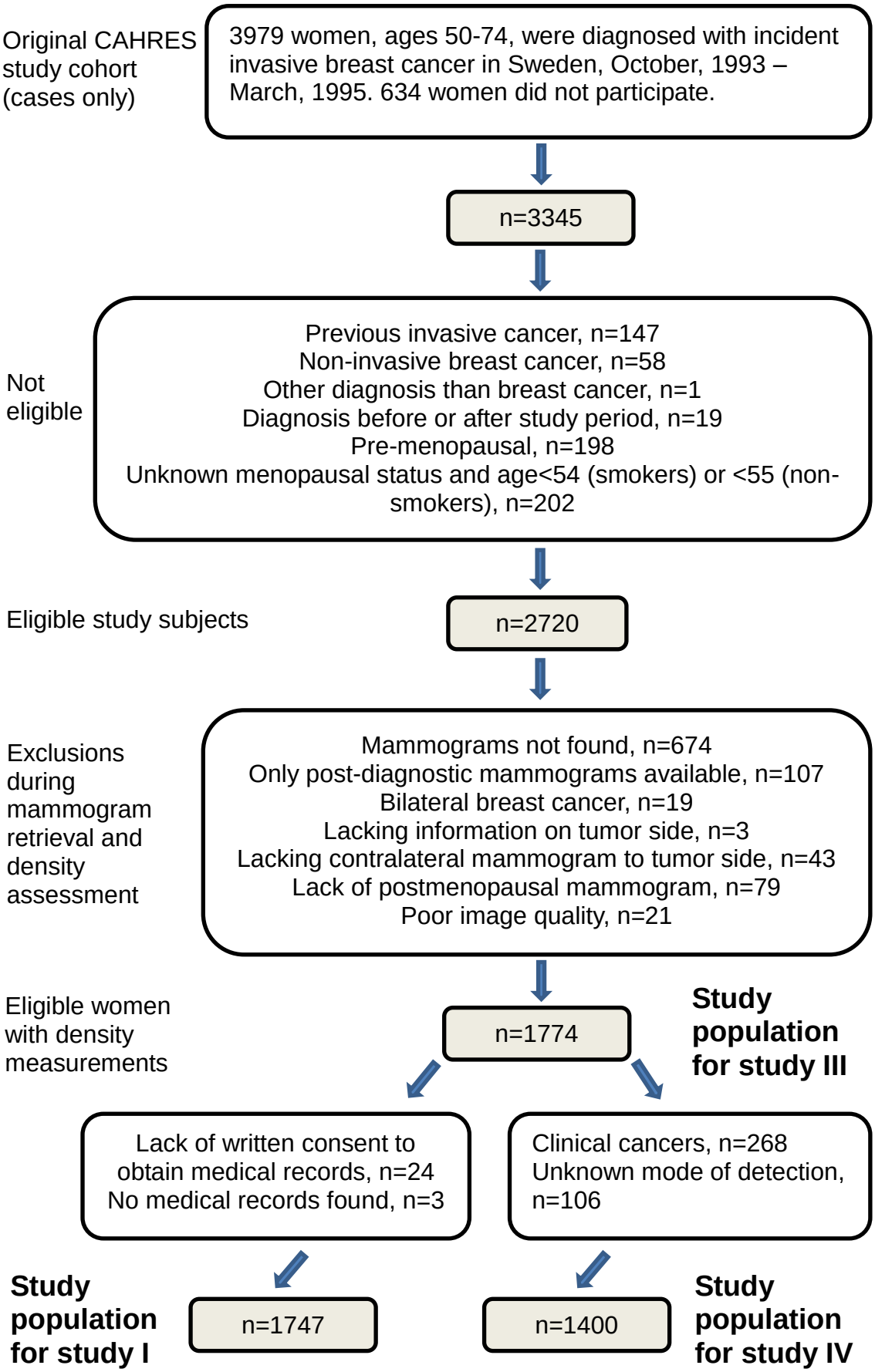
Tumors appear white on a mammogram and can thereby distort density measurements. Hence, we used the mammogram of the breast contralateral to the tumor, excluding women with missing information on tumor side (n=3) or lacking contralateral mammograms (n=62, of which 19 had bilateral breast cancer). Images of poor quality, including breasts with silicone implants, were also omitted, excluding 21 women. 1774 women thus comprised the study population of study III.

Following a decision of the Ethical Review Board of the University of Lund, written informed consent was sought to retrieve information from medical records. For study I, 24 women were excluded due to lack of written consent and another 3 women were excluded due to missing medical records. The final study population for this study thus included 1747 women.

For the fourth study, only women with screening-detected cancers and interval cancers were included. We thus used information collected from 66 of the 68 mammography units in Sweden from which we had gathered the following: Date and reason for the mammographies (screening or referral and reason for referral,) performed within 5 years before diagnosis, excluding 3 months just before diagnosis to avoid registering diagnostic examinations. We added this to the information on mode of detection (collected from the medical journals,) to assess interval cancer status for which the following definition was applied: Non-screening detected cancers (i.e. reason for diagnostic mammography was not screening), where a previous, negative (either negative directly at mammography or at follow-up) screening mammography, had been conducted 3-24 months before the diagnostic mammography. 24 months was the screening-interval in Sweden during the time period of the study. The three-month cut-off was used to avoid including mammography screenings with clinical work-up periods. Out of the 1774 eligible women, 268 women had non-screening detected cancers but either a previous screening mammography more than two years prior to diagnosis, or no previous screening mammography at all, and 106 individuals lacked information to be able to assess interval cancer status. All of these women were therefore excluded. Hence, there were 1115 women with screening-detected cancers and 285 women with interval cancers which comprised the study population.

Figure 9 depicts the process of study population selection and figure 10 illustrates the process of interval cancer assessment.

Figure 9. Flow chart of study populations



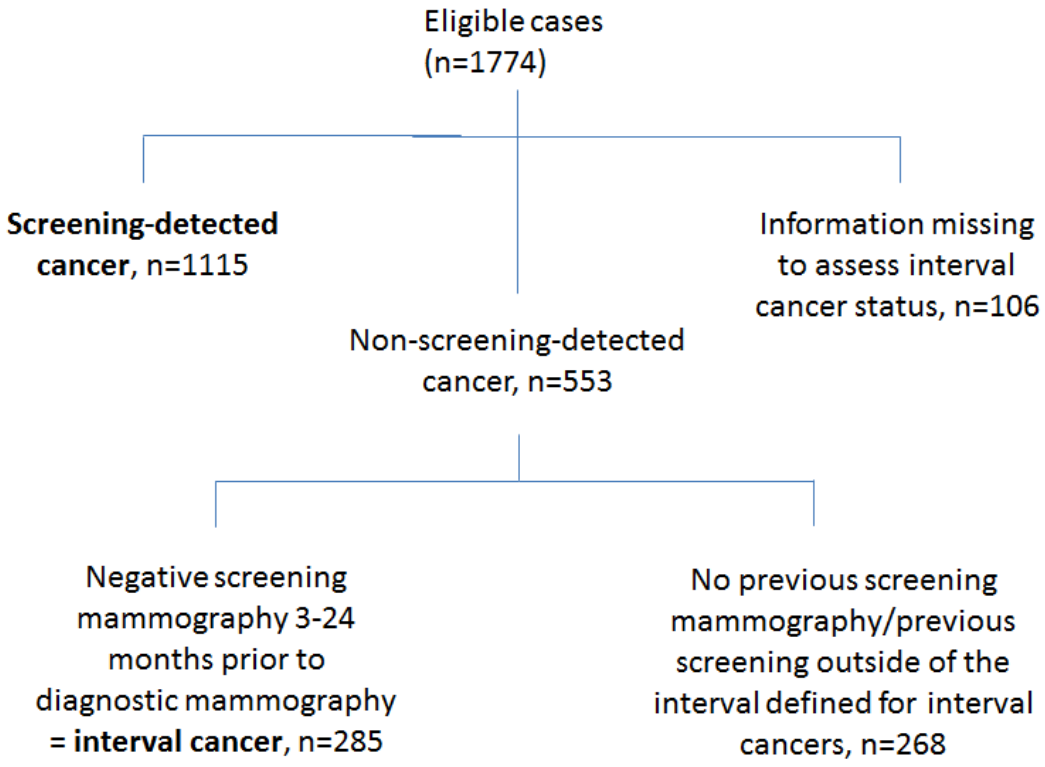


Figure 10. Flow chart of study population for study IV

4.6.2 Study II

Study II is also a case-only study. The source population was all women with breast cancer operated at a large university hospital in Stockholm (Karolinska University Hospital) between 1 January 1994 and 31 December 1996 (n=524), as previously described (198). The women were identified through the population-based Stockholm-Gotland Breast Cancer Registry. Exclusion was due to refusal of participation (n=6), emigration (n=7), lack of frozen tumor (n=231), insufficient amount or quality of RNA (n=89), lack of gene expression profiling on U133 A and B chips (n=14), neoadjuvant therapy (n=12), in situ cancer (n=5), or stage IV cancer (n=1).

The mammogram closest to diagnosis was retrieved for 141 out of the 159 subjects with gene expression profiling information. Women with bilateral breast cancer (n=10) and subjects with breast implants (n=3) were excluded. We thus had density measurements for 128 women. As in the original publication (198), it was not possible to assign a unique subtype for all samples. Consequently, n=18 patients were excluded, leaving n=110 patients for the final analysis.

4.7 DATA COLLECTION AND CLASSIFICATIONS

4.7.1 Studies I, III, and IV

4.7.1.1 Questionnaire data

Data on sociodemographic, anthropometric, hormonal and lifestyle factors, as well as family history of breast and ovarian cancer, previous benign breast disease, and previous mammography examinations were collected by means of a postal questionnaire. The questionnaire included detailed information on HRT including brand, dosage, and dates of first and last use of each treatment episode. Pictures of all brands marketed in Sweden 1950-1995 were also included to aid recall. If essential information was missing in the questionnaire, then cases (and controls) were contacted by telephone to obtain this. Approximately 50% of cases were contacted in this way.

Since date of mammography was prior to study entrance, the variables age, menopausal status, and HRT use were reassessed according to date of mammography. We were not able to do this with BMI as we only had information on BMI at study entrance and one year prior to study entrance. However, it has previously been shown that inter-individual variations in BMI are small (199) and the difference in BMI at study entrance and one year prior to this was 0.05 units (SD 1.2) for our study participants.

HRT was classified according to recency (current, former, and never use). Since the influence of HRT on MD may diminish within three weeks of cessation (176), former users were those who discontinued HRT-use more than one month prior to date of mammography. There were no individuals classified as never users who started using HRT after date of mammography. All compounds, modes of administration, and potencies were included in the HRT variable except for low potency, estrogen-only pharmaceuticals, since the latter have not been shown to increase breast cancer risk (200).

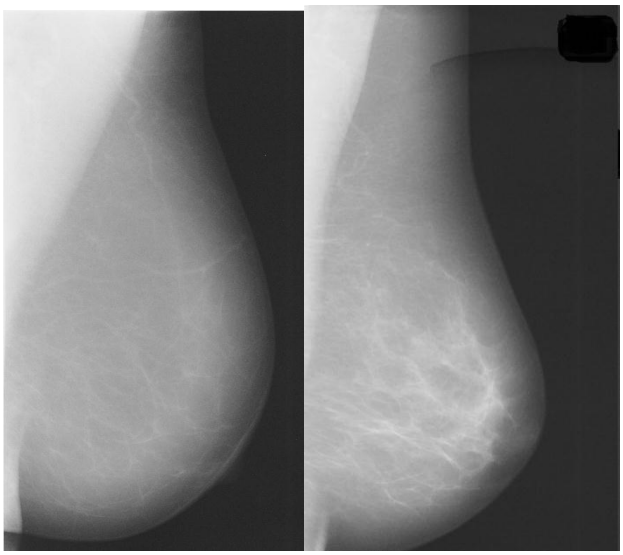
4.7.1.2 Mammographic density data

Film mammograms of the MLO view were digitized using an Array 2905HD Laser Film Digitizer (Array Corporation, Tokyo, Japan), which covers a range of 0 to 4.7 optical density. The MLO view was used since this was the routine view used at mammography screening in Sweden. The density resolution was set at 12-bit spatial resolution. We used Cumulus to assess density (152) of the mammogram contralateral to the tumor. For each image, a trained observer (LE) set the appropriate gray-scale threshold levels defining the edge of the breast and distinguishing dense from non-dense tissue. The software calculated the total number of pixels within the entire region of interest and within the region identified as dense. The percentage density was then calculated from these values (dense area/total breast area). The images were measured together with approximately the same amount of images for healthy women, and the reader was blinded to case-control status and also, naturally, to all variables associated with tumor phenotype/prognosis. A random 10% of the images were included as replicates to assess the intra-observer reliability, which was high with an R^2 of 0.92.

4.7.1.2.1 Mammographic density

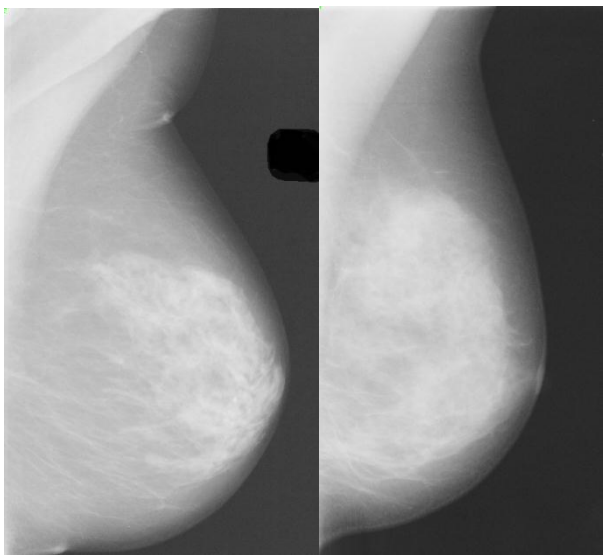
In study I, III, and IV, mammographic density was treated as a dichotomized variable with the cutoff of 25% for high mammographic density, defining the highest quartile in our cohort. (Figure 11 depicts mammograms with different PD values also categorized based on the 25% cut-off.) We used a categorical variable since this is more clinically relevant than a continuous measure and more easily interpreted. An alternative to our categorization would be to use the categories introduced by Boyd et al. (157). However, these did not fit our postmenopausal study population whose density is lower and more homogeneous than a combined pre- and postmenopausal population. N.B. that all analyses in study I and III were also performed using PD as a continuous variable.

Figure 11. Mammograms depicting breasts with different PD values, also categorized according to the 25% cut-off used in studies I, III, and IV.



PD<25%

a) PD=0% b) PD=10%



PD≥25%

c) PD=25% d) PD=50%

4.7.1.3 Data collection of clinical variables

We used the national registration numbers to retrieve patient records and register information. Between 2000 and 2002, we collected information on primary surgery, adjuvant treatment (endocrine therapy, chemotherapy, and radiotherapy), tumor characteristics, possible recurrence, and reason for diagnostic mammography from surgical and oncological patient records throughout Sweden.

Grade was classified according to the Nottingham histologic grade or the Bloom–Richardson scale into three groups. Tumors were considered ER-positive or progesterone receptor (PR)-positive if they contained ≥ 0.05 fmol receptor/μg DNA or ≥ 10 fmol receptor/mg protein.

Type of surgery was classified as partial or total mastectomy independent of axillary surgery since 98% of our study participants had axillary lymph node clearance carried out.

Mode of detection was treated as a binary variable – screening- vs. non-screening-detected cancer in studies I and III. This was due to few observations in the categories of asymptomatic women referred for mammographic examination, and they were thus combined with the group of women with symptomatic cancers.

For information on recurrence variables, see *4.7.1.5 Data on survival variables*.

4.7.1.4 Data on interval cancer status

We visited 66 of the 68 units performing mammographic examinations in Sweden, and collected information on the dates and reasons for the examinations (screening or referral and reason for referral) performed within 5 years before diagnosis, excluding 3 months just before diagnosis to avoid registering diagnostic examinations. To evaluate whether lack of information on pre-diagnostic mammographies was due to failure in finding the information or a true lack of previous mammographies, we compared the information retrieved from mammography units to the questionnaire. In the latter, women had stated how many mammographies they had undergone 5 years prior to diagnosis and this was thus used as reference. N.B., this was only done to distinguish between clinical cancers and cancers where we lacked information to assess interval cancer status (see Figure 10), and was carried out as quality control.

4.7.1.5 Data on survival variables

We collected information on emigrations from the Swedish National Population Register, and the date and cause of death until 31 December 2008 from the Swedish Causes of Death Registry.

Local recurrence included recurrence in the residual ipsilateral breast, scar tissue, or chest wall. Locoregional recurrence included both local recurrences and recurrences in regional lymph nodes. During the collection of follow-up information from clinical records, the study subjects' follow-up period was marked as complete or incomplete depending on whether follow-up information was complete or missing until date of retrieval. Follow-up time for recurrence variables was thus computed from date of diagnosis and, depending on outcome, as follows: If recurrence occurred, time to specific recurrence; if the specific recurrence did not occur and follow-up information

from medical records was considered complete, time to date of medical record retrieval or date of death, whichever came first; or, lastly, if information on follow-up was considered incomplete, time to last documented date in the medical record or date of death, whichever came first.

4.7.2 Study II

4.7.2.1 Data on patient and tumor characteristics

We collected information pertaining to status at diagnosis on age, menopausal status, HRT, family history, oral contraceptive use, and tumor characteristics from the medical records. Breast cancer heredity is defined as history of breast cancer in a first-degree relative. Menopausal status was self-assessed by the patient as either pre- or postmenopausal. Two women had unknown menopausal status. Both oral contraceptive use and HRT use were assessed according to status at time of referral to the Karolinska Hospital (former, current, and nonuse, collapsing former and current use into one category because of few observations). Non-users of HRT were postmenopausal women actively stating no current or previous use of HRT. Of the HRT users, approximately two out of three used a combined estrogen and progesterone regimen, and one out of three used estrogen only. Local estrogen treatment was not considered as HRT use. Oral contraceptive use included all preparations.

4.7.2.2 Gene expression data

Details on RNA preparation and microarray profiling have previously been described elsewhere (198). Briefly, frozen tumor was cut into minute pieces and transferred into test tubes with RLT buffer (RNeasy lysis buffer, Qiagen, Hilden, Germany), followed by homogenization. Proteinase K was then added. After this step, total RNA was isolated using Qiagen's microspin technology. DNase was added to some samples to further increase RNA quality. The quality of RNA was assessed by measuring the 28S:18S ribosomal RNA ratio. Preparation of in vitro transcription products and oligonucleotide array hybridization and scanning were performed according to the protocol of Affymetrix (Santa Clara, CA, USA). The molecular subtypes have been validated previously on a larger cohort of patients (201).

4.7.2.3 Mammographic density data

The process of digitization of images was the same as for studies I, III, and IV (see above).

Two independent observers (ISS and VM) carried out the density measurements blinded to the characteristics of the patients and their tumors. Both observers measured all of the images and a random repeat sample of 10% of the images. There was good inter- and intra-observer reliability with Pearson's correlation coefficients of 0.82 and 0.93, respectively, for absolute density. For our analysis, the density measurements from both observers were averaged to minimize random measurement error.

4.7.3 Statistical analyses

4.7.3.1 Brief description of statistical analyses used in this thesis

4.7.3.1.1 Linear regression

A linear regression model is used to model the relationship between an outcome that is continuous and a linear combination of a set of explanatory/predictive variables. In the simplest case of one explanatory variable, the model is based on the following equation: $y_i = a + bx_i + \varepsilon_i$, where y_i is the outcome, x_i is the explanatory variable of the ith individual/sampling unit, a is the intercept, and b is the slope of the model. The relationship is modeled using an error variable ε_i.

The model is most commonly based on the normal distribution curve – ε_i is assumed to be normally distributed, hence, the outcome is also assumed to be normally distributed. The supposed linear relationship should be tested when fitting the model to the data by e.g. plotting a simple scatter plot of the exposure and outcome. Furthermore, the variance in errors or the outcome variable should be independent of the values of the explanatory variables (homoscedasticity), which can be assessed by plotting the residual errors against the predicted values. This linear regression model can be used to make inferences about the mean value of y (e.g. BMI) given x.

4.7.3.1.2 Logistic regression

Logistic regression is used for modeling binary outcomes (0 or 1, where 0 could be the state *healthy* and 1 the state *diseased*). It can be used to estimate the odds ratio, which compares the odds of e.g. developing breast cancer given the exposure compared to the odds of developing breast cancer given not having the exposure. It models the natural logarithm of the odds, which is continuous, as a function of the linear predictor. Instead of the normal distribution, the binomial or Bernouilli distribution underlies this model.

The multinomial logistic regression model is a generalization of the binomial logistic regression model, and allows for the use of categorical outcomes, such as molecular subtypes. There is no implied order for the different values of the outcome.

4.7.3.1.3 The Cox proportional hazards model

The Cox proportional hazards model is a model used for survival analyses and is used to estimate the effect of an exposure on a time-to-event variable, such as recurrence or death. A hazards ratio (HR) is calculated. This is the ratio of the hazard rates (e.g. the instantaneous rates of dying/failing) for the exposed and unexposed groups. Basic Cox proportional hazards models assume that this ratio does not change over time (proportional). However, the model assumes nothing of the underlying baseline hazard which may vary with time.

4.7.3.2 Study I

We performed regression analyses treating tumor characteristics as outcomes and density as a covariate; we used linear regression for studying tumor size (tumor size was transformed by power of 0.2 to obtain an approximately normal distribution), multinomial logistic regression for grade and histological classification, and logistic regression for all other outcomes. Tests of heterogeneity in effect sizes across

subgroups were carried out for grade and histological classification. We included age (continuous), BMI (continuous), HRT-use (categorical), age at menarche (continuous), previous oral contraceptive use (binary), parity and age at first birth combined into one categorical variable (nulliparous, parity ≤2 and age at first birth ≤25, parity ≤2 and age at first birth >25, parity >2 and age at first birth ≤25, and lastly, parity >2 and age at first birth >25), breast feeding ever (binary), age at menopause (continuous), previous benign breast disease (binary), family history of breast cancer in a first-degree relative (binary), and mode of detection (to try to account for masking) as potential confounders. We did not adjust for tumor size, presence of lymph node metastasis, or grade since this would risk depleting a true, biological effect of density on tumor characteristics or over-adjustment, since mode of detection also is associated with these factors. The decision to use covariates as continuous or categorical variables was based on the goodness of fit of the model which was ascertained using AIC (202).

4.7.3.3 Study II

Our main variable of interest pertaining to MD was the AD area (measured in cm2), owing to the lack of information on BMI.

We analyzed AD as a continuous variable after transformation to the square-root of AD to make the density distribution more symmetric. The transformed density values were then standardized by subtracting the mean and dividing by the SD (Z scores), to be able to interpret the risk estimates in terms of the inherent variability of the density values, see below. The relationship between AD and molecular subtypes was modelled via multinomial logistic regression. The multinomial model is an extension of the logistic regression model that allows for more than two categories in the response variable (i.e., subtype). The luminal A subtype was set as reference category and we report risk estimates as relative risk ratios (RRRs) for the standardized, square-root-transformed AD values. The RRRs reported in Table 3 (page 44) measure the change in odds for a tumor falling into any of the reported categories relative to the reference category that is associated with an increase of square-root transformed AD by one SD (or somewhat less than 25% of the range of densities).

We fitted both an age-adjusted model as well as a fully adjusted model, which took into account known correlates of MD and breast cancer risk (age, menopausal status, HRT, family history, and oral contraceptive use) (157, 203) and tumor size. The latter adjustment was made to try to account for the masking bias and its possible influence on molecular subtypes. We have no prior knowledge of whether the factors adjusted for influence molecular subtypes, so we chose a conservative approach. Significance testing was conducted via likelihood ratio tests. Confidence intervals and P-values for individual parameters are based on Wald statistics. All tests were conducted at a nominal significance level of a=0.05.

4.7.3.4 Study III

Survival analyses were carried out using the Cox proportional hazards model to study the association between PD and prognostic variables. We included age (continuous), BMI (continuous), and HRT-use (categorical) as potential confounders, based on their strong association with MD and breast cancer risk (203) and influences on breast cancer survival (143, 149, 204). We viewed this as our "crude model" since we were not interested in their effect on tumor characteristics but rather the independent effect of PD. In the full model we further adjusted for tumor size (continuous), lymph node

metastasis (continuous), estrogen receptor(ER)-status (categorical), progesterone receptor(PR)-status (categorical), grade (categorical), and mode of detection (binary; screening- vs. non-screening-detected cancer). These factors were adjusted for since they all influence prognosis (111, 141). Whereas less than 5% of our study population were missing information on mode of detection, tumor size, and presence of lymph node metastasis, respectively, approximately 30% were missing information on grade, ER-, and PR-status. We thus added a missing category to these variables.

4.7.3.5 Study IV

We studied the influence of PD on survival in interval cancers versus screening-detected cancers by first comparing survival in interval cancers compared to screening-detected cancers independent of density, and then stratifying on density (<25% vs. ≥25%). The Cox proportional hazards model was used to calculate hazard ratios (HR) and their associated 95% confidence intervals (95% CI) for 5-year breast cancer-specific survival. Age at mammography (continuous), BMI (continuous), and HRT use (categorical) were included as possible confounders in the first model, further adjusting for tumor size (continuous) to try to account for time to diagnosis, or lead-time, in the second model.

5 MAIN RESULTS

5.1 STUDY I

Table 2 shows the results from the analyses of PD and tumor characteristics. PD was positively associated with tumor size (regression coefficient 0.031 for tumor size, p = 0.017). When we excluded mode of detection from the model, the regression coefficient increased slightly and became highly significant (regression coefficient 0.043, p = 0.001) (results not shown in Table 1). There was a borderline statistically significant association between PD and grade 3 tumors (OR 1.56 for grade 3, p = 0.069). However, a test of heterogeneity revealed that there was no difference in risk associated with PD between different categories of grade (p = 0.192).

Re-analyses using PD as a continuous variable gave similar results as those presented, i.e., PD was only associated with tumor size (data not shown).

Table 2: Association between percentage density (PD) (≥25% vs. <25%) and tumor characteristics*

Outcome variable	**Effect size**[a]	**95% CI**	**p-value**
	(Linear) regression coefficient		
Tumor size			
mm^0.2	0.031	0.005 0.056	0.017
	Odds ratio (OR)		
ER-status			
ER-negative	1.00 (Ref.)		
ER-positive	1.07	0.72 1.60	0.725
PR-status			
PR-negative	1.00 (Ref.)		
PR-positive	1.17	0.83 1.66	0.369
ERPR-status			
ERPR-negative	1.00 (Ref.)		
ERPR-positive	1.31	0.83 2.06	0.247
Lymph node metastasis			
Negative	1.00 (Ref.)		
Positive	0.88	0.65 1.19	0.394
Grade (WHO)			
1	1.00 (Ref.)		
2	1.36	0.85 2.19	0.205[b]
3	1.56	0.97 2.52	0.069[b]
Histology			
Ductal	1.00 (Ref.)		
Lobular	1.22	0.82 1.82	0.335
Other	1.08	0.74 1.57	0.689

[a] Odds ratios in all cases except for tumor size for which a (linear) regression coefficient is presented. PD coded as 0= <25%, 1= ≥25%.
[b] p=0.192 for heterogeneity across subgroups

5.2 STUDY II

Compared with the luminal A subtype (taken as the reference category), the relative risk of the luminal B, ERBB2, and normal breast-like subtypes increased with increasing AD both in the age adjusted (RRR 1.19, 95% CI 0.58–2.45; RRR 1.88, 95% CI 0.79–4.48; and RRR 1.51, 95% CI 0.78–2.92, respectively, for an increase in square-root-transformed density by one SD) and in the fully adjusted (RRR 1.22, 95% CI 0.53–2.83; RRR 1.74, 95% CI 0.62–4.85; and RRR 1.43, 95% CI 0.64–3.17, respectively) models (Table 3). The relative risk of the basal subtype was essentially the same as that of the luminal A subtype in the age-adjusted model (RRR 0.99, 95% CI 0.48–2.06), but decreased with increasing AD in the fully adjusted model (RRR 0.83, 95% CI 0.33–2.10) (Table 3). None of the individual associations were, however, statistically significant, nor was the association between AD and molecular subtype as a whole statistically significant (p=0.483 and p=0.651 for the age adjusted and fully adjusted models, respectively).

Table 3: Relative risk ratios (RRRs) for specific molecular subtypes of breast cancer compared to the luminal A subtype for an increase in square-root transformed AD by one standard deviation (SD).

	Age-adjusted*			**Fully adjusted****		
	RRR	**95% CI**	**p-value**	**RRR**	**95% CI**	**p-value**
Subtype						
Luminal A	1.00 (Ref.)			1.00 (Ref.)		
Luminal B	1.19	0.58-2.45	0.641	1.22	0.53-2.83	0.644
Basal-like	0.99	0.48-2.06	0.986	0.83	0.33-2.10	0.690
ERBB2	1.88	0.79-4.48	0.153	1.74	0.62-4.85	0.291
Normal breast-like	1.51	0.78-2.92	0.222	1.43	0.64-3.17	0.385

*Adjusted for age; p=0.483 for the association between AD and subtype as a whole based on the likelihood ratio test.

**Adjusted for age, oral contraceptive use, menopausal status, HRT use, family history, tumor size; p=0.651 for the association between AD and subtype as a whole based on the likelihood ratio test.

5.3 STUDY III

Results from survival analyses are reported in Table 4. PD was associated with local and locoregional recurrence both before (HR for local recurrence 1·99, 95% CI 1·09-3·66; HR for locoregional recurrence 1·84, 95% CI 1·16-2·91) and after adjustment for established prognosticators (HR for local recurrence 1·92, 95% CI 1·03-3·57; HR for locoregional recurrence 1·67, 95% CI 1·04-2·69). No associations between PD, distant metastasis, breast cancer-specific survival, and overall survival were observed.

Table 4: Association of percentage density (PD) with breast cancer recurrence and survival

	Adjusted model 1*			Adjusted model 2**		
	HR[1]	**CI (95%)**	**p-value**	**HR[1]**	**CI (95%)**	**p-value**
Local recurrence	1.99	1.09-3.66	0.026	1.92	1.03-3.57	0.039
Locoregional recurrence	1.84	1.16-2.91	0.010	1.67	1.04-2.69	0.033
Distant recurrence	1.28	0.90-1.82	0.170	1.08	0.74-1.56	0.698
5-year breast cancer-specific survival	1.25	0.82-1.92	0.299	0.97	0.61-1.54	0.908
10-year breast cancer-specific survival	1.34	0.97-1.85	0.080	1.08	0.77-1.53	0.644
Overall 5-year survival	1.07	0.73-1.55	0.742	0.89	0.60-1.32	0.554
Overall 10-year survival	1.14	0.87-1.48	0.343	0.99	0.75-1.29	0.921

*Adjusted for age, BMI, and hormone replacement therapy (HRT) use
** Adjusted for age, BMI, HRT use, mode of detection, tumor size, lymph node metastasis, ER-status, PR-status, and grade
[1] HR=Hazards ratio comparing PD≥25% to PD<25%

The associations described in Table 4 between PD and local and locoregional recurrence could be due to density masking residual disease in women who were operated with partial mastectomy. We therefore redid the analyses pertaining to PD, local, and locoregional recurrence, stratifying on surgical procedure (Table 5). For women exposed to total mastectomy, the HRs for local and locoregional recurrence were 2·93 (95% CI 1·03-8·39) and 2·16 (95% CI 1·02-4·54), respectively, after full adjustment. For partially mastectomized women, the HRs for local recurrence and locoregional recurrence were 1·48 (95% CI 0·67-3·25) and 1·57 (95% CI 0·83-2·97), respectively, after full adjustment. Although the HRs were highest and only statistically significant for the group treated with total mastectomy, there was no statistically significant difference in the effect of PD on risk of recurrence based on surgical procedure, reflected in the non-significant p-values for interaction.

Table 5: Association between percentage density (PD≥25% compared to PD<25%) and local recurrence stratified on type of surgery

	Partial mastectomy			**Total mastectomy**			**p-value for inter-action[2]**
	HR[1]	**CI**	**p-value**	**HR[1]**	**CI**	**p-value**	
Local recurrence							
Adjustment 1*	1.53	0.70-3.32	0.286	3.44	1.26-9.38	0.016	0.248
Adjustment 2**	1.48	0.67-3.25	0.334	2.93	1.03-8.39	0.045	0.314
Locoregional recurrence							
Adjustment 1*	1.45	0.78-2.70	0.242	2.69	1.34-5.43	0.006	0.238
Adjustment 2**	1.57	0.83-2.97	0.166	2.16	1.02-4.54	0.044	0.485

[1]HR=Hazards ratio

[2]p-value for interaction testing a possible difference in effects of PD on local and locoregional recurrence depending on type of surgery.

* Adjusted for age, BMI, and hormone replacement therapy (HRT)

** Adjusted for age, BMI, HRT, mode of detection, tumor size, lymph node metastasis, ER-status, PR-status, and grade

Although we used PD as a binary variable, we also carried out analyses using PD as a continuous measure. This did not change the interpretation of the results (data not shown).

5.4 STUDY IV

ICs had a particularly unfavorable phenotype compared to screening-detected cancers; they were larger (p<0.001), more often lymph node positive (p<0.001), ER-negative (p=0.020), PR-negative (0.008), of higher grade (<0.001), and of higher proliferation rate (p<0.001) (see Table 2 in manuscript IV). These results were reflected in the comparison of ICs in non-dense breasts with screening-detected cancers in non-dense breasts. Conversely, in dense breasts, ICs were similar to screening-detected cancers except that they were larger (p<0.001), more often presented with lymph node metastasis (p=0.001), and of higher grade (p=0.012). Despite the difference in tumor size, there was no statistically significant difference in proliferation rates (p=0.523).

According to the Cox proportional hazards model, ICs had a worse prognosis than screening-detected cancers, independent of adjustments (Table 6). Both types of ICs were associated with a HR of ~3 before adjustment for tumor size. After adjustment for tumor size, ICs in non-dense breasts still had a statistically significantly worse survival then screening-detected cancers in non-dense breasts (HR of 2.43, 95% CI 1.44-4.10), whereas the point estimate approached unity and was statistically non-significant in dense breasts (HR 1.41, 95% CI 0.53-3.74).

Table 6: Hazard ratios (HRs) comparing 5-year breast cancer-specific survival in 1) interval cancers (ICs) to screening-detected cancers 2) only including non-dense breasts 3) and only including dense breasts

	ICs compared to screening-detected cancers, all			**ICs compared to screening-detected cancers in non-dense breasts**			**ICs compared to screening-detected cancers in dense breasts**		
	HR	**CI (95%)**	**p-value**	**HR**	**CI (95%)**	**p-value**	**HR**	**CI (95%)**	**p-value**
Model[1]	3.50	2.25-5.44	<0.001	3.62	2.17-6.06	<0.001	3.00	1.26-7.17	0.013
Model[2]	2.17	1.36-3.47	0.001	2.43	1.44-4.10	0.001	1.41	0.53-3.74	0.486

[1]Adjusted for age, BMI, and hormone replacement therapy (HRT) use.
[2]Adjusted for age, BMI, HRT use, and tumor size.

6 DISCUSSION

6.1 METHODOLOGICAL CONSIDERATIONS

6.1.1 Study designs

6.1.1.1 Observational and experimental studies

Epidemiologic studies can be either observational or experimental. Observational studies can be purely descriptive, investigating e.g. the occurrence of disease in a certain population. They can also aspire to investigate causes of disease by studying associations which is the aim of this thesis.

In an observational study, the assignment of an individual to an exposure is outside of the control of the investigator. This leads to important drawbacks including confounding and bias (see below), hampering assumptions on causality. In an experimental study, on the other hand, a study subject is randomly allocated to a treatment (exposure) or control group by the investigator. If carried out properly, i.e. a correct randomization process with no cross-over between treatment arms, complete follow-up, and double-blind assessment of outcome, the experimental study is supposed to be free from systematic errors. Yet the role of chance can still operate in both types of studies. To minimize this risk, the study population has to be large, which is not always possible, especially in experimental studies. Probably the most important limitation of experimental studies, especially in humans, is that they can be unethical to perform; one cannot e.g. expose a group of people to a suspected carcinogen to study disease etiology. Moreover, certain exposures, e.g. mammographic density, are not possible to assign. In these cases we have to rely on animal models and observational studies to inform us of disease etiology but be aware of their limitations.

6.1.1.2 Cohort and case-control studies

A cohort is a group of individuals that is followed for a period of time. A cohort study consists of a group of individuals with a certain exposure and a group of individuals without the exposure. These two groups are followed for a certain time period and comparisons of different outcomes are then made between the two groups. Cohort studies are good when it comes to the study of rare exposures and has the advantage that one can study the association between the exposure and several different outcomes. Further, temporality is known, i.e. it is known that the exposure came before the outcome, and this is fundamental when trying to make inferences on causality.

Although a cohort study can initially be created to assess the impact of one exposure on different outcomes, it is not uncommon that information on other exposures is collected at the same time (or later on). The same cohort can then be re-classified as either exposed or non-exposed depending on the next exposure one would want to study.

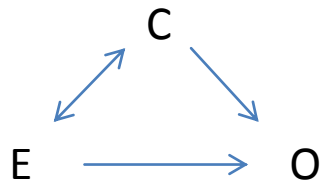
When an outcome is rare, e.g. breast cancer compared to a cold, it is much more efficient to create a study based on individuals with and without the outcome. This is called a case-control study and also gives the possibility of studying many different exposures at once. Individuals selected as controls have to be identified from the same study base as the cases were identified, i.e. if a control had developed the disease under

investigation, he/she would have been included as a case in the study. Sampling controls from anything but the study base, e.g. comparing female breast cancer cases to healthy males, will give rise to selection bias (see below). Since individuals are identified when they have developed the outcome (disease), collection of information on possible risk factors has to be done retrospectively, which is a limitation of case-control studies and may give rise to recall bias. This is a specific bias in case-control studies due to the ascertainment of exposure occurring after disease development. The result may be differences in the likelihood of cases and controls recalling an exposure.

Although studies I, III, and IV all are based on a case-control study, they could be considered cohort studies in the sense that they include a cohort of population-based breast cancer cases. In line with the benefits of a cohort study, the issue of control selection is also irrelevant for all studies. Both exposure and outcomes have been re-defined compared to the original study, thus, although information from the questionnaire was retrospective, recall bias cannot be present. Studies III and IV could also be considered prospective since the exposure and all possible confounders were assessed (or in the case of mammographic density, mammograms were taken) prior to the outcome (recurrence, distant metastasis, or death), thus, temporality is certain. However, retrospective collection of mammograms affected the selection of included women which is an important limitation of our studies (please see 6.1.2.2.1 Selection bias below).

6.1.2 Factors influencing validity

6.1.2.1 Confounding



Confounding occurs when the exposure and outcome have common causes. It affects the internal validity of the study. If we were to study coffee drinking and lung cancer we would most probably see an association since coffee drinkers tend to be smokers. To remove the effect of smoking on the relationship between coffee drinking on lung cancer, one could e.g. limit the study to non-smokers, adjust for smoking, or stratify on smoking. However, in reality, all confounders are not always known, and/or cannot always be accurately measured, resulting in residual confounding (33).

The probability of residual confounding due to lack of adjustment for possible confounders seems low in studies I, III, and IV, since we had detailed information on many related factors. However, we cannot completely rule out the possibility of residual confounding due to confounders that we do not yet know about or due to inaccurate assessment of possible confounders.

For study III, we lacked BMI and thus instead used absolute dense area as a measure of mammographic density instead of PD. This should have taken care of (or at least minimized) any possible confounding by BMI on the relationship between

mammographic density and molecular subtypes. As for the studies above, there may also be residual confounding due to unknown confounding factors or inaccurate assessment of confounders.

6.1.2.2 Bias

6.1.2.2.1 Selection bias

The most important threat to the internal validity of all the studies within this thesis is the possibility of selection bias, since all study participants had to have an available mammography. However, since only cases were included in the four studies, and mammography is one of the fundaments of detecting breast cancer, all cases would at least have a mammogram taken at diagnosis. However, the retrieval rate was affected for the following reasons: First, it is not uncommon that "healthy" mammograms older than 10 years are disposed. Secondly, we used analogue films so some films may have been misplaced. Thirdly, each time a woman has a mammogram carried out, old mammograms are studied for comparison, therefore, if a woman has an active disease, the likelihood of finding the mammograms in the storage area is lower. However, since we searched for mammograms in 2007-2009 i.e. >10 years after diagnosis, and events (recurrences and contralateral breast cancer) are most common within the first five years, it is unlikely that this will be a frequent problem. Further, a second retrieval round was carried out in 2010, minimizing this risk. The last and probably most important source of possible selection bias lies in the management of deceased women's' mammograms. These are often stored separately, and not always in order, making them more difficult to find. In some cases they are even discarded. This would lead to an exclusion of older cases and cases with a more aggressive disease than those included.

We did see a small, yet significant age difference between women with and without mammograms (62.9 for included women compared to 63.6 for excluded women, p=0.015). Nevertheless, selection bias can be controlled for by adjusting for the factors influencing selection (33) which we did with age. We could not adjust for vital status. Although selection bias can influence the estimation of the strength of the association it does not influence the validity of the hypothesis testing. In study III, the finding of a null association between mammographic density and survival is valid. In study IV, the p-values are also valid, though there is a risk that the strength of the association comparing survival in interval cancers to screening-detected cancers may be slightly biased due to a selection of survivors.

In study II we had the opposite problem; here subjects were excluded because of lack of frozen tumor. Excluded women therefore had a lower mean tumor diameter and less often metastatic lymph nodes. There was thus a selection of higher stage tumors than that of the source population. This was inevitable because of the harvest of RNA requiring a certain amount of tumor tissue. Hence, the results of the study may not be generalizable to women with smallest tumors. However, as pointed out above, the testing of the null hypothesis is still valid. We found no evidence of an association between mammographic density and molecular subtypes, i.e. could not reject the null hypothesis.

6.1.2.2.2 Lead-time and length bias

Comparisons between screening- and non-screening-detected cancers to evaluate the efficiency of mammography screening suffer from several biases e.g. selection bias,

lead-time bias, and length bias (see 2.2.5.2.4 Biases in estimating the effect of screening on mortality). In study IV, we compare the survival of interval cancer cases to screening-detected cases, to see if there are true biological differences between them. With this aim, only lead-time bias will be a concern. Selection bias will not be present since the interval cancer cases originate from the same population as the screening-detected cancers. The length bias is a true difference in biology - differences in tumors' proliferation rates - and thereby a difference we did not wish to remove. Screening-detected cancers and interval cancers will, however, have differences in lead-time, resulting in an artificial survival advantage for the former compared to the latter. This is because interval cancers will be diagnosed at a symptomatic stage, i.e. later in their natural history than screening-detected cancers, which mostly will be pre-symptomatic and diagnosed earlier in their natural history. Hence, we had to take lead-time into account and tried to do so by adjusting for tumor size, a proxy for time to diagnosis.

6.1.2.2.3 Misclassification/information bias

We used Cumulus, a semi-automated thresholding technique to minimized exposure misclassification. Further, the intra- and inter-observer reliability was high in all studies. For studies I, III, and IV the R^2 was 0.92 and the reader (LE) regularly calibrated herself against the gold standard of Cumulus measurement, test images previously read by Professor Norman Boyd (who was one of the developers of Cumulus). For study II the Pearson's rank correlation coefficients were 0.82 and 0.93 also high for inter- and intra-observer reliability, respectively. Furthermore, we took the mean of the density values to minimize measurement error. Nonetheless, exposure may have been misclassified in certain individuals. For this misclassification to be differential and have biased results, the misclassification of exposure must be associated with the outcome (tumor characteristics, molecular subtypes, and prognostic variables, respectively), which there is no reason to believe. Yet we cannot completely rule out a non-differential misclassification which is a misclassification of exposure not related to outcome. This would attenuate results and would therefore be of importance in the studies in which we have null findings (e.g. mammographic density and tumor characteristics, molecular subtypes, and survival). However, since inter- and intra-correlation was high throughout, the influence of misclassification should be low.

6.1.2.3 Chance

Not only can systematic errors influence the results in epidemiologic studies, but chance also plays an important role. Increasing the number of participants in a study is the best way to reduce its impact. We can also assess the uncertainty of results by calculating p-values and confidence intervals (CI).

6.1.2.3.1 The p-value

In all studies, we construct a null hypothesis (H_0) (e.g. that there is no association between exposure E and outcome O), and an alternative hypothesis (H_A) (there is an association between E and O). A p-value measures the certainty with which we can say that an observed association (or a more extreme association than the observed) would appear by chance. We often use the cut-off of 0.05 for the p-value to discriminate between "true" and chance findings. This means that if the p-value<0.05, then there is a

5% chance that the observed association (or a more extreme one) would appear by chance. If we are only conducting one test, are not conducting multiple tests of the same association, and have a probable *à priori* hypothesis about the relationship, we usually deem this a satisfactory cut-off at which to reject the null hypothesis.

The p-value says nothing of the strength of an association; it can be small in a large study although the effect of an exposure is weak, and it can be large in a small study, although the effect is strong.

6.1.2.3.2 The confidence interval

In contrast to the p-value, the confidence interval (CI) incorporates both the significance and strength of an association. Throughout this thesis, the 95% CI is used which means that one can be 95% confident that the "true" value lies within that range. If the CI was set at 99%, the confidence that we have captured the true estimate of the association would increase, but the width of the CI would also increase. The width of the CI is also affected by the size of the study, so that the width increases with smaller study samples and vice versa.

6.1.2.4 Type I and type II errors

Type I and type II errors refer to the erroneous acceptance or rejection of a given null hypothesis. A type I error is the rejection of a null hypothesis that is actually true. In study III, for example, we found an association between mammographic density and local recurrence. However, if there really is no association, we have falsely rejected the null hypothesis and have committed a type I error. A type II error occurs if one instead erroneously accepts the null hypothesis when it should be rejected. The probability of not making a type II error is called the statistical power and is dependent on study size and effect size. Study II is an example where a type II error may have occurred. In study II, we did not find any association between mammographic density and molecular subtypes which could be a result of the relatively small study population. Study III is also an example of a possible type II error; in study III we found an association between mammographic density and local and locoregional recurrence but no association with breast cancer specific survival. If the effect of mammographic density on survival is small, we may not have had a large enough study sample to observe the association and may therefore have made a type II error. Indeed, it was not until the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) conducted a large meta-analysis in 2005 that it became clear that locoregional recurrences influence survival (48).

6.2 FINDINGS AND INTERPRETATION

6.2.1 Study I

Among the tumor characteristics studied, we only found an association between PD and tumor size. Interestingly, neither presence of lymph node metastasis nor hormone receptor status was associated with PD.

Although there was a statistically significant, positive association between PD and tumor size, the effect size was relatively small. The association became more pronounced when mode of detection was excluded from the model, lending support to

the hypothesis that it is due to masking. However, dense breasts may also give rise to more highly proliferative tumors which may be the reason why there still is an association between PD and tumor size even after adjustment for mode of detection. To try to disentangle the cause of this relationship in an epidemiological study is difficult, but an investigation of the association between PD and proliferation rate could give a clue to the etiology. Unfortunately, we were not able to study this due to too many missing values for proliferation rate (approximately 70% missing values).

Since MD decreases screening sensitivity (90, 205), our *à priori* hypothesis was that density would be correlated with both tumor size and lymph node metastasis. However, we found no evidence of an association between PD and the latter. Previous studies have been inconsistent (206-210). The ability for breast tumors to metastasize is not only dependent on time but also on the acquired ability to invade vessels, survive migration, extravasate, and colonize distant sites (45). In agreement with this, Jatoi et al. observed that lymph node metastasis was a marker of both time to diagnosis and tumor aggressiveness (211). Hence, although density may cause a delay in diagnosis, it does not per se imply a direct association with lymph node metastasis.

Although MD is a hormonally responsive trait (212), we saw no association between density and hormone receptor status. This is in line with most previous studies (157, 213) but is in contrast with another study of postmenopausal women in which higher density was associated with increased risk of ER-negative cancers (210). However, mode of detection was not accounted for and the proportions of ever-users of HRT (76% of the study population) and women with previous benign breast disease (59% of the study population) were very high, both of which contribute to increased PD (203, 214) and increased risk of interval cancers (215, 216). In our study, 50% of the participants were ever-users of HRT (including all compounds and modes of administration) and 14% had previous benign breast disease. The proportions were the same as for women lacking density measurements as was the proportion of ER-negative cancers. Thus, we believe that our results are representative.

Heusinger et al. recently published a large study on PD and tumor characteristics, including both pre-and postmenopausal women, in which they found that PD was associated with lower ER-expression (217). However, although highly statistically significant, differences in PD between categories of ER-expression were very small, and there was no dose-response relationship.

A limitation of this study is that it was solely composed of postmenopausal women; hence, our results may not be applicable to premenopausal populations. MD may differ histologically in pre- and postmenopausal women (164), and has been shown to be influenced by different hormonal factors depending on menopausal status (196, 197). Furthermore, both age and menopausal status influence tumor characteristics (146). Thus, well-conducted studies are needed to investigate the relationship between MD and tumor characteristics in premenopausal women.

6.2.2 Study II

We found no associations between AD and Sorlie-Perou subtypes; neither between AD and individual subtypes, nor between AD and subtype as a whole. However, our study population was relatively small and the null findings could simply be due to low power. Hence, larger studies are needed to confirm our results.

A couple of studies have previously attempted to investigate the association between MD and Sorlie-Perou subtypes using receptor status (ER-, PR-, and HER-2-status) as proxies for the different molecular subtypes (213, 218-221). Ma et al. (219) studied the association between PD and the luminal A and basal-like subtypes and found no relationship in case-only analyses. In case-control analyses they observed positive associations between PD and both the luminal A and basal-like subtypes, but as these associations were of a similar magnitude got they are likely to simply reflect the general increase in breast cancer risk associated with PD. Phipps et al. (213) also conducted a case-control study investigating the association between density (assessed using a visual categorical classification, BI-RADS), ER+, ER-/PR-/HER2+, and triple-negative breast cancers. They achieved the same results as Ma et al. (219) i.e. that density was similarly, positively associated with all subtypes. They concluded that although the different subtypes are distinct biological entities, this is not a result of differences in the association with MD. We thus believe that our results are in agreement with both studies. Arora et al. (218) studied the association between density and the luminal A, luminal B, basal-like, and ERBB-2 subtypes also using BI-RADS to assess density. They observed that women with extremely dense breasts had a higher frequency of luminal A tumors ($p=0.05$). However, as only age was adjusted for in this analysis the findings might have been affected by confounding.

Our findings are in line with most studies relating to the association between MD and other tumor characteristics such as hormone-receptor status (157, 222), HER2-status (157, 210, 222, 223), and Ki67 (206, 224). Furthermore, most studies addressing the relationship between MD and survival have found no association (207, 225-227). We find this to be in indirect support of the null association between MD and Sorlie-Perou subtypes shown in this study, since molecular subtypes are associated with both tumor characteristics and prognosis (228, 229).

6.2.3 Study III

We found that mammographic density increases risk of local and locoregional recurrence after diagnosis of invasive breast cancer. There was no association between density and distant metastasis or between density and survival in our postmenopausal study population.

A PD of ≥25% nearly doubled the risk of local recurrence compared to women with PD<25% and increased the risk of locoregional recurrence by almost 70%, independent of established breast cancer prognosticators. The association between PD and locoregional recurrence is in agreement with two of three previously published studies investigating this relationship (227, 230), although they were both limited to women who had undergone breast-conserving surgery.

The relationship between PD and local recurrence could be caused by PD masking residual disease after breast-conserving surgery. This reservation is especially important in our study since we lacked information on the status of the surgical margin. We thus stratified on type of breast surgery and found that the effect was most pronounced in the group of women exposed to total mastectomy, undermining this hypothesis.

We did not see an association between density and distant recurrence nor with density and survival. Distant metastasis is closely linked with breast cancer mortality whereas locoregional recurrence has a moderate impact on survival (48). Since we only had 10-

year follow-up of survival, we cannot exclude that there may be a relationship between PD and long term survival, nor can we exclude that there is a small effect of PD on survival requiring even larger studies to identify the association. However, our findings of a lack of association between MD, distant metastasis, and survival is in accordance with both previous studies pertaining to MD and distant metastasis (227, 230), and two of three studies on MD and survival (207, 225, 227). The conflicting study relating to density and prognosis (226) showed a lower case fatality for women with high density compared to women with low density. However, since only age was adjusted for, this relationship is most likely due to confounding.

A limitation of our study was that the study population was solely composed of postmenopausal women; hence, our results may not be applicable to pre-menopausal populations. We also lacked information on HER2 status and Ki67 since they were not clinically in use in Sweden at the time our participants were diagnosed with breast cancer. Both HER2-status and Ki67 influence prognosis (121, 231). However, for these factors to have influenced our results, they must be associated with MD, which there is currently no evidence of (210, 219, 223, 224).

6.2.4 Study IV

We found that ICs, overall, have a worse prognosis than screening-detected cancers. Both interval cancers in non-dense and dense breasts (defined as PD<25% and PD≥25%, respectively,) had a poorer survival than corresponding screening-detected cancers before adjusting for tumor size, a proxy for time to diagnosis. After adjusting for tumor size, ICs in non-dense breasts still had a HR of >2 compared to corresponding screening-detected cancers. For women with dense breasts, however, there was no longer evidence of a statistically significant difference between ICs and screening-detected cancers. We thus conclude that ICs in non-dense breasts seem to be truly aggressive tumors, whereas the poorer prognosis of ICs in dense breasts compared to corresponding screening-detected cancers seems primarily due to differences in time to diagnosis.

In support of the above, ICs in non-dense breasts had particularly unfavorable tumor characteristics, correlating well with their poorer prognosis. Tumor characteristics in ICs in dense breasts were more similar to their corresponding screening-detected cancers. However, ICs in dense breasts were larger, more often positive for lymph node metastasis, and of higher grade; the first two differences could be in line with the hypothesis that these primarily are masked tumors.

A limitation of our study is that our study population was solely composed of postmenopausal women. Our results may not be applicable to premenopausal populations; hence, studies are needed to investigate the relationship between mammographic density, ICs, and survival in premenopausal women.

In conclusion, women with interval cancers had a worse prognosis than women with screening-detected cancers. Women diagnosed with ICs in non-dense breasts seemed to have a particularly aggressive phenotype and poor prognosis. In women diagnosed with ICs in dense breasts, the poorer prognosis seemed primarily due to differences in time to diagnosis rather than differences in tumor aggressiveness. Since this is the first study of its kind, it is, however, important to verify our findings. If they hold true, future studies should focus on establishing the risk factors associated with the more aggressive

entity of interval cancers in non-dense breasts, hopefully, allowing for primary preventive measures to be taken in the future.

6.3 OVERALL CONCLUSIONS

Mammographic density does not affect lymph node metastasis, ER-status, PR-status, grade, or histopathological classification, but is positively associated with tumor size.

Mammographic density is not associated with molecular subtypes of breast cancer, but larger studies are needed to validate our findings.

Mammographic density is an independent risk factor of local and locoregional recurrence.

Although high mammographic density increases risk of local and locoregional recurrence, it is not associated with distant metastasis or survival. Nonetheless, larger studies are needed to rule out a weak association between density and survival.

Interval cancers diagnosed in non-dense breasts seem to be more aggressive than corresponding screening-detected cancers.

In dense breasts, there is no evidence of a statistically significantly worse prognosis for interval cancers than screening-detected cancers, after taking time to diagnosis into account.

In conclusion, out of all phenotypic variables studied, we only found a relationship between mammographic density and tumor size, as well as mammographic density and local and locoregional recurrence. The latter associations were independent of type of surgery; the relationships were actually more pronounced in totally mastectomized patients. This is quite surprising – how does mammographic density at diagnosis, which is excised, influence recurrence years later? I have two theories: First, there will probably be women for whom some breast tissue is left despite a total mastectomy. If this tissue is dense rather than non-dense, it may increase the risk of local recurrence in the same way it increased the risk of primary breast cancer. Secondly, and more speculatively, maybe mammographic density increases risk of self-seeding, which I will describe in detail below.

Tumor dissemination was long viewed as a one-way process. However, in 2009, Kim et al. (232) showed that disseminated tumor cells could return to the primary tumor and re-infiltrate it, a process they referred to as self-seeding. This capacity of re-infiltrating the primary tumor was also seen for cells from metastatic colonies.

The self-seeding hypothesis was originally generated when the same research group saw that some tumors grew faster than others, despite that the faster-growing tumors had lower proliferation rates. Hence, they hypothesized that disseminated tumor cells re-colonized the primary tumor. The self-seeded primary tumor could then grow more quickly than other unseeded tumors, despite the slower growth rate. They were also able to prove this later on (232).

It is known that large quantities of tumor cells leave the primary tumor and enter the blood stream and that this may occur at early stages (232). However, only a minority of these cells give rise to metastasis due to many obstacles, among other things, unfavorable conditions at the new site. Self-seeding, on the other hand, does not require all of these capacities, since the circulating tumor cell returns "home", where vessels are leaky (easy to extravasate) and the environment familiar (no/little adaptation is required) (233).

The primary tumor can either be self-seeded by circulating tumor cells coming directly from the primary tumor or from metastatic sites (232). N.B. metastatic tumor cells may reside in a dormant manner in a new organ (233), thus, seeding of the primary site does not require manifest metastasis. Kim et al. therefore hypothesized that this could explain why local recurrence can follow a complete eradication of the primary tumor (232).

Mammographic density has been hypothesized to give rise to more aggressive tumors due to the increased stromal composition of the breast and increased deposition of collagen. Despite this, we and others (227, 230) were not able to show an association between mammographic density and risk of distant metastasis. Perhaps, however, density does increase risk of dissemination, but maybe these circulating tumor cells prefer to colonize or are only capable of colonizing the primary tumor/breast.

The self-seeding hypothesis could thus explain the observed association between mammographic density and tumor size, despite that mammographic density has not yet been shown to be associated with increased proliferation rate. It would further explain the relationship between mammographic density and local and locoregional recurrence, in spite of the lacking relationship with distant metastasis. Finally, the strong relationship of mammographic density and recurrence in totally mastectomized patients would not be as surprising.

7 FUTURE PERSPECTIVE

Imagine the endless depth of information a mammography unit holds – mammograms for countless women, spanning several years of their lives. Some of these women will develop breast cancer and others will not. Some will be cured the minute they get off the surgery table, while others will die in spite of all treatments. Who is who? Neither our tools of risk prediction nor our prognostic tools are satisfactory. Yet, for many years, we neglected the information available within mammograms and are just now realizing their potential.

It is now well-established that high mammographic density increases breast cancer risk. However, we still lack knowledge of what mammographic density really is and how it relates to breast cancer. Not surprisingly, mammographic density seems to be a reflection of the number of epithelial and stromal cells in the breast. Some therefore believe that density purely is a reflection of the number of cells at risk, but I do not think it is that simple.

As both John Wolfe and Laszlo Tabár noted, mammographic density isn't just white. It appears to have different textures – from a flat, homogeneous mass, to multiple round little islands, to a rugged, streaky appearance, and so on. Although we did not find that quantitative mammographic density was differentially associated with different tumor phenotypes, it seems more plausible that qualitative mammographic density is. This would be in agreement with studies showing e.g. that the organization of collagen affects the degree of malignancy in breast tumors. However, when investigating this, it is important that the qualitative measure of density is assessed objectively in order to minimize exposure misclassification. Our research group is currently working on an automated density assessment tool which incorporates an objective evaluation of texture that was shown to add predictive value to breast cancer risk estimates. Hopefully, this will also add valuable information pertaining to prognosis.

In clinical practice, we are using a mammogram as a simple means of answering the question whether a woman has cancer or not. I believe that a mammogram holds much more information than that. We don't just get a binary value of a tumor, we see the qualitative specifics of the tumor and we also see the tumor environment - an important factor of carcinogenesis and tumor progression. Furthermore, we can compare the mammogram to previous mammograms and, when appropriate, assess interval cancer status, and evaluate possible changes in the structure and amounts of the different tissues. I think that these variables contain important prognostic information independent of established prognosticators. Furthermore, if we could combine this information with e.g. histological, cytological, and genetic examinations, I believe that we would come a long way in our understanding of breast cancer biology.

We know so much and so little. It appears that for every new answer we get, only more questions arise. The complexity of cancer is intriguing, and, at times, also overwhelming. Within each field of cancer research we are digging deeper and deeper, but we also have to be able to see the big picture and cross the gaps between the

disciplines to create new and greater research. Only then can we understand cancer and remove the obstacles impeding us from curing and preventing it.

8 ACKNOWLEDGEMENTS

First of all, I would like to thank all the women who contributed to the studies within this thesis. With your help, we have added another piece to the puzzle, and for every little step we take, we are coming closer to being able to eradicate breast cancer.
I would also like to thank all the personnel at the mammography units for being so kind in lending us their mammograms. Furthermore, I would like to express my gratitude to the following individuals who all helped me in different ways during this journey:

Per Hall, my main supervisor –It's been a privilege to work with such a creative mind. The way you pursue and execute your high-flying aims in research is inspiring. Thank you for letting me work with you, for all the encouragement, and all the fun times throughout these years!
Kamila Czene, my co-supervisor – For all your support – both when things were up and when they were down. Your warmth, frankness, and brilliant intellect make it a pleasure to work with you!
Keith Humphreys, my co-supervisor –For being so patient and for always taking your time to sit down and talk all the statistical mumbo jumbo through. I appreciate your kind and calm way, and for giving me perspective when everything seemed frantic.
Lena Rosenberg, my co-supervisor –For all the phone talks and e-mails when you've helped me with both CAHRES and life at large! I appreciate your studious ways and always felt confident when you'd read a manuscript.
Alex Ploner – For all your help and for taking your time to explain statistical issues in a very pedagogical way.
Isabel dos Santos Silva and Valerie McCormack – For your meticulous work and insightful critique!
Hans-Olov Adami, Nancy Pedersen, and Henrik Grönberg – For creating a stimulating research environment that I have had the opportunity to enjoy and grow in.
Tomas Sonnenfeld, my mentor - For your wholehearted support, for taking time from your busy schedule to listen to my ideas and problems, and for all the hysterical lunches and coffees we've had. I can't imagine having a better mentor than you.
Hasse Walum – For all our discussions about both the large and small aspects of life. I will miss them and you.
My roomies, Edoardo Colzani and Thang Trinh– For putting up with me and my bad habits (although I probably don't have many ☺), for all the discussions and all the nonsense, and for making me smile when I've been down.
Mattias Hammarström – For all the lunches and talks – it's been a pleasure!
To the rest of the breast cancer group, Maria Sandberg, Anki Kristensen, Ann-Sofie Andersson, Milka Krestelica, Agneta Lönn, Gerd Agerberg, Ami Rönnberg, Krystyna Håkansson, Hatef Darabi, Jingmei Li, Linda Lindström, Ulf Hannelius, Daniel Klevebring, Mikael Karlsson, and the list goes on… - For making my time at MEB memorable! And to the data collectors – thank you for your tremendous work!
The MEB faculty, especially Olof Nyrén, Paul Dickman, and Arvid Sjölander – For teaching me many issues in epidemiology and biostatistics, and for making it enjoyable.

IT, especially Gunnar Petersson, Gastón Nuñez, Zack Yusof, and Torbjörn Gustafsson – For all your help and for including me in the "IT-corridor-gang" when I first came to MEB ☺.
To the administrative staff, especially Camilla Ahlqvist – For all your help, always delivered with a smile.
To all other colleagues at MEB – For contributing to the friendly and stimulating atmosphere.
Fuat Celebioglu – For believing in me and for showing me what a great surgeon is.
Jerzy Lewin – If it weren't for you, I would never have embarked upon this path – thank you!
Mikael Hartman – For bringing me to MEB to meet Pelle.
All my colleagues at Bröstcentrum, Södersjukhuset – Thank you for making my clinical work so enjoyable!

To all my friends – For true friendship - thank you.
Lena, Olav, and all my new siblings – For including me in your family, for always making me feel welcome, and for caring.
Martin and Michael – For humbling me ☺ and for being the best big brothers a little sister could have.
Mamma and pappa – For opening the world to me and showing me all the possibilities life has to offer. Thank you also for all your endless love and support. Lastly, thank you for all your help with Eyvind; without it, this thesis would never have been possible.
David – You are the love of my life. Thank you for everything.
Eyvind - Eyvind-peyvind, snuse, buse, and älskling - you are my sunshine.

9 REFERENCES

1. U. S. National Institutes of Health NCI. SEER Training Modules, Breast Cancer. http://training.seer.cancer.gov/breast/anatomy; 2012.
2. Sleeman KE, Kendrick H, Robertson D, Isacke CM, Ashworth A, Smalley MJ. Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. The Journal of cell biology. 2007;176(1):19-26. Epub 2006/12/28.
3. Booth BW, Smith GH. Estrogen receptor-alpha and progesterone receptor are expressed in label-retaining mammary epithelial cells that divide asymmetrically and retain their template DNA strands. Breast Cancer Res. 2006;8(4):R49. Epub 2006/08/03.
4. Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardiner-Garden M, Hilton HN, et al. The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes & development. 2008;22(5):581-6. Epub 2008/03/05.
5. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, et al. Generation of a functional mammary gland from a single stem cell. Nature. 2006;439(7072):84-8. Epub 2006/01/07.
6. Sleeman KE, Kendrick H, Ashworth A, Isacke CM, Smalley MJ. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast Cancer Res. 2006;8(1):R7. Epub 2006/01/19.
7. Swedish Breast Cancer Group S. National guidelines for treatment of breast cancer (Nationella riktlinjer för behandling av bröstcancer). 2012 [updated 2012-02-112012-08-20]; 3.0:[Available from: http://www.swebcg.se/Files/Docs/Nationella_riktlinjer120211.pdf.
8. Gusterson BA, Stein T. Human breast development. Seminars in cell & developmental biology. 2012;23(5):567-73. Epub 2012/03/20.
9. Russo J, Russo IH. Development of the human breast. Maturitas. 2004;49(1):2-15. Epub 2004/09/08.
10. Birkenfeld A, Kase NG. Functional anatomy and physiology of the female breast. Obstetrics and gynecology clinics of North America. 1994;21(3):433-44. Epub 1994/09/01.
11. Hassiotou F, Geddes D. Anatomy of the human mammary gland: Current status of knowledge. Clinical anatomy. 2012. Epub 2012/09/22.
12. Milanese TR, Hartmann LC, Sellers TA, Frost MH, Vierkant RA, Maloney SD, et al. Age-related lobular involution and risk of breast cancer. J Natl Cancer Inst. 2006;98(22):1600-7. Epub 2006/11/16.
13. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010;127(12):2893-917. Epub 2011/02/26.
14. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol. 2001;2(9):533-43. Epub 2002/03/22.
15. Hoover RN. That recognised risk factors can explain past and present international differences in breast cancer incidence: misconceptions 5. Br J Cancer. 2012;107(3):408-10. Epub 2012/07/26.
16. Tyczynski J, Tarkowski W, Parkin DM, Zatonski W. Cancer mortality among Polish migrants to Australia. Eur J Cancer. 1994;30A(4):478-84. Epub 1994/01/01.
17. Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AM, West DW, et al. Migration patterns and breast cancer risk in Asian-American women. J Natl Cancer Inst. 1993;85(22):1819-27. Epub 1993/11/17.
18. Bray F, McCarron P, Parkin DM. The changing global patterns of female breast cancer incidence and mortality. Breast Cancer Res. 2004;6(6):229-39. Epub 2004/11/13.

19. Cancer Statistics (Cancerstatistik) [database on the Internet]. 2012 [cited 2012-10-07]. Available from: http://192.137.163.49/sdb/can/val.aspx www.socialstyrelsen.se.
20. Cancer Incidence in Sweden 2010: Socialstyrelsen; 2011 [cited 2012 2012-06-26]. Available from: http://www.socialstyrelsen.se/publikationer2011/2011-12-15.
21. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. Jama. 2002;288(3):321-33. Epub 2002/07/19.
22. Lambe M, Wigertz A, Holmqvist M, Adolfsson J, Bardage C, Fornander T, et al. Reductions in use of hormone replacement therapy: effects on Swedish breast cancer incidence trends only seen after several years. Breast Cancer Res Treat. 2010;121(3):679-83. Epub 2009/11/07.
23. Bouchardy C, Usel M, Verkooijen HM, Fioretta G, Benhamou S, Neyroud-Caspar I, et al. Changing pattern of age-specific breast cancer incidence in the Swiss canton of Geneva. Breast Cancer Res Treat. 2010;120(2):519-23. Epub 2009/07/28.
24. Canfell K, Banks E, Moa AM, Beral V. Decrease in breast cancer incidence following a rapid fall in use of hormone replacement therapy in Australia. The Medical journal of Australia. 2008;188(11):641-4. Epub 2008/06/03.
25. Clarke CA, Glaser SL, Uratsu CS, Selby JV, Kushi LH, Herrinton LJ. Recent declines in hormone therapy utilization and breast cancer incidence: clinical and population-based evidence. J Clin Oncol. 2006;24(33):e49-50. Epub 2006/11/23.
26. Hemminki E, Kyyronen P, Pukkala E. Postmenopausal hormone drugs and breast and colon cancer: Nordic countries 1995-2005. Maturitas. 2008;61(4):299-304. Epub 2008/11/11.
27. Katalinic A, Rawal R. Decline in breast cancer incidence after decrease in utilisation of hormone replacement therapy. Breast Cancer Res Treat. 2008;107(3):427-30. Epub 2007/04/25.
28. Parkin DM. Is the recent fall in incidence of post-menopausal breast cancer in UK related to changes in use of hormone replacement therapy? Eur J Cancer. 2009;45(9):1649-53. Epub 2009/02/17.
29. Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, et al. The decrease in breast-cancer incidence in 2003 in the United States. N Engl J Med. 2007;356(16):1670-4. Epub 2007/04/20.
30. Seradour B, Allemand H, Weill A, Ricordeau P. Changes by age in breast cancer incidence, mammography screening and hormone therapy use in France from 2000 to 2006. Bulletin du cancer. 2009;96(4):E1-6. Epub 2009/05/14.
31. Vankrunkelsven P, Kellen E, Lousbergh D, Cloes E, Op de Beeck L, Faes C, et al. Reduction in hormone replacement therapy use and declining breast cancer incidence in the Belgian province of Limburg. Breast Cancer Res Treat. 2009;118(2):425-32. Epub 2009/02/25.
32. Li CI, Daling JR. Changes in breast cancer incidence rates in the United States by histologic subtype and race/ethnicity, 1995 to 2004. Cancer Epidemiol Biomarkers Prev. 2007;16(12):2773-80. Epub 2007/12/19.
33. Adami H-O, Hunter D, Trichopoulos D. Textbook of Cancer Epidemiology. Second ed: Oxford University Presss; 2008.
34. Welfare' SNBoHa. Causes of Death 20102011 2012-09-22.
35. Dödsorsaksstatistik (Statistics of cause of death) [database on the Internet]. 2012.
36. Cancer Statistics [database on the Internet]. 2011 [cited 120615]. Available from: www.socialstyrelsen.se/register/halsodataregister/cancerregistret/inenglish.
37. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70. Epub 2000/01/27.
38. Lazebnik Y. What are the hallmarks of cancer? Nat Rev Cancer. 2010;10(4):232-3. Epub 2010/04/01.
39. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74. Epub 2011/03/08.

40. Sonnenschein C, Soto AM. The Society of Cells: Cancer and Control of Cell Proliferation. New York, New York: Springer Verlag; 1999.
41. Weaver VM, Gilbert P. Watch thy neighbor: cancer is a communal affair. J Cell Sci. 2004;117(Pt 8):1287-90. Epub 2004/03/17.
42. Moinfar F, Man YG, Arnould L, Bratthauer GL, Ratschek M, Tavassoli FA. Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. Cancer Res. 2000;60(9):2562-6. Epub 2000/05/16.
43. Paszek MJ, Weaver VM. The tension mounts: mechanics meets morphogenesis and malignancy. Journal of mammary gland biology and neoplasia. 2004;9(4):325-42. Epub 2005/04/20.
44. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. Cell. 2009;139(5):891-906. Epub 2009/11/26.
45. Weinberg R. The Biology of Cancer. 1st edn. ed: Garland Science; 2006.
46. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell. 2011;147(2):275-92. Epub 2011/10/18.
47. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. Science. 2011;331(6024):1559-64. Epub 2011/03/26.
48. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans E, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;366(9503):2087-106. Epub 2005/12/20.
49. Welfare SNBoHa. Cancer incidence in Sweden 20102011 2012-06-21.
50. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. Nature. 1983;303(5920):767-70. Epub 1983/06/30.
51. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006;15(6):1159-69. Epub 2006/06/16.
52. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer. 1997;71(5):800-9. Epub 1997/05/29.
53. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. Br J Cancer. 2000;83(10):1301-8. Epub 2000/10/25.
54. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet. 2003;72(5):1117-30. Epub 2003/04/05.
55. Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. Nature genetics. 2008;40(1):17-22. Epub 2007/12/29.
56. Fanale D, Amodeo V, Corsini LR, Rizzo S, Bazan V, Russo A. Breast cancer genome-wide association studies: there is strength in numbers. Oncogene. 2012;31(17):2121-8. Epub 2011/10/15.
57. Key T, Appleby P, Barnes I, Reeves G, Endogenous H, Breast Cancer Collaborative G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst. 2002;94(8):606-16. Epub 2002/04/18.
58. Eliassen AH, Missmer SA, Tworoger SS, Spiegelman D, Barbieri RL, Dowsett M, et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. J Natl Cancer Inst. 2006;98(19):1406-15. Epub 2006/10/05.
59. Missmer SA, Eliassen AH, Barbieri RL, Hankinson SE. Endogenous estrogen, androgen, and progesterone concentrations and breast cancer risk among postmenopausal women. J Natl Cancer Inst. 2004;96(24):1856-65. Epub 2004/12/17.
60. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiologic reviews. 1993;15(1):36-47. Epub 1993/01/01.
61. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and

108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. Lancet. 1997;350(9084):1047-59. Epub 1999/04/23.
62. Ewertz M, Duffy SW, Adami HO, Kvale G, Lund E, Meirik O, et al. Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries. Int J Cancer. 1990;46(4):597-603. Epub 1990/10/15.
63. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. Lancet Oncol. 2001;2(3):133-40. Epub 2002/03/21.
64. Collaborative Group on Hormonal Factors in Breast C. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. Lancet. 2002;360(9328):187-95. Epub 2002/07/23.
65. Key TJ, Appleby PN, Reeves GK, Roddam A, Dorgan JF, Longcope C, et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. J Natl Cancer Inst. 2003;95(16):1218-26. Epub 2003/08/21.
66. Judd HL, Shamonki IM, Frumar AM, Lagasse LD. Origin of serum estradiol in postmenopausal women. Obstet Gynecol. 1982;59(6):680-6. Epub 1982/06/01.
67. Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V, et al. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. Lancet Oncol. 2011;12(8):785-94. Epub 2011/07/26.
68. Narod SA. Hormone replacement therapy and the risk of breast cancer. Nature reviews Clinical oncology. 2011;8(11):669-76. Epub 2011/08/03.
69. Lee SA, Ross RK, Pike MC. An overview of menopausal oestrogen-progestin hormone therapy and breast cancer risk. Br J Cancer. 2005;92(11):2049-58. Epub 2005/05/19.
70. Chlebowski RT, Kuller LH, Prentice RL, Stefanick ML, Manson JE, Gass M, et al. Breast cancer after use of estrogen plus progestin in postmenopausal women. N Engl J Med. 2009;360(6):573-87. Epub 2009/02/07.
71. Kerlikowske K, Cook AJ, Buist DS, Cummings SR, Vachon C, Vacek P, et al. Breast cancer risk by breast density, menopause, and postmenopausal hormone therapy use. J Clin Oncol. 2010;28(24):3830-7. Epub 2010/07/21.
72. Trichopoulos D. Hypothesis: does breast cancer originate in utero? Lancet. 1990;335(8695):939-40. Epub 1990/04/21.
73. Michels KB, Xue F. Role of birthweight in the etiology of breast cancer. Int J Cancer. 2006;119(9):2007-25. Epub 2006/07/11.
74. Park SK, Kang D, McGlynn KA, Garcia-Closas M, Kim Y, Yoo KY, et al. Intrauterine environments and breast cancer risk: meta-analysis and systematic review. Breast Cancer Res. 2008;10(1):R8. Epub 2008/01/22.
75. Silva Idos S, De Stavola B, McCormack V, Collaborative Group on Pre-Natal Risk F, Subsequent Risk of Breast C. Birth size and breast cancer risk: re-analysis of individual participant data from 32 studies. PLoS Med. 2008;5(9):e193. Epub 2008/10/03.
76. Bodian CA. Benign breast diseases, carcinoma in situ, and breast cancer risk. Epidemiologic reviews. 1993;15(1):177-87. Epub 1993/01/01.
77. Longnecker MP. Alcoholic beverage consumption in relation to risk of breast cancer: meta-analysis and review. Cancer Causes Control. 1994;5(1):73-82. Epub 1994/01/01.
78. Palmer JR, Rosenberg L. Cigarette smoking and the risk of breast cancer. Epidemiologic reviews. 1993;15(1):145-56. Epub 1993/01/01.
79. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. Breast Cancer Res. 2005;7(1):21-32. Epub 2005/01/12.
80. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst. 1989;81(24):1879-86. Epub 1989/12/20.
81. Gail MH, Costantino JP. Validating and improving models for projecting the absolute risk of breast cancer. J Natl Cancer Inst. 2001;93(5):334-5. Epub 2001/03/10.

82. Schonfeld SJ, Pee D, Greenlee RT, Hartge P, Lacey JV, Jr., Park Y, et al. Effect of changing breast cancer incidence rates on the calibration of the Gail model. J Clin Oncol. 2010;28(14):2411-7. Epub 2010/04/07.
83. Claus EB, Risch N, Thompson WD. The calculation of breast cancer risk for women with a first degree family history of ovarian cancer. Breast Cancer Res Treat. 1993;28(2):115-20. Epub 1993/11/01.
84. Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Lancet. 1994;343(8899):692-5. Epub 1994/03/19.
85. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Statistics in medicine. 2004;23(7):1111-30. Epub 2004/04/02.
86. Amir E, Evans DG, Shenton A, Lalloo F, Moran A, Boggis C, et al. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. Journal of medical genetics. 2003;40(11):807-14. Epub 2003/11/25.
87. Assi V, Warwick J, Cuzick J, Duffy SW. Clinical and epidemiological issues in mammographic density. Nature reviews Clinical oncology. 2012;9(1):33-40. Epub 2011/12/07.
88. Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. Breast Cancer Res Treat. 2005;94(2):115-22. Epub 2005/11/02.
89. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006;98(17):1204-14. Epub 2006/09/07.
90. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. Jama. 1996;276(1):33-8. Epub 1996/07/03.
91. Burke W, Daly M, Garber J, Botkin J, Kahn MJ, Lynch P, et al. Recommendations for follow-up care of individuals with an inherited predisposition to cancer. II. BRCA1 and BRCA2. Cancer Genetics Studies Consortium. Jama. 1997;277(12):997-1003. Epub 1997/03/26.
92. Saslow D, Boetes C, Burke W, Harms S, Leach MO, Lehman CD, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. CA: a cancer journal for clinicians. 2007;57(2):75-89. Epub 2007/03/30.
93. Robson M, Offit K. Clinical practice. Management of an inherited predisposition to breast cancer. N Engl J Med. 2007;357(2):154-62. Epub 2007/07/13.
94. Mavaddat N, Barrowdale D, Andrulis IL, Domchek SM, Eccles D, Nevanlinna H, et al. Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). Cancer Epidemiol Biomarkers Prev. 2012;21(1):134-47. Epub 2011/12/07.
95. Yeatman TJ, Cantor AB, Smith TJ, Smith SK, Reintgen DS, Miller MS, et al. Tumor biology of infiltrating lobular carcinoma. Implications for management. Annals of surgery. 1995;222(4):549-59; discussion 59-61. Epub 1995/10/01.
96. Hamilton LJ, Evans AJ, Wilson AR, Scott N, Cornford EJ, Pinder SE, et al. Breast imaging findings in women with BRCA1- and BRCA2-associated breast carcinoma. Clinical radiology. 2004;59(10):895-902. Epub 2004/09/29.
97. Tilanus-Linthorst M, Verhoog L, Obdeijn IM, Bartels K, Menke-Pluymers M, Eggermont A, et al. A BRCA1/2 mutation, high breast density and prominent pushing margins of a tumor independently contribute to a frequent false-negative mammography. Int J Cancer. 2002;102(1):91-5. Epub 2002/09/28.
98. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. Jama. 2005;293(10):1245-56. Epub 2005/03/10.
99. Olsson S, Andersson I, Karlberg I, Bjurstam N, Frodis E, Hakansson S. Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme. J Med Screen. 2000;7(1):14-8. Epub 2000/05/12.

100. Welfare SNBoHa. Nationella Riktlinjer för bröstcancersjukvård (Swedish National Guidelines for Breast Cancer Healthcare)2007 120923.
101. Porter PL, El-Bastawissi AY, Mandelson MT, Lin MG, Khalid N, Watney EA, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst. 1999;91(23):2020-8. Epub 1999/12/02.
102. Chu KC, Smart CR, Tarone RE. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. J Natl Cancer Inst. 1988;80(14):1125-32. Epub 1988/09/21.
103. Duffy SW, Tabar L, Chen HH, Holmqvist M, Yen MF, Abdsalah S, et al. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer. 2002;95(3):458-69. Epub 2002/09/05.
104. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002;359(9310):909-19. Epub 2002/03/29.
105. Tabar L, Vitak B, Chen HH, Yen MF, Duffy SW, Smith RA. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer. 2001;91(9):1724-31. Epub 2001/05/04.
106. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane database of systematic reviews. 2011(1):CD001877. Epub 2011/01/21.
107. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol. 2007;8(12):1129-38. Epub 2007/12/07.
108. Gotzsche PC. On the benefits and harms of screening for breast cancer. International journal of epidemiology. 2004;33(1):56-64; discussion 9-73. Epub 2004/04/13.
109. Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. N Engl J Med. 2010;363(13):1203-10. Epub 2010/09/24.
110. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer. 1989;63(1):181-7. Epub 1989/01/01.
111. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Senn HJ. Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. Seventh International Conference on Adjuvant Therapy of Primary Breast Cancer. J Clin Oncol. 2001;19(18):3817-27. Epub 2001/09/18.
112. Greenough RB. Varying degrees of malignancy in cancer of the breast. J Cancer Research. 1925(9):453.
113. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer. 1957;11(3):359-77. Epub 1957/09/01.
114. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19(5):403-10. Epub 1991/11/01.
115. Meyer JS, Alvarez C, Milikowski C, Olson N, Russo I, Russo J, et al. Breast carcinoma malignancy grading by Bloom-Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. Mod Pathol. 2005;18(8):1067-78. Epub 2005/05/28.
116. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. Lancet Oncol. 2010;11(2):174-83. Epub 2010/02/16.
117. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ, et al. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. Ann Oncol. 2009;20(8):1319-29. Epub 2009/06/19.
118. Ferno M, Stal O, Baldetorp B, Hatschek T, Kallstrom AC, Malmstrom P, et al. Results of two or five years of adjuvant tamoxifen correlated to steroid receptor

and S-phase levels. South Sweden Breast Cancer Group, and South-East Sweden Breast Cancer Group. Breast Cancer Res Treat. 2000;59(1):69-76. Epub 2000/04/07.
119. Stendahl M, Ryden L, Nordenskjold B, Jonsson PE, Landberg G, Jirstrom K. High progesterone receptor expression correlates to the effect of adjuvant tamoxifen in premenopausal breast cancer patients. Clin Cancer Res. 2006;12(15):4614-8. Epub 2006/08/11.
120. Mansour EG, Ravdin PM, Dressler L. Prognostic factors in early breast carcinoma. Cancer. 1994;74(1 Suppl):381-400. Epub 1994/07/01.
121. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235(4785):177-82. Epub 1987/01/09.
122. Dawood S, Broglio K, Buzdar AU, Hortobagyi GN, Giordano SH. Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: an institutional-based review. J Clin Oncol. 2010;28(1):92-8. Epub 2009/11/26.
123. Metzger-Filho O, Tutt A, de Azambuja E, Saini KS, Viale G, Loi S, et al. Dissecting the heterogeneity of triple-negative breast cancer. J Clin Oncol. 2012;30(15):1879-87. Epub 2012/03/29.
124. Carey L, Winer E, Viale G, Cameron D, Gianni L. Triple-negative breast cancer: disease entity or title of convenience? Nature reviews Clinical oncology. 2010;7(12):683-92. Epub 2010/09/30.
125. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52. Epub 2000/08/30.
126. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869-74. Epub 2001/09/13.
127. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. Lancet. 2011;378(9805):1812-23. Epub 2011/11/22.
128. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome biology. 2007;8(5):R76. Epub 2007/05/12.
129. Farmer P, Bonnefoi H, Becette V, Tubiana-Hulin M, Fumoleau P, Larsimont D, et al. Identification of molecular apocrine breast tumours by microarray analysis. Oncogene. 2005;24(29):4660-71. Epub 2005/05/18.
130. Doane AS, Danso M, Lal P, Donaton M, Zhang L, Hudis C, et al. An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen. Oncogene. 2006;25(28):3994-4008. Epub 2006/02/24.
131. Sorlie T. Molecular classification of breast tumors: toward improved diagnostics and treatments. Methods in molecular biology. 2007;360:91-114. Epub 2006/12/19.
132. Fisher B, Jeong JH, Anderson S, Bryant J, Fisher ER, Wolmark N. Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation. N Engl J Med. 2002;347(8):567-75. Epub 2002/08/23.
133. Giuliano AE, Hunt KK, Ballman KV, Beitsch PD, Whitworth PW, Blumencranz PW, et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. Jama. 2011;305(6):569-75. Epub 2011/02/10.
134. Orr RK. The impact of prophylactic axillary node dissection on breast cancer survival--a Bayesian meta-analysis. Annals of surgical oncology. 1999;6(1):109-16. Epub 1999/02/25.
135. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans E, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;366(9503):2087-106. Epub 2005/12/20.

136. Lind PA, Wennberg B, Gagliardi G, Fornander T. Pulmonary complications following different radiotherapy techniques for breast cancer, and the association to irradiated lung volume and dose. Breast Cancer Res Treat. 2001;68(3):199-210. Epub 2001/12/01.
137. Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365(9472):1687-717. Epub 2005/05/17.
138. Dowsett M, Cuzick J, Ingle J, Coates A, Forbes J, Bliss J, et al. Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen. J Clin Oncol. 2010;28(3):509-18. Epub 2009/12/02.
139. Mieog JS, van der Hage JA, van de Velde CJ. Preoperative chemotherapy for women with operable breast cancer. Cochrane database of systematic reviews. 2007(2):CD005002. Epub 2007/04/20.
140. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med. 2002;347(16):1233-41. Epub 2002/10/24.
141. Joensuu H, Lehtimaki T, Holli K, Elomaa L, Turpeenniemi-Hujanen T, Kataja V, et al. Risk for distant recurrence of breast cancer detected by mammography screening or other methods. Jama. 2004;292(9):1064-73. Epub 2004/09/02.
142. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst. 2005;97(16):1195-203. Epub 2005/08/18.
143. Ewertz M, Jensen MB, Gunnarsdottir KA, Hojris I, Jakobsen EH, Nielsen D, et al. Effect of obesity on prognosis after early-stage breast cancer. J Clin Oncol. 2011;29(1):25-31. Epub 2010/12/01.
144. Ballard-Barbash R, Friedenreich CM, Courneya KS, Siddiqi SM, McTiernan A, Alfano CM. Physical activity, biomarkers, and disease outcomes in cancer survivors: a systematic review. J Natl Cancer Inst. 2012;104(11):815-40. Epub 2012/05/10.
145. Sparano JA, Wang M, Zhao F, Stearns V, Martino S, Ligibel JA, et al. Race and hormone receptor-positive breast cancer outcomes in a randomized chemotherapy trial. J Natl Cancer Inst. 2012;104(5):406-14. Epub 2012/01/18.
146. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. Jama. 2006;295(21):2492-502. Epub 2006/06/08.
147. Rosenberg L, Thalib L, Adami HO, Hall P. Childbirth and breast cancer prognosis. Int J Cancer. 2004;111(5):772-6. Epub 2004/07/15.
148. Daling JR, Malone KE, Doody DR, Anderson BO, Porter PL. The relation of reproductive factors to mortality from breast cancer. Cancer Epidemiol Biomarkers Prev. 2002;11(3):235-41. Epub 2002/03/16.
149. Chlebowski RT, Anderson GL. Changing concepts: Menopausal hormone therapy and breast cancer. J Natl Cancer Inst. 2012;104(7):517-27. Epub 2012/03/20.
150. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer. 1976;37(5):2486-92. Epub 1976/05/01.
151. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. AJR Am J Roentgenol. 1976;126(6):1130-7. Epub 1976/06/01.
152. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. Phys Med Biol. 1994;39(10):1629-38. Epub 1994/10/01.
153. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med. 2007;356(3):227-36. Epub 2007/01/19.
154. Byrne C, Schairer C, Wolfe J, Parekh N, Salane M, Brinton LA, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst. 1995;87(21):1622-9. Epub 1995/11/01.

155. Gram IT, Funkhouser E, Tabar L. The Tabar classification of mammographic parenchymal patterns. European journal of radiology. 1997;24(2):131-6. Epub 1997/02/01.
156. Kerlikowske K, Grady D, Barclay J, Frankel SD, Ominsky SH, Sickles EA, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. J Natl Cancer Inst. 1998;90(23):1801-9. Epub 1998/12/05.
157. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. Breast Cancer Res. 2011;13(6):223. Epub 2011/11/26.
158. Heine JJ, Scott CG, Sellers TA, Brandt KR, Serie DJ, Wu FF, et al. A novel automated mammographic density measure and breast * cancer risk. J Natl Cancer Inst. 2012;104(13):1028-37. Epub 2012/07/05.
159. Li J, Szekely L, Eriksson L, Heddson B, Sundbom A, Czene K, et al. High-throughput mammographic density measurement: a tool for risk prediction of breast cancer. Breast Cancer Res. 2012;14(4):R114. Epub 2012/08/01.
160. Thijssen JH. Local biosynthesis and metabolism of oestrogens in the human breast. Maturitas. 2004;49(1):25-33. Epub 2004/09/08.
161. Haars G, van Noord PA, van Gils CH, Grobbee DE, Peeters PH. Measurements of breast density: no ratio for a ratio. Cancer Epidemiol Biomarkers Prev. 2005;14(11 Pt 1):2634-40. Epub 2005/11/15.
162. Stone J, Ding J, Warren RM, Duffy SW, Hopper JL. Using mammographic density to predict breast cancer risk: dense area or percentage dense area. Breast Cancer Res. 2010;12(6):R97. Epub 2010/11/23.
163. Stone J, Warren RM, Pinney E, Warwick J, Cuzick J. Determinants of percentage and area measures of mammographic density. Am J Epidemiol. 2009;170(12):1571-8. Epub 2009/11/17.
164. Guo YP, Martin LJ, Hanna W, Banerjee D, Miller N, Fishell E, et al. Growth factors and stromal matrix proteins associated with mammographic densities. Cancer Epidemiol Biomarkers Prev. 2001;10(3):243-8. Epub 2001/04/17.
165. Li T, Sun L, Miller N, Nicklee T, Woo J, Hulse-Smith L, et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. Cancer Epidemiol Biomarkers Prev. 2005;14(2):343-9. Epub 2005/03/01.
166. Alowami S, Troup S, Al-Haddad S, Kirkpatrick I, Watson PH. Mammographic density is related to stroma and stromal proteoglycan expression. Breast Cancer Res. 2003;5(5):R129-35. Epub 2003/08/21.
167. Fisher ER, Palekar A, Kim WS, Redmond C. The histopathology of mammographic patterns. Am J Clin Pathol. 1978;69(4):421-6. Epub 1978/04/01.
168. Ghosh K, Brandt KR, Reynolds C, Scott CG, Pankratz VS, Riehle DL, et al. Tissue composition of mammographically dense and non-dense breast tissue. Breast Cancer Res Treat. 2012;131(1):267-75. Epub 2011/08/31.
169. Wiseman BS, Werb Z. Stromal effects on mammary gland development and breast cancer. Science. 2002;296(5570):1046-9. Epub 2002/05/11.
170. Ghosh K, Hartmann LC, Reynolds C, Visscher DW, Brandt KR, Vierkant RA, et al. Association between mammographic density and age-related lobular involution of the breast. J Clin Oncol. 2010;28(13):2207-12. Epub 2010/03/31.
171. Ghosh K, Vachon CM, Pankratz VS, Vierkant RA, Anderson SS, Brandt KR, et al. Independent association of lobular involution and mammographic breast density with breast cancer risk. J Natl Cancer Inst. 2010;102(22):1716-23. Epub 2010/11/03.
172. Maskarinec G, Pagano I, Lurie G, Kolonel LN. A longitudinal investigation of mammographic density: the multiethnic cohort. Cancer Epidemiol Biomarkers Prev. 2006;15(4):732-9. Epub 2006/04/15.
173. Boyd N, Martin L, Stone J, Little L, Minkin S, Yaffe M. A longitudinal study of the effects of menopause on mammographic features. Cancer Epidemiol Biomarkers Prev. 2002;11(10 Pt 1):1048-53. Epub 2002/10/12.
174. Boyd NF, Lockwood GA, Martin LJ, Knight JA, Byng JW, Yaffe MJ, et al. Mammographic densities and breast cancer risk. Breast disease. 1998;10(3-4):113-26. Epub 2005/02/03.

175. Warren R. Hormones and mammographic breast density. Maturitas. 2004;49(1):67-78. Epub 2004/09/08.
176. Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S. Changes in breast density associated with initiation, discontinuation, and continuing use of hormone replacement therapy. Jama. 2001;285(2):171-6. Epub 2001/02/15.
177. Colacurci N, Fornaro F, De Franciscis P, Mele D, Palermo M, del Vecchio W. Effects of a short-term suspension of hormone replacement therapy on mammographic density. Fertility and sterility. 2001;76(3):451-5. Epub 2001/09/05.
178. Cuzick J, Warwick J, Pinney E, Duffy SW, Cawthorn S, Howell A, et al. Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: a nested case-control study. J Natl Cancer Inst. 2011;103(9):744-52. Epub 2011/04/13.
179. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. Breast Cancer Res. 2008;10(1):201. Epub 2008/01/30.
180. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst. 2007;99(15):1178-87. Epub 2007/07/27.
181. Boyd NF, Martin LJ, Sun L, Guo H, Chiarelli A, Hislop G, et al. Body size, mammographic density, and breast cancer risk. Cancer Epidemiol Biomarkers Prev. 2006;15(11):2086-92. Epub 2006/11/23.
182. Brisson J, Morrison AS, Kopans DB, Sadowsky NL, Kalisher L, Twaddle JA, et al. Height and weight, mammographic features of breast tissue, and breast cancer risk. Am J Epidemiol. 1984;119(3):371-81. Epub 1984/03/01.
183. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. Br J Cancer. 1998;78(9):1233-8. Epub 1998/11/20.
184. Boyd NF, Rommens JM, Vogt K, Lee V, Hopper JL, Yaffe MJ, et al. Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol. 2005;6(10):798-808. Epub 2005/10/04.
185. Rosner B, Colditz GA, Willett WC. Reproductive risk factors in a prospective study of breast cancer: the Nurses' Health Study. Am J Epidemiol. 1994;139(8):819-35. Epub 1994/04/15.
186. Rosner B, Colditz GA. Nurses' health study: log-incidence mathematical model of breast cancer incidence. J Natl Cancer Inst. 1996;88(6):359-64. Epub 1996/03/20.
187. Colditz GA, Rosner B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. Am J Epidemiol. 2000;152(10):950-64. Epub 2000/11/25.
188. Boyd NF, Dite GS, Stone J, Gunasekara A, English DR, McCredie MR, et al. Heritability of mammographic density, a risk factor for breast cancer. N Engl J Med. 2002;347(12):886-94. Epub 2002/09/20.
189. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). Cancer Causes Control. 2000;11(7):653-62. Epub 2000/09/08.
190. Varghese JS, Thompson DJ, Michailidou K, Lindstrom S, Turnbull C, Brown J, et al. Mammographic breast density and breast cancer: evidence of a shared genetic basis. Cancer Res. 2012;72(6):1478-84. Epub 2012/01/24.
191. Lindstrom S, Vachon CM, Li J, Varghese J, Thompson D, Warren R, et al. Common variants in ZNF365 are associated with both mammographic density and breast cancer risk. Nature genetics. 2011;43(3):185-7. Epub 2011/02/01.
192. Vachon CM, Scott CG, Fasching PA, Hall P, Tamimi RM, Li J, et al. Common breast cancer susceptibility variants in LSP1 and RAD51L1 are associated with mammographic density measures that predict breast cancer risk. Cancer Epidemiol Biomarkers Prev. 2012;21(7):1156-66. Epub 2012/03/29.
193. Odefrey F, Stone J, Gurrin LC, Byrnes GB, Apicella C, Dite GS, et al. Common genetic variants associated with breast cancer and mammographic density measures that predict disease. Cancer Res. 2010;70(4):1449-58. Epub 2010/02/11.

194. Helgesson O, Bengtsson C, Lapidus L, Merck C, Sparen P. Malignant disease observed in a cohort of women. A validation of Swedish Cancer Registry data. Scandinavian journal of social medicine. 1994;22(1):46-9. Epub 1994/03/01.
195. Nystrom L, Larsson LG, Rutqvist LE, Lindgren A, Lindqvist M, Ryden S, et al. Determination of cause of death among breast cancer cases in the Swedish randomized mammography screening trials. A comparison between official statistics and validation by an endpoint committee. Acta Oncol. 1995;34(2):145-52. Epub 1995/01/01.
196. Byrne C, Colditz GA, Willett WC, Speizer FE, Pollak M, Hankinson SE. Plasma insulin-like growth factor (IGF) I, IGF-binding protein 3, and mammographic density. Cancer Res. 2000;60(14):3744-8. Epub 2000/08/05.
197. Tamimi RM, Hankinson SE, Colditz GA, Byrne C. Endogenous sex hormone levels and mammographic density among postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2005;14(11 Pt 1):2641-7. Epub 2005/11/15.
198. Pawitan Y, Bjohle J, Amler L, Borg AL, Egyhazi S, Hall P, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res. 2005;7(6):R953-64. Epub 2005/11/11.
199. Livshits G, Malkin I, Williams FM, Hart DJ, Hakim A, Spector TD. Longitudinal study of variation in body mass index in middle-aged UK females. Age (Dordr). 2011. Epub 2011/08/20.
200. Magnusson C, Baron JA, Correia N, Bergstrom R, Adami HO, Persson I. Breast-cancer risk following long-term oestrogen- and oestrogen-progestin-replacement therapy. Int J Cancer. 1999;81(3):339-44. Epub 1999/04/21.
201. Calza S, Hall P, Auer G, Bjohle J, Klaar S, Kronenwett U, et al. Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. Breast Cancer Res. 2006;8(4):R34. Epub 2006/07/19.
202. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974;19:716-23.
203. Vachon CM, van Gils CH, Sellers TA, Ghosh K, Pruthi S, Brandt KR, et al. Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res. 2007;9(6):217. Epub 2008/01/15.
204. Narod SA. Breast cancer in young women. Nature reviews Clinical oncology. 2012;9(8):460-70. Epub 2012/06/27.
205. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst. 2000;92(13):1081-7. Epub 2000/07/06.
206. Aiello EJ, Buist DS, White E, Porter PL. Association between mammographic breast density and breast cancer tumor characteristics. Cancer Epidemiol Biomarkers Prev. 2005;14(3):662-8. Epub 2005/03/16.
207. Chiu SY, Duffy S, Yen AM, Tabar L, Smith RA, Chen HH. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. Cancer Epidemiol Biomarkers Prev. 2010;19(5):1219-28. Epub 2010/04/22.
208. Roubidoux MA, Bailey JE, Wray LA, Helvie MA. Invasive cancers detected after breast cancer screening yielded a negative result: relationship of mammographic density to tumor prognostic factors. Radiology. 2004;230(1):42-8. Epub 2003/12/26.
209. Sala E, Solomon L, Warren R, McCann J, Duffy S, Luben R, et al. Size, node status and grade of breast tumours: association with mammographic parenchymal patterns. Eur Radiol. 2000;10(1):157-61. Epub 2000/02/09.
210. Yaghjyan L, Colditz GA, Collins LC, Schnitt SJ, Rosner B, Vachon C, et al. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. J Natl Cancer Inst. 2011;103(15):1179-89. Epub 2011/07/29.
211. Jatoi I, Hilsenbeck SG, Clark GM, Osborne CK. Significance of axillary lymph node metastasis in primary breast cancer. J Clin Oncol. 1999;17(8):2334-40. Epub 1999/11/24.

212. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density: a hormonally responsive risk factor for breast cancer. J Br Menopause Soc. 2006;12(4):186-93. Epub 2006/12/21.
213. Phipps AI, Buist DS, Malone KE, Barlow WE, Porter PL, Kerlikowske K, et al. Breast Density, Body Mass Index, and Risk of Tumor Marker-Defined Subtypes of Breast Cancer. Ann Epidemiol. 2012. Epub 2012/03/01.
214. Byrne C, Schairer C, Brinton LA, Wolfe J, Parekh N, Salane M, et al. Effects of mammographic density and benign breast disease on breast cancer risk (United States). Cancer Causes Control. 2001;12(2):103-10. Epub 2001/03/15.
215. Brekelmans CT, Peeters PH, Faber JA, Deurenberg JJ, Collette HJ. The epidemiological profile of women with an interval cancer in the DOM screening programme. Breast Cancer Res Treat. 1994;30(3):223-32. Epub 1994/01/01.
216. Kavanagh AM, Mitchell H, Giles GG. Hormone replacement therapy and accuracy of mammographic screening. Lancet. 2000;355(9200):270-4. Epub 2000/02/16.
217. Heusinger K, Jud SM, Haberle L, Hack C, Adamietz BR, Meier-Meitinger M, et al. Association of mammographic density with hormone receptors in invasive breast cancers - results from a case-only study. Int J Cancer. 2012. Epub 2012/03/07.
218. Arora N, King TA, Jacks LM, Stempel MM, Patil S, Morris E, et al. Impact of breast density on the presenting features of malignancy. Ann Surg Oncol. 2010;17 Suppl 3:211-8. Epub 2010/10/01.
219. Ma H, Luo J, Press MF, Wang Y, Bernstein L, Ursin G. Is there a difference in the association between percent mammographic density and subtypes of breast cancer? Luminal A and triple-negative breast cancer. Cancer Epidemiol Biomarkers Prev. 2009;18(2):479-85. Epub 2009/02/05.
220. Arora N, King TA, Jacks LM, Stempel MM, Patil S, Morris E, et al. Impact of breast density on the presenting features of malignancy. Annals of surgical oncology. 2010;17 Suppl 3:211-8. Epub 2010/10/01.
221. Phipps AI, Buist DS, Malone KE, Barlow WE, Porter PL, Kerlikowske K, et al. Breast density, body mass index, and risk of tumor marker-defined subtypes of breast cancer. Ann Epidemiol. 2012;22(5):340-8. Epub 2012/03/01.
222. Eriksson L, Czene K, Rosenberg L, Humphreys K, Hall P. The influence of mammographic density on breast tumor characteristics. Breast Cancer Res Treat. 2012;134(2):859-66. Epub 2012/06/20.
223. Heusinger K, Jud SM, Haberle L, Hack CC, Adamietz BR, Meier-Meitinger M, et al. Association of mammographic density with hormone receptors in invasive breast cancers: Results from a case-only study. Int J Cancer. 2012. Epub 2012/03/07.
224. Heusinger K, Jud SM, Haberle L, Hack CC, Fasching PA, Meier-Meitinger M, et al. Association of mammographic density with the proliferation marker Ki-67 in a cohort of patients with invasive breast cancer. Breast Cancer Res Treat. 2012. Epub 2012/09/01.
225. Gierach GL, Ichikawa L, Kerlikowske K, Brinton LA, Farhat GN, Vacek PM, et al. Relationship between mammographic density and breast cancer death in the breast cancer surveillance consortium. J Natl Cancer Inst. 2012;104(16):1218-27. Epub 2012/08/23.
226. Olsen AH, Bihrmann K, Jensen MB, Vejborg I, Lynge E. Breast density and outcome of mammography screening: a cohort study. Br J Cancer. 2009;100(7):1205-8. Epub 2009/03/19.
227. Cil T, Fishell E, Hanna W, Sun P, Rawlinson E, Narod SA, et al. Mammographic density and the risk of breast cancer recurrence after breast-conserving surgery. Cancer. 2009;115(24):5780-7. Epub 2009/11/11.
228. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869-74. Epub 2001/09/13.
229. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a

population-based study. Proc Natl Acad Sci U S A. 2003;100(18):10393-8. Epub 2003/08/15.
230. Park CC, Rembert J, Chew K, Moore D, Kerlikowske K. High mammographic breast density is independent predictor of local but not distant recurrence after lumpectomy and radiotherapy for invasive breast cancer. Int J Radiat Oncol Biol Phys. 2009;73(1):75-9. Epub 2008/08/12.
231. de Azambuja E, Cardoso F, de Castro G, Jr., Colozza M, Mano MS, Durbecq V, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. Br J Cancer. 2007;96(10):1504-13. Epub 2007/04/25.
232. Kim MY, Oskarsson T, Acharyya S, Nguyen DX, Zhang XH, Norton L, et al. Tumor self-seeding by circulating cancer cells. Cell. 2009;139(7):1315-26. Epub 2010/01/13.
233. Comen E, Norton L, Massague J. Clinical implications of cancer self-seeding. Nature reviews Clinical oncology. 2011;8(6):369-77. Epub 2011/04/28.