From the Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

Determinants of interval cancer and tumor size among breast cancer screening participants

Fredrik Strand

Stockholm 2018

Cover image by the author, pseudo-colored mammographic image of a breast cancer.

Published by Karolinska Institutet.

ISBN 978-91-7831-004-3

Printed by Eprint AB 2018

INSTITUTIONEN FÖR MEDICINSK EPIDEMIOLOGI OCH BIOSTATISTIK

Determinants of interval cancer and tumor size among breast cancer screening participants

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i hörsal Petrén, Nobels Väg 12 B, Karolinska Institutet, Solna

Fredagen den 8 juni 2018, kl 09:00

av

Fredrik Strand
MD, MSc

Huvudhandledare

Professor Kamila Czene
Karolinska Institutet
Institutionen för Medicinsk Epidemiologi och Biostatistik

Bihandledare

Universitetslektor Keith Humphreys
Karolinska Institutet
Institutionen för Medicinsk Epidemiologi och Biostatistik

Professor Per Hall
Karolinska Institutet
Institutionen för Medicinsk Epidemiologi och Biostatistik

Docent Edward Azavedo
Karolinska Institutet
Institutionen för Molekylär Medicin och Kirurgi

Fakultetsopponent

Professor Emily Conant
University of Pennsylvania, U.S.A.
Department of Radiology

Betygsnämnd

Docent Torkel Brismar
Karolinska Institutet
Institutionen för Klinisk vetenskap, Intervention och Teknik

Professor Olof Akre
Karolinska Institutet
Institutionen för Molekylär Medicin och Kirurgi

Docent Henrik Lindman
Uppsala Universitet
Institutionen för Immunologi, Genetik och Patologi

Stockholm, 2018

to Teodor and Ingrid

Abstract

Breast cancer is the most common cancer of women in Sweden and globally. In the more affluent countries, mammography screening has been in place for a few decades and has successfully reduced mortality. However, there is increasing interest in enhancing the impact of screening by going from the current age-based screening system to a risk-based system. There are two risk components that must be taken into account – the underlying breast cancer risk and the risk of delayed detection. Mammographic density, the amount of dense tissue in the breast, has been shown to be a risk factor for both. In this thesis, my aim was to identify novel determinants of delayed breast cancer detection by studying observed cases of interval cancer or large cancer at diagnosis. The potential risk factors for delayed detection were based on negative mammograms and other data that can be determined before diagnosis. Study I to III, were based on a retrospective case-only population, while Study IV was based on a prospective cohort.

In **Study I**, we developed an estimate of the longitudinal fluctuation in mammographic percent density between screenings. Based on our results, we concluded that women that were subsequently diagnosed with interval cancer had higher density fluctuations than women with screen-detected cancer.

In **Study II**, we went beyond density and examined 32 other image features which were computer-extracted from digitized mammograms. We identified two novel features that were associated with an increased risk of interval cancer compared to screen-detected cancer. One feature seemed to be related to the shape of the entire dense area, being flat rather than round increased the risk of interval cancer, possibly due to making clinical detection easier. The other feature seemed to be related to whether the density was more concentrated or instead was interspersed with fatty streaks. When density was more concentrated, the risk of interval cancer increased, possibly by making mammographic detection more difficult.

In **Study III**, we determined risk factors for the cancer diagnosis being delayed until the cancer had reached a size larger than 2 cm. High density and high body mass index (BMI) were already known risk factors in general. Our aim was to understand if different factors were involved depending on the detection mode, screen-detection or interval cancer detection. We found that high BMI increased the risk of large cancer markedly among interval cancers and somewhat among screen-detected cancers. High density was associated with large cancer only among screen-detected cases. In survival analysis, we showed that high BMI increased the risk of disease progression, but only among women with interval cancer.

In **Study IV**, we found that the localized density category at the site of the subsequent cancer was often different compared to the overall density. We examined the effect of high localized density, independent of overall density, and found that it was strongly associated with large cancer at diagnosis. In addition, it was associated with interval cancer among the less aggressive node-negative cases. It remains to be elucidated whether this effect is purely due to visual masking or also due to an association with biological characteristics of the tumor microenvironment.

In conclusion, we have identified several novel determinants of delayed breast cancer detection, which could be further validated in trials of risk-stratified screening.

List of Publications

I. **Fredrik Strand**, Keith Humphreys, Mikael Eriksson, Jingmei Li, Therese ML Andersson, Sven Törnberg, Edward Azavedo, John Shepherd, Per Hall, Kamila Czene
Longitudinal fluctuation in mammographic percent density differentiates between interval and screen-detected breast cancer.
International Journal of Cancer. 2017;140(1):34-40.

II. **Fredrik Strand**, Keith Humphreys, Abbas Cheddad, Sven Törnberg, Edward Azavedo, John Shepherd, Per Hall, Kamila Czene
Novel mammographic image features differentiate between interval and screen-detected breast cancer: a case-case study.
Breast Cancer Research. 2016;18(1):100.

III. **Fredrik Strand**, Keith Humphreys, Johanna Holm, Mikael Eriksson, Sven Törnberg, Per Hall, Kamila Czene
Long-term prognostic implications of risk factors associated with tumor size: a case study of women regularly attending screening.
Breast Cancer Research. 2018; 20(1); 31.

IV. **Fredrik Strand**, Edward Azavedo, Roxanna Hellgren, Keith Humphreys, Mikael Eriksson, John Shepherd, Per Hall, Kamila Czene
The effect of overall and localized mammographic density on breast cancer detection: a prospective cohort study.
Manuscript

Contents

Abbreviations 1
1. Background 3
1.1 Breast Biology 3
1.1.1 Development 3
1.1.2 Anatomy 3
1.1.3 Changes over Time 4
1.2 Breast Cancer 4
1.2.1 Cancer Development 4
1.2.2 Epidemiology 5
1.2.3 Risk Factors 5
1.2.4 Staging 6
1.2.5 Histopathology and Molecular classifications 6
1.2.6 Treatment and Prognosis 7
1.3 Breast Imaging 7
1.3.1 Mammography 7
1.3.2 Mammographic Density 9
1.3.3 Supplemental Imaging Methods 11
1.3.4 Screening 11
1.4 Interval Cancer 14
1.4.1 Definitions 14
1.4.2 Prevention 15
1.4.3 Determinants 15
1.4.4 Prognosis 16
1.5 Tumor Size 16
1.5.1 Definitions 16
1.5.2 Prevention 17
1.5.3 Determinants 17
1.5.4 Prognosis 17
2. Aims and Hypotheses 18
3. Patients and Methods 19
3.1 Underlying Study Populations 19
3.1.1 LIBRO-1 19
3.1.2 CAHRES 19
3.1.3 KARMA 19
3.2 Data 20
3.2.1 Register Data (all Studies) 20
3.2.2 Cancer Detection Mode (all Studies) 20
3.2.3 Mammograms and Density (all Studies) 20

3.2.4 Image Feature Extraction and Selection (Study II) 21
3.2.5 Tumor Characteristics (all Studies) 21
3.3 Epidemiological Study Design 21
3.3.1 Cohort Study (part of Study III and IV) 21
3.3.2 Case-control Study (all Studies) 22
3.4 Statistical methods 22
3.4.1 Linear Regression (Study II and III) 22
3.4.2 Logistic Regression (all Studies) 23
3.4.3 Mixed Effects Model – Mammographic Density Fluctuation (Study I) 23
3.4.4 Cox Regression (Study III) 24
4. Results 24
4.1 Study I 24
4.2 Study II 25
4.3 Study III 26
4.4 Study IV 28
5. Discussion 29
5.1 Study I 29
5.2 Study II 29
5.3 Study III 29
5.4 Study IV 30
6. Methodological Considerations 31
6.1 Bias and Confounding 31
6.2 Study I 32
6.3 Study II 33
6.4 Study III 33
6.5 Study IV 34
7. Ethical Considerations 35
8. Concluding Remarks 36
9. Future Perspectives 37
Svensk sammanfattning (abstract in Swedish) 38
Acknowledgements 39
References 41

Abbreviations

95%CI	95 percent confidence interval
BI-RADS density	A four-category visual classification of mammographic density issued by the American College of Radiology, where 'A' is the least and 'D' the most dense
BMI	Body Mass Index
BRCA1/2	Breast Cancer Susceptibility Gene 1 or 2
CAHRES	Cancer and Hormone Replacement Study – one of my study populations
cBIRADS	Computer-generated score mimicking the BI-RADS density classification
Dnr	Reference number in public archives; spelled out in Swedish: 'diarienummer'
GWAS	Genome-Wide Association Studies
HER2	Human Epidermal growth factor Receptor 2
HRT	Hormone Replacement Therapy (mainly for menopausal symptoms)
IC	Interval Cancer
KARMA	Karolinska Mammography cohort – one of my study populations
LIBRO-1	Linné-Bröst study 1 – one of my study populations
MRI	Magnetic Resonance Imaging
OR	Odds Ratio
p	the Probability that an observation might be explained by chance alone
PD	Percent mammographic density = The proportion of dense tissue out of total breast tissue as estimated by analysing mammograms
SDC	Screen-Detected Cancer
SNP	Single Nucleotide Polymorphisms
TDLU	Terminal Duct Lobular Unit
TNM	A staging system for malignant tumors

"Pick a star on the dark horizon
And follow the light"

Regina Spektor

1. Background

1.1 Breast Biology

1.1.1 Development

The development of the breast is identical in males and females during the first two years of life. During puberty, in females, a new development stage is initiated as the mammary tissue beneath the areola starts growing and a two-layer structure of the epithelium develops (1). This process, between age 8 and 18, is gradually controlled by the hypothalamus which in turn stimulates the anterior pituitary gland that increases the levels of follicle-stimulating hormone and luteinizing hormone. Follicle-stimulating hormone activates ovarian follicles that secrete estrogen. In the breast, estrogen stimulates the formation of ducts, connective tissue and blood vessels. Later, progesterone exposure contributes to the formation of lobular, potentially milk-producing, structures (2).

1.1.2 Anatomy

The dorsal part of the breast rests on the fascia of the underlying muscles, e.g., the pectoral muscle. Ductal structures can appear in a large area of the thorax, from the axilla down to the sub-costal region, and from the fascia to very close to the skin border.

The breast gland contains glandular, stromal and fatty tissue. The proportions of each tissue type vary greatly between women. The stroma consists of connective tissue, mainly collagen, that provides a 'soft skeleton' for the breast maintaining its shape and internal structure. The term 'fibroglandular tissue' is often used to refer to the glandular and the stromal tissue together.

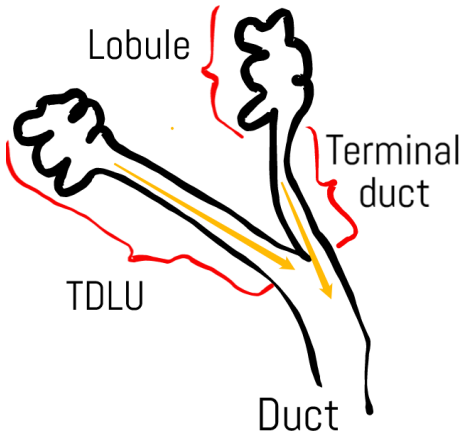

Figure 1. Anatomy of the fundamental glandular structures of the breast. TDLU = Terminal Duct Lobular Unit.

The glandular tissue consists of the lobules, potentially milk-producing, and the ducts, leading to the nipple. The lining of the ducts is composed of an inner luminal layer of epithelial cells, and an outer basal layer of myoepithelial cells (3). These are enclosed by a surrounding basement membrane. The nipple contains around 10 openings, or terminal ducts, each one connected to

a lactiferous sinus that receives a lobar collecting duct. The lobar ducts are formed by the merger of ducts from many lobules (Figure 1). The lobules and their connecting terminal duct are together called TDLU (terminal duct lobular unit). The TDLU is the likely starting point for the most common form of breast cancer, ductal carcinoma (4).

In adults, lobules can appear in different stages of development, from bud to complete differentiation. The earlier differentiation stages have shown to be more vulnerable to carcinogenic insults compared to the more differentiated stages (2).

1.1.3 Changes over Time

Menstrual Cycle

The menstrual cycle is characterized by periodic changes of estrogen and progesterone hormone serum levels. The proportion of proliferation and apoptosis in the lobules of the breast varies with the stage of the menstrual cycle (5). In addition, it has been shown that certain collagen types in the stroma of the breast show a similar periodic variation, most pronounced near the basal layer and the ductules (6). A study of magnetic resonance imaging (MRI) by Fowler et al. (7) showed that the increased breast volume in the luteal phase was mainly caused by water retention and to lesser extent by other stromal and epithelial changes.

Pregnancy

During pregnancy, there is a hormone-induced increase of all parenchymal components, ductal and lobular (2). A histological study by Russo et al. showed that there is a larger proportion of highly differentiated lobules, less vulnerable to carcinogenic insults, in the breasts of parous women compared to non-parous women (8).

Menopause

Menopause is initiated by atresia of the ovarian follicles. Declining ovarian hormone levels, including estrogen, leads to involution of the breast. Ducts remain but lobules shrink. Stromal tissue regresses and is replaced by fat tissue (2). After involution, in the remaining smaller parenchyma, there can be a high proportion of less differentiated lobules in the breasts of both parous and non-parous women, but it is unclear whether these confer the same risk of breast cancer regardless of parity status (8). After menopause, the enzyme aromatase is involved in the local production of estrogen from peripheral fatty tissue which explains why the postmenopausal estrogen levels can be much higher in the breast tissue than in the plasma (9).

1.2 Breast Cancer

1.2.1 Cancer Development

The pathway towards breast cancer is believed to start with carcinogenic insults to the epithelial cells resulting in hyperplasia of benign atypical cells, which later can become malignant non-invasive cells and, finally, malignant invasive cells. Some mutations of the cancer genome confer selective advantages to the host cell in this local evolutionary process, and are thus called 'driver mutations', while other mutations are 'passenger mutations' (10). Nik-Zainal et al. studied the genome of 21 breast cancers and reconstructed their genomic history. They showed that step-wise mutational processes evolve across the lifespan of a cancer, and that there is large diversity in the cellular genomes within the same tumor (11). During the evolution of the genome leading

to malignancy, the phenotype progressively acquire the so-called 'hallmarks of cancer' proposed by Hanahan and Weinberg: Sustaining proliferative signalling, Evading growth suppressors, Activating invasion and metastasis, Enabling replicative immortality, Inducing angiogenesis, and Resisting cell death (12). Malignant cells that have not yet acquired the invasive capability are called 'in situ cancer'.

The stroma, or connective tissue, is involved in cancer development through paracrine and mechanical interactions with the epithelial cells (13). Stromal changes in reorganization, cell types and gene expression as well as signalling cascades seem to affect tumor progression and patient survival.

1.2.2 Epidemiology

For women, breast cancer is the most common cancer globally (14). In Sweden, during 2016, there were 7558 women who were diagnosed with breast cancer, of which 7000 were first-time cancers (15). As can be seen in Figure 2 below, there has been a long-term trend of increasing incidence and decreasing mortality. Based on a similar mortality decrease in the U.S., Plevritis et al. estimated that 63% was attributable to improved treatment and 37% to improved screening (16).

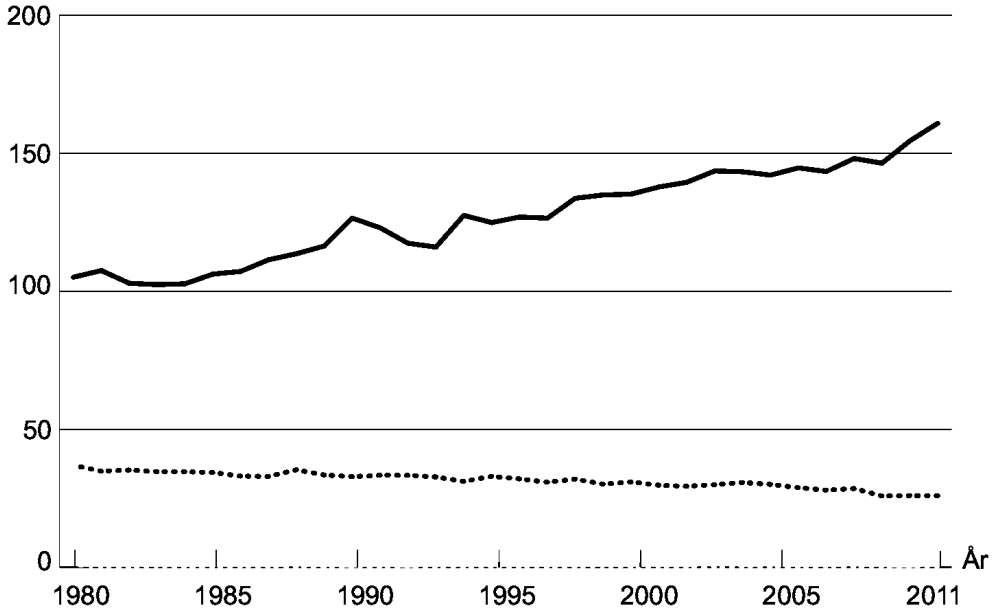

Figure 2. Swedish breast cancer statistics. The solid line shows breast cancer incidence per 100,000 women. The dashed line shows age-standardized breast cancer deaths per 100,000 women. *Source*: National Board of Health and Welfare.

1.2.3 Risk Factors

Breast cancer is a complex disease that have both genetic and non-genetic causes. The heritable component has been estimated to be between 25 and 27% in Scandinavia (17). The large environmental component is nicely illustrated by a study comparing the Japanese domestic population with the Japanese emigrants living in the San Francisco Bay Area in the USA. For the 35-64 age group, the age-adjusted incidence rates were 32, 94 and 116 for domestic, first-

generation emigrants and second-generation emigrants, compared to 179 for Caucasian women in the Bay Area (18). In general, the risk is about three times higher in more developed countries compared to less developed ones (14).

Female sex and increasing age are the strongest risk factors for breast cancer. After menopause the age-related risk increase slows down (19). Many breast cancer risk factors are linked to estrogen exposure(20). Risk factors associated with high endogenous estrogen exposure are: younger age at menarche, none or low number of childbirths, and older age at menopause (21). Risk factors associated with high exogenous estrogen exposure are: oral contraception pills (22) and hormone replacement therapy (HRT) for menopausal symptoms (23). Among lifestyle factors, breast cancer risk is increased by alcohol consumption and decreased by physical activity (24, 25). Having a diagnosis of benign proliferative breast lesions increases the risk (26). Body mass index (BMI) plays a dual role, high BMI increases the risk for post-menopausal women, but seems to decrease the risk for pre-menopausal women (27). Last, but not least, a high mammographic density is an important risk factor for breast cancer (28, 29). Mammographic density will be discussed in more detail in the 'Breast Imaging' section.

In terms of genetic risk factors, there are two mutations, BRCA1 and BRCA2, that incur a very high risk of breast cancer to the carriers, with a cumulative risk of developing breast cancer of 65% and 45%, respectively, by age 70 (30). Through genome-wide association studies (GWAS), more than 180 single nucleotide polymorphisms (SNP) associated with smaller risk increases have been identified (31, 32). It has been estimated that there may be around 1000 loci remaining to be identified (33).

1.2.4 Staging

Stage is one of the strongest prognostic predictors for breast cancer (34). Stage is defined based on the TNM system in which T concerns the primary tumor, N regional spread, and M distant metastasis (35). As a few examples, T1 is a tumor 2 cm or less in the greatest dimension, while T2 is a tumor more than 2 cm but nor more than 5 cm in the greatest dimension. Nearly all tumors have T stage different from T0, while only some have N stage different from N0 (between 32 and 35 % in the study above had lymph node metastasis) and fewer have M stage different from M0. Thus, for the majority of patients with N0 and M0, it is the T stage, i.e., the tumor size (and local invasion) that differentiates the stages and consequently the prognoses.

1.2.5 Histopathology and Molecular classifications

In addition to stage, for therapeutic guidance, breast cancers are often classified into groups based on histological origin or molecular subtypes as described below.

Histopathology

Histological analysis, i.e., microscopy of tissue specimens, shows that the most common histological types of invasive breast cancer are ductal carcinoma and lobular carcinoma. In a study by Martinez et al, there was around 80% ductal carcinoma and 15% lobular carcinoma (36). The lobular carcinomas are harder to identify on mammography, and are generally larger at diagnosis (37). Lobular carcinoma has a better prognosis than ductal carcinoma, and is often homogenous of low nuclear grade with limited desmoplastic reaction in the surrounding breast tissue (38). Absence of such dense fibrous reaction makes them harder to detect both clinically and mammographically.

Molecular Subtype

The molecular subtype has recently become an important consideration when assigning appropriate oncological treatment for each woman with breast cancer. Based on a gene expression clustering technique Sorlie et al. defined four molecular subtypes of breast cancer: 'Luminal A', 'Luminal B', 'HER2-overexpressing', and 'Basal-like' (39). The Luminal subtypes have an estrogen-driven gene expression. The HER2-overexpressing tumors have estrogen receptor negative status and instead has an amplified expression of the human epidermal growth factor receptor 2 (HER2). Finally, the basal-like tumors are negative for both estrogen receptor and HER2 receptor expression. Since not all hospitals have access to gene expression analysis, Cheang et al. defined an algorithm for how to use results from immunohistochemistry staining of receptors to assign proxy molecular subtypes (40). This latter approach was used in my studies, and the exact algorithm used has been detailed under the Patient and Methods section about 'Tumor Characteristics'.

1.2.6 Treatment and Prognosis

The prognosis for breast cancer patients is very good compared to most other malignant diseases, having an overall 5-year survival of around 90% in the Nordic countries (41). Primary breast cancer, except at the most advanced stage, are surgically excised either by a total mastectomy removing the entire breast gland or by partial mastectomy removing the identified tumor with a small margin of normal tissue (41). Depending on patient tolerability, type of mastectomy, and tumor characteristics including molecular subtype, the appropriate oncological treatment is chosen. Treatment options include chemotherapy, anti-hormonal drugs and radiotherapy. It has become increasingly popular to use neoadjuvant therapy, i.e., delaying surgery and instead administering oncological treatment to examine how the tumor responds while in the breast. To determine the effect of neoadjuvant treatment, the tumor is repeatedly examined by clinical palpation, by histopathology, and by various radiological modalities further described in the 'Breast Imaging' section.

1.3 Breast Imaging

1.3.1 Mammography

The mammographic image is formed by having an X-ray emitting tube on one side of the breast and a detector on the other. X-rays are electromagnetic radiation with a relatively high frequency and energy, which can pass through human tissue. For mammography, the X-ray photon energy is around 20 keV, relatively low compared to skeletal and chest x-ray imaging, which serves to maximize the contrast between an invasive carcinoma and adipose tissue (42). For the mammographic examination, a radiographer positions the breast between two plates, one compression plate transparent to X-rays and a larger plate containing the detector. Some X-ray photons pass unhindered through the tissue while others interact with the atoms of the tissue resulting in a decrease, or attenuation, of the x-rays. Due to a lower density and fewer electrons per atom, fat attenuates less of the X-rays and is thus 'non-dense' and depicted as dark in the mammogram. Fibroglandular tissue is 'dense' and depicted as bright pixels. Tumors are also dense and appear bright on a mammogram (43, 44).

Mammography was for a long time based on analog film-screen technology but has during the last decades shifted to digital technology. In analog mammography, the detector consisted of an

intensifying screen converting x-rays to a directly proportional amount of light reaching a film, which was then chemically developed. Film had a rather narrow dynamic range, and once a film had been exposed the contrast characteristics were fixed. An automatic exposure control terminated the radiation exposure once enough photons had penetrated the tissue of interest to reach a pre-set intensity. Nowadays, in digital mammography, the detector is based on semi-conductor technology and has a very large dynamic range allowing adjustment of image contrast and brightness in post-processing. Differences in the automatic exposure control complicates comparisons of automated density measurements between analog and digital mammograms(42).

With the introduction of digital mammography, the dose could be lowered by 25 to 30 percent (42). The diagnostic performance was similar to film mammography overall, but better for younger women with more dense breasts (45). In Stockholm, digital mammography was gradually introduced over a few years and the transition was completed by May 2008.

Normal mammograms of two women of the same age can look dramatically different in terms of the volume, organization and pattern of the dense and non-dense tissue. In addition, the appearance of breast cancer varies widely; examples include certain types of calcifications, architectural distortion, indistinct and spiculated masses (Figure 3). To increase the accuracy of mammographic interpretation two views are acquired per breast, 'cranio-caudal' view in which the breast is compressed horizontally and 'mediolateral oblique' (MLO) view in which the breast is compressed along an oblique line passing through the nipple towards the axillary region to follow the pectoral muscle and most of the glandular tissue.

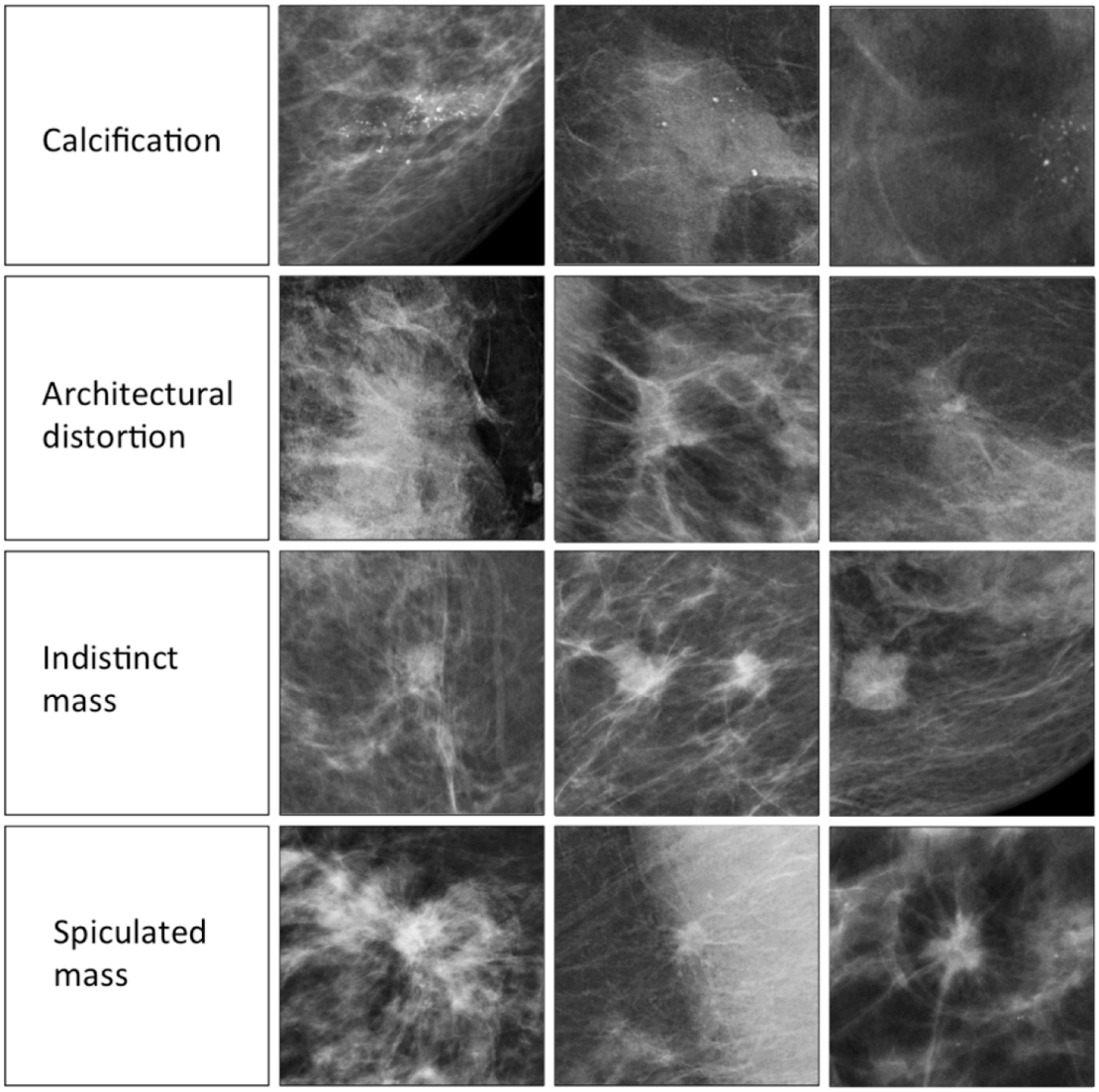

Figure 3. Examples of tumor appearances at mammography. In addition to the four categories above another potential tumor sign are 'asymmetrical densities' in which there are dense areas in one breast but not in the corresponding location in the other one.

1.3.2 Mammographic Density

The amount of mammographic density, the brighter pixels in the image, is an important risk factor for breast cancer and also for being diagnosed with an interval cancer (IC; defined in section 'Interval Cancer' below) compared to a screen-detected cancer (SDC) (46-49). In addition to being a primary risk factor, mammographic density seems to mediate other risk factors, i.e., other risk factors affect density which in turn affects the risk of breast cancer. Density is a partial mediator for nulliparity, age at first birth, hormone therapy and having breast biopsies (50, 51). Related to this, it was determined that mammographic density mediates some of the non-genetic geographical differences in breast cancer incidence (52).

In 1976, Wolfe first proposed the idea that the mammographic appearance of the healthy breast categorized into four different pattern groups could be associated with different levels of breast cancer risk (28). An alternative pattern categorization was proposed by Tabar et al. (53) Later,

Boyd et al. developed a semi-automated method, called 'Cumulus', to quantify mammographic density as a percentage of bright pixels out of the total breast area. This measure is called 'percent density' (PD) (29). Cumulus requires manual delineation of the relevant breast area in each mammogram which limits it use in large research volumes.

A fully automated density calculation method based on the ImageJ software was developed in our group with the aim of mimicking Cumulus (54, 55). It has later been refined under the name of Stratus (56). Other methods, e.g., Volpara™ (57), Libra, Quantra and Single X-ray Absorptiometry, estimate *volumetric* density, i.e., PD per cm^3. In contrast to these purely quantitative methods, the American College of Radiology has developed a widely used classification system based on visual assessment by four categories that also takes a qualitative aspect of detection sensitivity into account. It is often referred to as BI-RADS density, which stands for Breast Imaging-Reporting And Data System and includes different classification systems for density and for suspicious lesions. The density definitions according to the fifth edition of the BI-RADS guidelines are as follows from least to most dense (58):

A. The breasts are almost entirely fatty

B. There are scattered areas of fibroglandular density

C. The breasts are heterogeneously dense, which may obscure small masses

D. The breasts are extremely dense, which lowers the sensitivity of mammography

In a study by Li et al. [11] it was shown that high PD was associated with higher total nuclear area, both for epithelial and for non-epithelial cells, a higher proportion of collagen, and a larger area of glandular structures (59). When they analysed the relative amounts of each tissue type, collagen was the most abundant and explained 29% of the between-individual variation in PD.

Cross-sectional studies of mammographic density have shown that older women have lower mammographic density than younger women (60, 61). Part of the apparent mammographic density decrease with age may be an effect of later birth cohorts having higher age-adjusted density compared to earlier ones, as determined in a Danish study (62). Nevertheless, longitudinal studies have confirmed that there is a decreasing trend of mammographic density with increasing age - at least between age 40 and 65 (63, 64). That menopause has a density decreasing effect, in addition to what is explained by increasing age, was demonstrated in a study comparing those who transitioned from pre- to postmenopausal status with age-matched women who remained premenopausal (65).

In addition to long-term trends of decrease, there has been some evidence of short-term fluctuations in mammographic density and the corresponding fibroglandular tissue (66-68). The study by Ursin et al. showed large individual differences in the magnitude of fluctuation. Possibly, such fluctuations in density might be related to the periodic changes in endogenous hormones as suggested by a study of urinary estradiol metabolites in the ovulatory phase (69). The previously mentioned MRI study of eight women showed that the parenchymal volume on average increased by nearly 40% in total, and water content by 25%, over the menstrual cycle (7). Examining temporal, or longitudinal, fluctuations in density in relation to IC was the focus of my Study I.

A study comparing five different density measurement techniques showed that a visual-analog scale had a stronger association with breast cancer risk than the computer-based fully automated methods (70). In a study by Kerlikowske et al. it was shown that combining a risk model that include a visual density score with computer-based measures of density increased the ability to identify high-risk women (71). Applying a visual method in large retrospective research studies is often not feasible, but the finding shows that there are risk-relevant aspects of the image in addition to computer-calculated PD measure. Using computer programs to extract features of the parenchymal texture, in addition to density, has been gaining ground during recent years (72-76). The focus of my Study II was to identify computer-extracted features of the dense, parenchymal, tissue that would indicate an increased risk of IC compared to SDC.

1.3.3 Supplemental Imaging Methods

The sensitivity of mammography in detecting breast cancer is markedly reduced for women with high mammographic density. In a large study by Carney et al., it was shown that the sensitivity for detecting breast cancer was 63% in extremely dense breasts and 87% in the almost fatty breasts (77).

Recently, a variant of mammography has been introduced into clinical practice – digital breast tomosynthesis – sometimes called '3D mammography' (78). Digital breast tomosynthesis requires an adaptation of the regular mammography equipment. By having the detector move in an arc, images are acquired from various angles and the three-dimensional structure of the breast can be better appreciated. Tomosynthesis has a higher sensitivity compared to regular mammography, but its effect on recall rate is less clear (79-81). Reassuring for the ability to use mammographic density as a risk factor when transitioning to tomosynthesis, it has been shown that density measured from tomosynthesis images is highly correlated with that measured from regular mammographic images (82).

In addition, there are two non-X-ray imaging modalities that are routinely used for breast examinations: ultrasound and MRI. The basis for ultrasound images are the sound reflection characteristics of different tissues in the breast. It is often performed manually by the radiologist, even if semi-automated equipment also exists. The basis for MRI emerges from the interaction between radiofrequency signals and the characteristic magnetic properties of different tissues in the breast. For MRI, intravenous contrast media is commonly administered to examine the uptake and temporal dynamics of potentially malignant lesions. It has been shown that ultrasound plus mammography has a higher sensitivity than mammography alone; it has also been shown that the sensitivity of MRI is higher than mammography alone, and higher than mammography plus ultrasound (83-85).

1.3.4 Screening

Screening mammography is effective in reducing breast cancer mortality (86-89). The current national recommendation in Sweden stipulates that women should be invited for screening starting at age 40 and ending at age 74. All women fulfilling the age criteria are invited to screening, and they will continue to receive invitations whether they choose to attend or not. The national recommendation further stipulates that the time interval between invitations for screening should be between 18 and 24 months, with the shorter time interval suggested for

younger women. There is some scientific support for having more frequent screening between 40 to 49 years, but not for differentiating between women 50 to 74 years old (90-92).

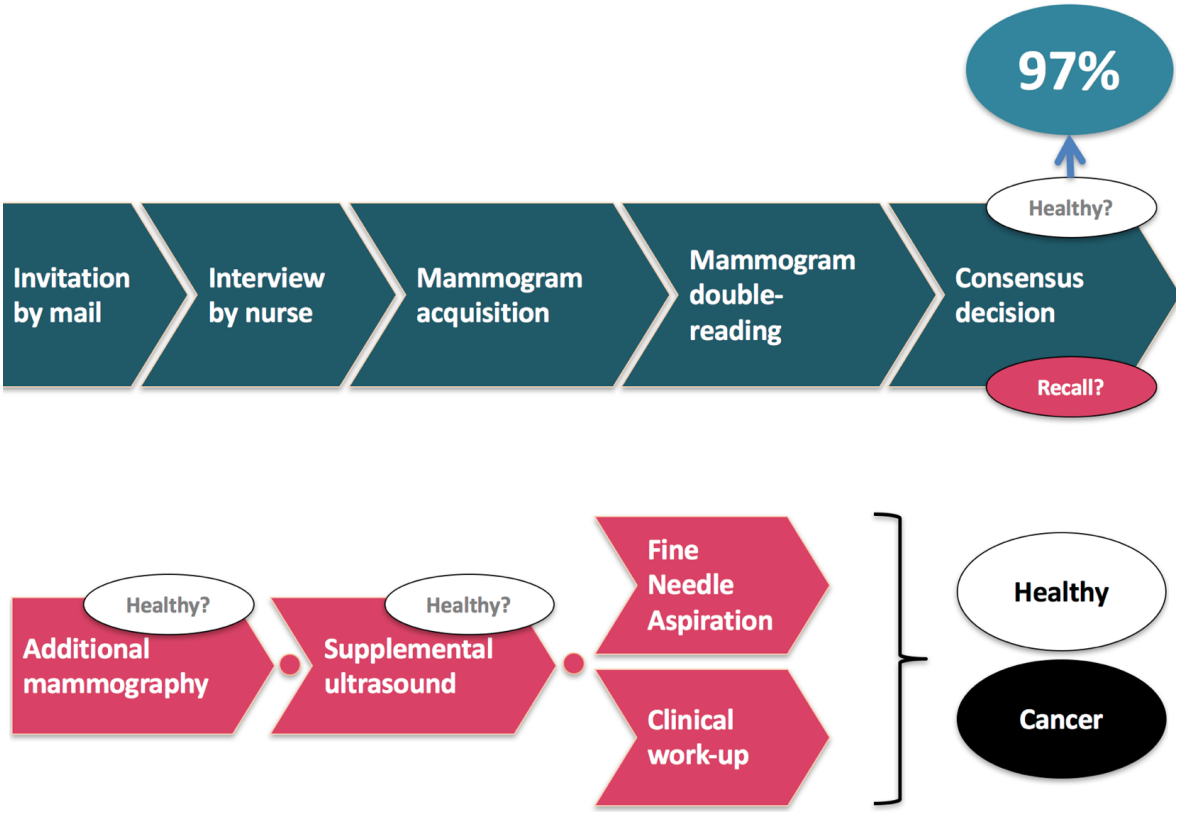

Figure 4. The screening process. The process until recall decision is shown in the top row. Around 97% are assessed to lack signs of malignancy, while around 3% are recalled for further assessments according to the bottom row. Around 0.5% are diagnosed with breast cancer.

The first pilot study of mammography screening in Sweden was performed in 1976, and by 1997 general screening had been established across all counties of Sweden (93). In Stockholm, general screening was introduced in 1989 for women aged 50 to 69 years with a 24-month interval between screenings. Between 2005 and 2010 women aged 40 to 49 years were gradually included in the screening program with an interval of 18 months between screenings; later changed to 24 months in 2014. The participation rate in Stockholm exceeds 70% (94). For women with a known high penetrance genetic mutations or strong family history of breast cancer, there are special screening regimens, often called surveillance. These women may start screening at an earlier age, have shorter time intervals between screenings, and be examined with other radiological modalities such as ultrasound and MRI as well as clinical breast examination.

The screening process is described in Figure 4 above. Double-reading means that two radiologists independently assess each case and identify suspicious findings which are 'flagged' for consensus discussion. Interpreting screening mammograms requires a substantial amount of experience and visual cognitive ability; specialized breast radiologists recall fewer women and find more cancers per 1000 women screened compared to general radiologists (95). To improve accuracy, the radiologist compares the current mammogram with prior ones. Comparison

improves specificity, while the effect on sensitivity is less evident (96, 97). After double-reading, the flagged cases are brought up in a consensus discussion, for arbitration, to ensure that the specificity remains high while sensitivity increases compared to a single reader (98). If a lesion cannot confidently be considered benign, the woman is recalled for further assessment. Another reason for recall is if the woman reports breast symptoms at the time of screening and especially if the symptom is a lump in the breast.

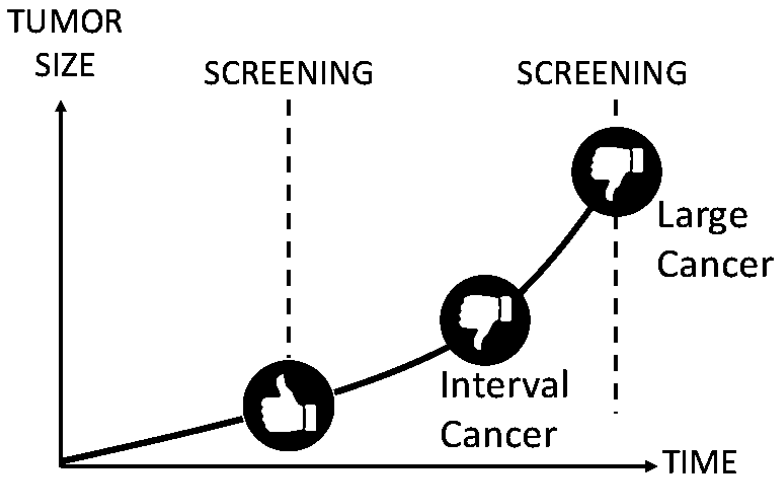

Figure 5. Illustrative graph of mode and opportunities of cancer detection. Interval cancer denotes a cancer clinically diagnosed after a negative screening. The cancer grows over time (solid black curve). Various opportunities of detection are marked (thumb icons). At periodic intervals, there are screening rounds when mammographic detection is possible (dashed lines).

In my studies, I have used observed cases of interval cancer and large cancer as proxies for delayed detection (Figure 5). Interval cancers are breast cancers clinically detected after a negative screening (further described in the 'Interval Cancer' section below). For studies I and II, the focus was on comparing IC with SDC. For study III, the focus was on comparing large (> 2cm) to small cancers. In Study IV, both comparisons were examined.

Risk-stratified Screening

In current population-based screening systems all women of a certain age are invited to screening with the same time interval between screenings and often using the same radiological method, e.g., mammography in Sweden. This one-size-fits-all approach for breast cancer screening is increasingly being questioned, and alternative risk-stratified systems have been proposed (99, 100). An accurate breast cancer risk estimation is important. Recently, models which take a broad range of risk factors, including mammographic density and genetics, into account have been proposed (101, 102). However, to accurately identify women at risk of interval cancer it is important to combine a general breast cancer risk model with a model for the risk of delayed detection, or mammographic masking, as shown by Kerlikowske et al. (103) Women with a high breast cancer risk and a high risk of mammographic masking had the highest incidence of interval cancer. As has been thoroughly discussed in a review article by Onega et al., it is important that the development of risk-stratified screening systems is accompanied by

systems to continuously evaluate individual benefits and harms, best practices for shared decision-making and comparative measures for different imaging methods (104).

1.4 Interval Cancer

1.4.1 Definitions

IC is defined according to the European Union guidelines as "primary breast cancer, which is diagnosed in a woman who had a screening test, with/without further assessment, which was negative for malignancy, either: before the next invitation to screening, or within a time period equal to a screening interval for a woman who has reached the upper age limit for screening" (105). Simplified, an IC is breast cancer that is clinically detected in the interval after a negative screening.

For each diagnosed breast cancer, the 'detection mode' should be determined: IC, SDC, or 'non-attender' (women who did not attend the prior screening). In a pooled analysis of six European countries, including Sweden, published in 2010, IC constituted 28% of IC and SDC cancers (106). The IC rate for the first 12 months after screening was 5.9 per 10,000 women and another 12.6 per 10,000 women the following 12 months. According to a review of several interval cancer studies, , by Houssami et al., the proportions are generally between 17% and 30% (107). In one study of annual screening it was 15%, and in a few studies of 3-year intervals it was 32% to 38%.

Most ICs are clinically detected due to the woman noticing a breast-related symptom. In one American study, cancer was diagnosed in 6.2% of 372 women seeking primary care for breast symptoms (108). The most common symptom that led to a diagnosis of cancer was palpation of a mass (88% of cancer cases). The most common symptoms that did **not** lead to a diagnosis of cancer were pain (42% of false alarms) and skin changes or nipple discharge (13% of false alarms).

When analysing ICs for quality control of mammographic screening, the following classification is often used. It is based on comparing the current positive mammogram with the prior negative mammogram (see Figure 6 on the next page). In my studies, this classification was applied in Study IV.

1. **True**: The prior screening mammogram was normal; there was no reason for assessment. This category may include occult ICs, which were not mammographically visible even at the time of diagnosis.
2. **Minimal sign**: There is a possible subtle abnormality on the prior screening mammogram, which would not necessarily have warranted further assessment.
3. **False negative (missed):** An abnormality is clearly visible in the prior screening mammogram and should have warranted further assessment.
4. **Occult:** A category sometimes included in 'True'. Denotes a cancer that was not visible on mammography even at the time of diagnosis.

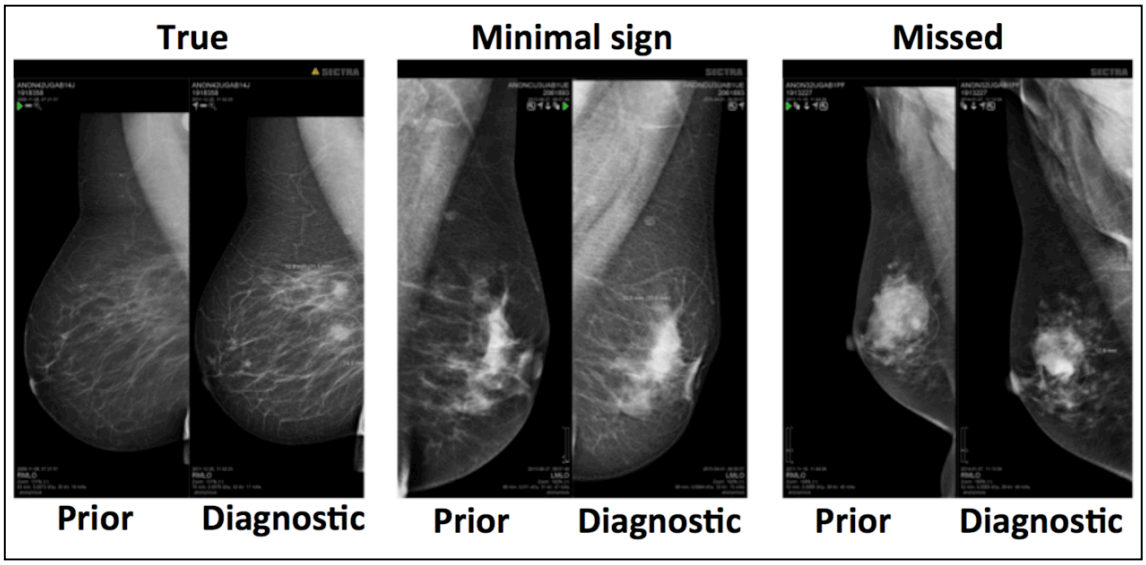

Figure 6. Pairs of prior and diagnostic mammogram for three review categories for interval cancer: True negative prior (left), Minimal sign (middle), and Missed or false negative (right).

In a Norwegian study of 231 ICs, the proportions were: 42% true, 23% minimal sign, and 35% false negative (109). There is large variation in the estimation of the false negative proportion. In a review from the UK, of IC classification in 15 different screening centers using film mammograms, the false negative rate ranged from 4% to 56% depending on whether the criteria was that all, a majority, or at least one of the reviewing radiologists were required to identify the prior abnormality (110). A Swedish study has shown that the proportion of examinations that were classified as 'missed' in a single dataset was 7% or 34% depending on whether the review was performed in a mix of cases and healthy women or in cases only. It has been shown that recalling a higher proportion of women for further assessment lowers the number of ICs (111).

1.4.2 Prevention

Preventing an IC does not mean that the woman will not have breast cancer at all, but that the cancer will instead be screen-detected, preferably at the prior rather than the subsequent screening. Preventive measures have been suggested: increasing the screening frequency to enable earlier detection of fast-growing tumors, and using supplemental imaging methods in addition to mammography, e.g., MRI or ultrasound, for increased screen-detection of otherwise masked tumors as described above in the 'Supplemental Imaging' section.

1.4.3 Determinants

High mammographic density increases the risk of IC. One mechanism is through masking, i.e., making it hard to visualize an incident cancer in the mammogram (112, 113). Table 1 below shows a summary of potential interval cancer risk factors. Three of the included studies focused on comparing IC to healthy women: 'JTL' by JT Lowery et al., 'MP' by M Pollán et al., and 'KK' by K Kerlikowske et al. (103, 114, 115). Four studies focused on comparing IC to SDC cases: 'JTL' by JT Lowery et al., 'JB' by Jordi Blanch et al., 'JH', by J Holm et al., and 'JL' by J Li et al. (115-118).

Table 1. Interval cancer determinants. Positive and negative associations as identified in six different studies are summarized.

Risk factor	Examined outcome and Association per study (*identified by first author initials below*)						
	Interval Cancer vs. Healthy			Interval Cancer vs. Screen-detected Cancer			
	JTL	MP	KK	JTL	JB	JH	JL
Mammographic density	↑	↑	↑	↑	?	↑	
BMI						↓	
Age	↑			↓			
Age at menarche							
Age at first birth		⊘					
Parity		⊘					
Age at menopause		⊘			↑		
Hormone replacement therapy	↑			⊘		↑	
Family history	↑	↑		⊘	↑	↑	
Previous breast biopsy		↑			↑		
Previous false positive screening					↑		
Early recall					↓		
Breast cancer risk by BCSC model			↑				
Hispanic vs. Non-hispanic white	↓			⊘			
Genetic susceptibility (PRS)							↓

↑ Positive association ⊘ No significant association
↓ Negative association (empty) Not examined

In summary, the risk factors consistently associated with increased risk are: high mammographic density, use of HRT, positive family history of breast cancer and prior breast biopsy. HRT is known to reduce the mammographic sensitivity, which is partly mediated by an increase in mammographic density (77). Recently, there was a study by Hofvind et al. confirming the association between prior false positive findings and subsequent IC (119). The following risk factors were identified in at least one of the studies: low BMI, high age at menopause, previous false positive screening, and less use of early recall after screening.

1.4.4 Prognosis

Women with IC have worse prognosis than women with screen-detected breast cancer (120, 121). In the latter study, by Eriksson et al. from our group, the higher hazard ratio for women with non-dense compared to dense breasts suggest that the proportion of intrinsically aggressive IC is highest among women with non-dense breasts. My Study III, shed further light on prognostic determinants for ICs and SDCs separately.

1.5 Tumor Size

1.5.1 Definitions

The tumor size is defined as the largest tumor diameter as measured at histopathological microscopy in the surgically excised breast specimen. The specimen is prepared and cut in thin slices, which need to be pieced together to contain the largest diameter of a cross-section of the entire tumor. Without using several large sections demonstrating the tumor on different levels there is a risk that the tumor size is underestimated (122).

Size is a continuous measure and there is no general agreement when the cancer should be considered large. However, having a size above 2.0 cm (20 mm) is used as a cut-off between stage T1 and T2 in the TNM staging system (123). In Study III and IV, we used the term 'large cancer' for cancers above 2.0 cm; for women with multiple tumor foci the size was measured on the largest invasive focus.

1.5.2 Prevention

Analysis of the Swedish two-county trial by Tabar et al. (89) showed that by participating in screening, women could expect the cancer size at diagnosis to be smaller than for non-participants. In a study from 2016 by Welch et al., it was shown that the cancer size at diagnosis became markedly smaller after the introduction of screening programs, and that this was due to a sharp increase in small-sized cancers, but only a modest decrease in large-sized cancers (124). An intervention that enhances the chance of self-detection would prevent large cancers. However, it would at the same time nominally increase the number of ICs by making a cancer clinically detected before the next planned screening where it would otherwise have been screen-detected.

1.5.3 Determinants

For women who participate in screening, having high compared to low mammographic density is a risk factor for large cancer size (125). In addition, having high BMI is a risk factor for large cancers, in general, according to previous studies (126, 127). These studies were not stratified by detection mode, and it is conceivable that the mechanism for reduced detectability, or 'masking', differ between mammography and clinical examination.

1.5.4 Prognosis

In one of the largest studies of the association between tumor size and survival among 24,740 breast cancer cases, published in 1989, it was estimated that the 5-year survival rate was 80 percent for women with cancers between 2.0 and 4.9 cm compared to 93 percent for tumors less than 2.0 cm (128). A later modelling study, by Michelson et al., based on three different study populations showed that there is very high correlation between tumor size and survival (129). The authors noted that not only did the introduction of screening decrease cancer size among screen-detected tumors, but there was also a decrease in cancer size among clinically detected ones.

2. Aims and Hypotheses

My overall aim of the thesis work has been to contribute to future trials of risk-stratified screening. The focus of my studies has been on identifying determinants of delayed detection by studying observed outcomes of interval cancer and large cancer at diagnosis. Interval cancer is breast cancer clinically diagnosed after a negative screening. Mammographic density was already known to be an important risk factor for both outcomes. To improve our understanding of density and to identify novel determinants, the specific aims of my four studies were:

I. **To examine whether large fluctuations in mammographic density were associated with an increased risk of IC compared to SDC.**
Our hypothesis was that large fluctuations might increase the relative risk of IC by making it harder to notice a subtle tumor when the background of normal breast tissue changes markedly between examinations.

II. **To identify novel computer-extracted image features, beyond mammographic density, associated with an increased risk of IC compared to SDC.**
Our hypothesis was that it should be possible to extract more information from the mammogram than a single density measure related to the risk of IC compared to SDC.

III. **To identify risk factors, for SDC and IC separately, of having a large cancer at diagnosis and to examine long-term prognostic implications.**
Our hypothesis was that there would be different risk factors for the tumor being large depending on the mode of detection. We thought mammographic density might be a more pronounced risk factor for tumor size among SDC cases and that other characteristics might be more pronounced among the mostly clinically detected IC cases.

IV. **To examine the effect of mammographic density, localized at the site of subsequent cancer, on the risk of being diagnosed with IC or large cancer.**
Our hypothesis was that density localized at the site of the subsequent cancer would be strongly associated with both IC and large cancer, especially for the slower growing less aggressive cancers.

3. Patients and Methods

3.1 Underlying Study Populations

The descriptions below refer to the original study populations from which the specific study populations for each of my studies were then selected based on criteria stated in the 'Results' section for each study I-IV.

3.1.1 LIBRO-1

In Study I-III, the underlying study population was based on the Linné-Bröst1 (LIBRO-1) cohort which was originally created by identifying all 11,696 female breast cancer cases incident in Stockholm from 2001 to 2008 (118). After exclusions due to being outside the age range 40 to 72 years at diagnosis, being deceased or without contact address, 9348 women remained. 61% gave informed consent after receiving an invitation sent by mail in early 2009, resulting in 5715 women being included in the final LIBRO-1 study population. Information on BMI, HRT use, and other socio-demographic, anthropometric, hormonal, and lifestyle factors was obtained through questionnaires. The median time between diagnosis and enrollment was 4.8 years (interquartile range: 3.0 to 6.6 years). BMI was calculated from self-reported length and weight.

Images

After exclusions of women without mammography screening within a regular screening interval, or without information on detection mode, 2901 women with IC or SDC remained in the underlying cohort. Both digital and analog mammograms were retrieved. The analog ones were digitized with a 12-bit dynamic range.

3.1.2 CAHRES

In Study I and II, the validation population was selected from the Cancer and Hormone Replacement Study (CAHRES) (130). In summary, CAHRES was based on women 50 to 74 years old when diagnosed with breast cancer, 1993 to 1995, and reported to any of the six Swedish Regional Cancer Registries. They were asked to participate through their physicians shortly after diagnosis. 84% consented to participate. A similar number of healthy control women were randomly selected, age-matched, from the Swedish population register. After exclusions, 2818 cases and 3111 controls remained in the underlying study population.

Images

Mammograms were retrieved for around 75% of participants (125). The mammograms were originally analog film-screen and were digitized as described above for LIBRO-1. After exclusions for poor image quality, lack of extended written consent and lack of postmenopausal images, 1747 cases of incident breast cancer remained in the underlying cohort. Descriptive statistics did not differ between the ones included and excluded, except for an age difference of less than one year.

3.1.3 KARMA

In Study IV, the main study population was based on the Karolinska Mammography (KARMA) cohort. KARMA is a prospective cohort created by inviting all women attending mammography at any of the recruiting mammography departments in Stockholm and Skåne between 2011 and 2013. In total, 210,233 women were invited and 70,877 (34%) gave informed consent to

participate (131). The age range was 21 to 95 years. The consent rate was highest, 39%, in the 65-69 age group and lowest, 30% in the 40-44 age group. Participants were asked to report on reproductive history, use of oral contraceptives and hormone replacement therapy, previous benign breast disease and family history of breast cancer. BMI was calculated based on self-reported height and weight. Baseline blood samples were collected.

Images

Mammograms were retrieved for 70,785 (99.8%) of the participants. All images were full-field digital mammograms in raw (a few initial ones missing) and processed format.

3.2 Data

3.2.1 Register Data (all Studies)

Population-based registers have a centuries-old tradition in Sweden, and national personal identity numbers have been in use since 1947. The personal number is assigned at birth, and can only be changed under very rare circumstances. For my studies, data were retrieved by linking personal numbers to following registers:

- *The Swedish Cancer Register* which contains information on type of cancer, date of diagnosis, invasiveness, TNM stage, and histological type. Already in 1978, 98% of all breast cancer diagnoses were reported (132).
- *The Breast Cancer Quality Register* which contains additional data on tumor receptor status, histological grade, et cetera.
- *The Cause of Death Register* which contains data on the cause of death for each individual since 1952.
- *The Screening Register at Regional Cancer Centre Stockholm-Gotland* which contains data on participation status and recall decisions for each screening mammography.

3.2.2 Cancer Detection Mode (all Studies)

Detection mode was ascertained using the population-based Screening Register at the Regional Cancer Centre Stockholm-Gotland. The cancer was defined as screen-detected if the woman was recalled from screening and diagnosed during the following diagnostic work-up. The cancer was defined as IC if it was clinically diagnosed within a normal screening interval after a negative screening. Remaining cancer cases were defined as diagnosed outside regular screening. At the time of the studies, the common regular interval between screenings was 24 months.

3.2.3 Mammograms and Density (all Studies)

For Study I-III, based on the cohorts CAHRES and LIBRO-1, analog film mammograms were collected from radiology departments and digitized with an Array 2905HD Laser Film Digitizer (Array Corp, Tokyo, Japan). For Study I-III, PD was measured using an automated validated method developed in our group (54). Briefly, the method mimics the gold standard PD measurement method Cumulus, which is based on a semi-automated thresholding procedure (133). In LIBRO-1 there were both digitized film-based mammograms as well as full-field digital mammograms. For Study I-II, in which the focus of investigation were features of the mammographic images beyond baseline PD, we included only digitized film-based mammograms.

For Study IV, overall mammographic density was assessed on a percent-scale using the validated automated Stratus software (56). It was then categorized on PD scale cut-points (2%, 18%, 49%) into four groups (cBIRADS) reflecting the definitions by the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) lexicon, fifth edition (58), see definitions in a previous section on 'Mammographic Density'. Localized density was assessed by visual assessment and categorized according to the BI-RADS standard based on two radiologists' consensus decision. First, the radiologists localized the tumor in the mammogram from the diagnostic time-point and then the corresponding location in the prior negative screening mammogram was assessed.

3.2.4 Image Feature Extraction and Selection (Study II)

In Study II, we examined a panel of 32 different image features that had previously been identified in a study from our group focusing on density prediction (134). We defined the dense area from which they would be extracted by three alternative thresholding methods. Before further analyses, we performed a global test of association to conclude whether there were any association with the outcome of IC vs. SDC status (further described under the section 'Methodological Considerations'). The features with the strongest associations with IC vs. SDC status were included in a multiple logistic regression model adjusted for age, BMI and use of hormone replacement therapy.

3.2.5 Tumor Characteristics (all Studies)

Data on tumor characteristics were obtained from the registers described above. In addition, missing data on tumor characteristics required to assign surrogate subtypes were retrieved from medical records.

For Study III, the surrogate subtype was defined based on the consensus of the 13th St Gallen International Breast Cancer Conference (2013) Expert Panel (135). The tumor was assigned the luminal subtype if it was estrogen receptor (ER) or progesterone receptor (PR) positive (or both), and if it was HER2 negative. For these luminal cancers, if both ER and PR were positive and the proliferation measure Ki-67 was < 14% we assigned it to Luminal A subtype, otherwise to Luminal B subtype. The tumor was assigned to HER2-overexpressing subtype if it was ER and PR negative and HER2 positive, and assigned to Basal-like subtype if it was ER, PR, and HER2 negative.

3.3 Epidemiological Study Design

The ideal research study would be having each study person simultaneously subject to different levels of the exposure of interest (e.g., having high and low BMI), and then evaluate the outcome. The influence of all other factors would be the same, and the outcome would then only depend on the exposure of interest. Since this counterfactual experiment is not possible, there are various epidemiological study types that aim to mimic the ideal situation as closely as possible. In the sections below, I will briefly describe and discuss the study designs used in my studies.

3.3.1 Cohort Study (part of Study III and IV)

A cohort study involves the observation of a study population over time, with different individuals having different levels of exposure of a potential risk factor (136), e.g., BMI or

mammographic density. The occurrences of the outcome of interest, e.g., breast cancer or progression of a disease, are continuously monitored. In a cohort study, the risk can be expressed "per person" or "per person-time". Person-time involves counting the individual time that each person is at risk of obtaining the outcome. The latter alternative is preferable when there are large differences between study persons regarding the time they are at risk. A person is not at risk when the potential outcome can no longer be monitored e.g., by the person moving abroad, dying, or otherwise becoming unavailable. It is called 'censoring' when such unobservable person-time is excluded and there are special statistical methods to handle this.

In my thesis, parts of Study III and Study IV were cohort studies. In Study III, cohort design was used for the follow-up of future breast-cancer events after diagnosis. In Study IV, cohort design was used to model the risk of interval cancer, and large cancer, compared to healthy women. Cohorts can rely on retrospectively collected data, e.g., when a study person is asked to remember smoking habits 10 years ago, or on prospectively collected data such as the mammograms in my studies that were recorded and stored as they were acquired and later collected for the studies. One advantage of prospectively collected data is to lower the risk of information bias, described in the section below on 'Bias and Confounding' in 'Methodological Considerations'.

3.3.2 Case-control Study (all Studies)

Sometimes a cohort design is not practically feasible, e.g., when a disease or outcome is rare. Performing a 'case-control study' makes the collection of exposure information more efficient since it can be limited to the individuals with the outcome of interest and only a limited number of healthy controls. The starting point for a case-control study is the collection of individuals who have obtained the outcome of interest. Then, controls without the outcome are sampled from the source population i.e., the individuals originally at risk. An important consideration is that the probability of being included in the sample should not depend on the exposure of interest otherwise there will be selection bias resulting in an incorrect estimation of the association between exposure and outcome.

The odds ratio (OR) is an often-used risk measure in case-control studies. The OR is defined as the ratio of exposed to non-exposed individuals among individuals with the outcome divided by the corresponding ratio among individuals without the outcome. Simplified, the OR shows how over- or under-represented the exposure is among those who obtained the outcome.

The term 'case-case study' is sometimes applied when the disease or outcome itself can be subdivided and analysed by different categories, e.g., breast cancer divided into SDC and IC. In principle, it is a type of case-control study but using the term case-case study highlights the fact that it is a comparison between two different categories of a disease rather than between diseased and healthy individuals. Studies I and II, and parts of Studies III and IV were case-case studies comparing IC with SDC, or small with large cancer.

3.4 Statistical methods

3.4.1 Linear Regression (Study II and III)

To examine the potential association between a determinant and an outcome, regression analysis is commonly applied. An advantage of a regression model is that it can be adjusted for, or take into account, multiple risk factors in a single model. In **linear** regression, a continuous outcome

is assumed to be a linear function of the determinants. For each determinant, a coefficient is estimated based on the line that best fits the observed data. If a coefficient is 0.5 it means that for every one unit change in the corresponding determinant, the outcome increases by 0.5. If the confidence interval of a linear regression coefficient includes both positive and negative numbers, i.e., that the 'true' slope of the regression line might be downward as well as upward, there is no statistically significant association.

Linear regression modelling was used in Study II and III with tumor size as the outcome, stratified by SDC and IC status.

3.4.2 Logistic Regression (all Studies)

Logistic regression is similar to linear regression, but is used when the outcome is binary, i.e., that it can have only two different values. This is particularly popular in medical research where regression models are used to examine or predict patients having vs. not having a certain disease or condition. When presenting the result in a medical context, the result is often presented as the estimated OR rather than the actual model coefficients. If the confidence interval of the OR includes 1, there is no statistically significant association.

Logistic regression modelling was used in all studies, either for the outcome of IC compared to SDC (Study I-IV) or for the outcome of large compared to small cancer (Study III-IV). Analyses were generally carried out crude and adjusted for age, BMI, HRT and other breast cancer risk factors as specified. In Study III, multinomial logistic regression, having the ability to handle multiple outcomes, was used for the analysis of tumor subtypes.

3.4.3 Mixed Effects Model – Mammographic Density Fluctuation (Study I)

For Study I, the long-term trend of mammographic density for each individual was estimated by a mixed effects model. A mixed effects model explicitly distinguishes two sources of variation, between individuals and between time points within each individual. Based on model estimation, one example from Study I of estimated and model-predicted PD measures is shown in Figure 7 below.

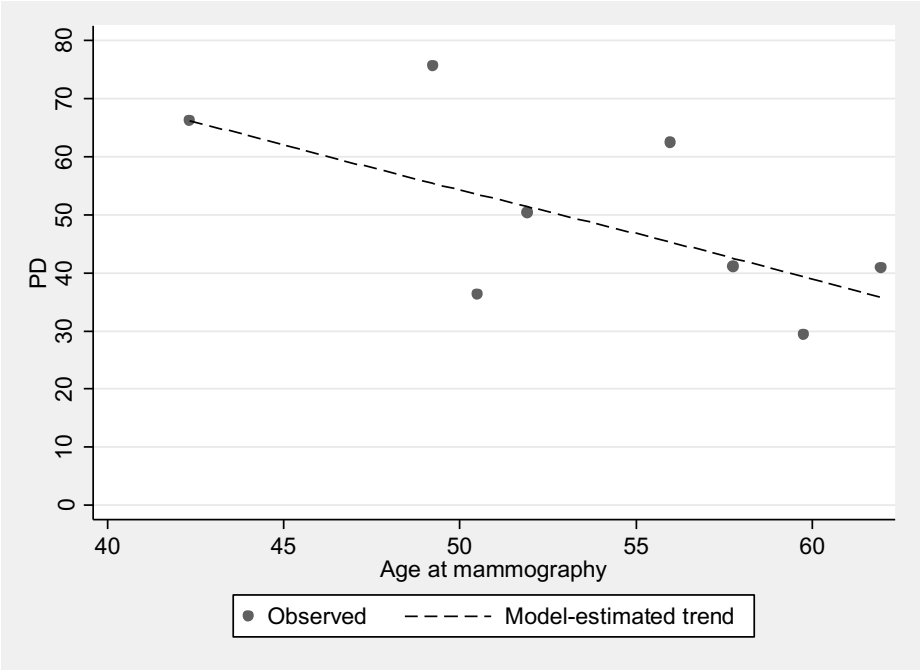

Figure 7. Example from one study person. The dashed line corresponds to the model-estimated PD values. The dots correspond to observed values at repeated mammogram examinations.

The model used in Study I, had both random intercepts and random slopes which lends the most flexibility to fit the line to each individual. As the second stage in Study I, the fluctuation measure for each individual was calculated as the root-mean-square of the absolute difference between each observed and model-estimated PD (i.e., the distance between each dot and the dashed line in the Figure above).

3.4.4 Cox Regression (Study III)

When analysing time-to-event data there is often a need for censoring, i.e., to take into account that some study subjects might leave, and re-enter, the study population at certain time-points. An outcome taking place during such absence would not be properly observed. These circumstances can be handled by Cox regression, a proportional hazards model, by which the ratio of one hazard to another can be estimated. In the basic model, the ratio is assumed to be constant regardless of time. In Study III, when examining the progression-free survival, the outcome of interest was defined as the first occurrence of either local recurrence, distant metastasis or death due to breast cancer. Patients were followed from the date of diagnosis until a breast cancer-event, death by other cause, emigration, or end of study period (December 31, 2015) whichever came first.

4. Results

4.1 Study I

In the primary study population, LIBRO-1, there were 1064 women with three or more digitized film-based mammograms, 178 women with two mammograms, and 172 women with one mammogram, among postmenopausal women without prior breast surgery. In total, there were 385 IC and 1029 SDC cases included. Women with IC had a higher average mammographic

density fluctuation than women with SDC, 0.44 and 0.41 respectively (p=0.031). There were no significant differences between IC and SDC cases regarding the following potential confounders: total number of mammograms per person, average time between mammograms, mean age at mammography or time from first to last mammogram.

Mammographic density fluctuation was, after multiple adjustments, associated with the risk of interval compared to screen-detected breast cancer with an estimated per-standard deviation OR of 1.17 (95% CI = 1.03 to 1.33). In the independent validation cohort, CAHRES, the corresponding OR was 1.19 (95% CI = 1.04 to1.38). After categorizing the fluctuation into quartiles, we estimated the probability of IC for each quartile (Figure 8). The probability of IC, compared to SDC, was 27% in the top quartile and 19% in the bottom quartile of PD fluctuation.

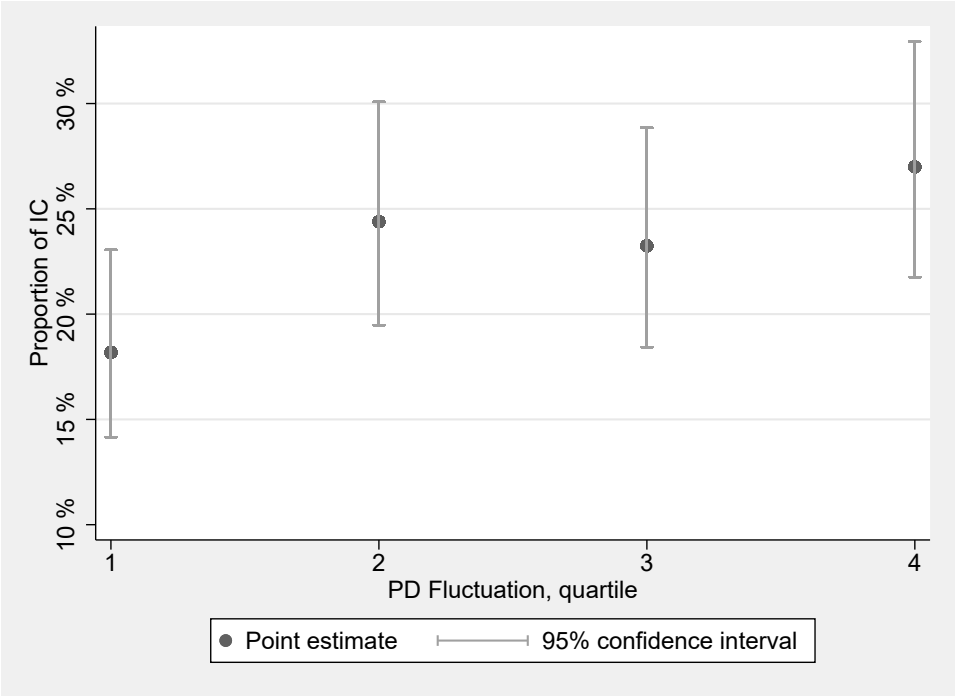

Figure 8. Proportion, or probability, of IC for each quartile of mammographic density fluctuation. Quartile 4 includes women with highest fluctuations and quartile 1 women with lowest fluctuations. Women in quartile 1 have a lower probability of being detected as interval cancer than women in the other three quartiles.

4.2 Study II

In the primary study population, LIBRO-1, there were 1403 women without prior cancer or benign breast surgery who had at least one pre-diagnostic digitized film-based mammogram. There were 394 IC and 1009 SDC included. Two image features, 'skewness of the intensity gradient' and 'eccentricity', were associated with the risk of interval compared to SDC. The two features were validated in the independent study population CAHRES (p=0.05, p=0.038). In secondary analyses, the first feature was associated with the tumor size at screen-detection (p=0.00001), while the second feature was associated with the tumor size at interval detection (p=0.018). In Figure 9 below there are examples of mammograms with high and low values of each feature.

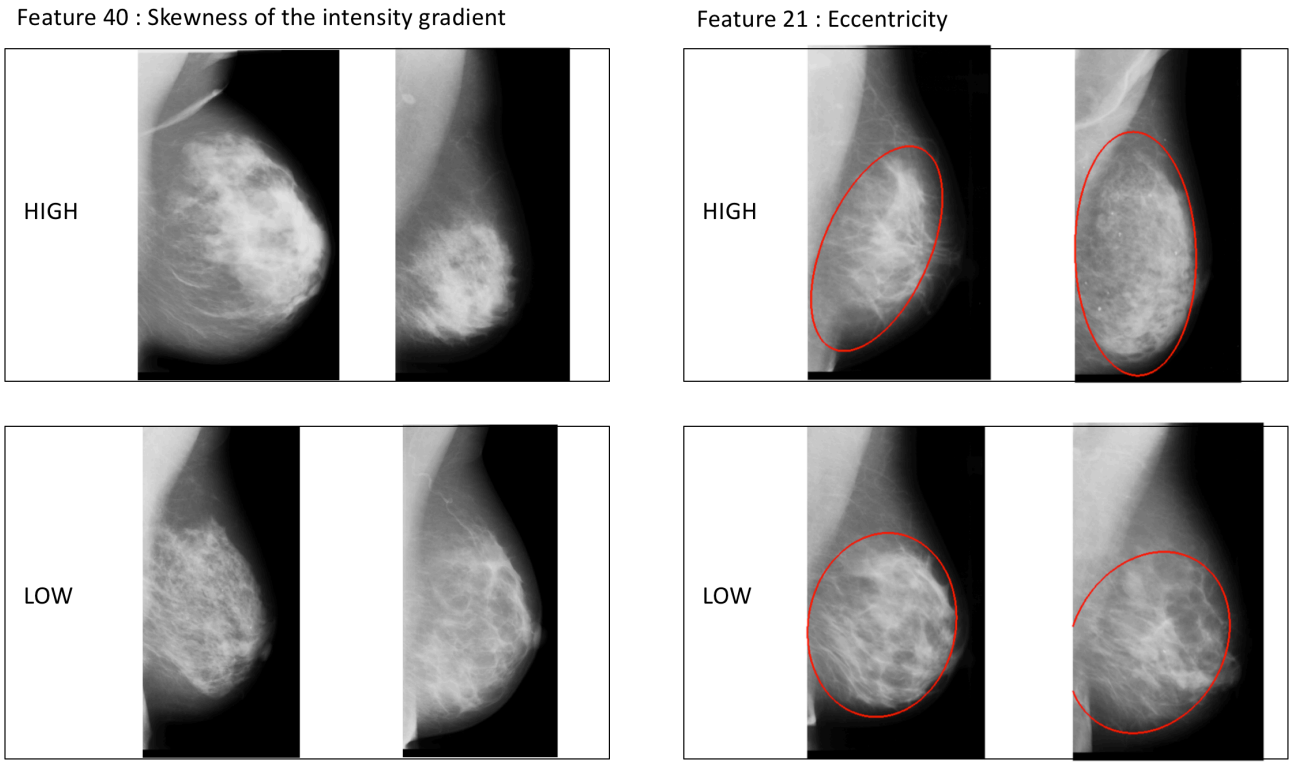

Figure 9. Example mammograms to illustrate high and low values of each of the two identified image features.

4.3 Study III

There were 2012 women in the study population that had information on PD, cancer size and detection mode and were over 50 years of age. There were 546 IC and 1466 SDC. Among ICs, there were 185 (34%) large cancers, and among SDCs there were 281 (19%) large cancers.

Overall, BMI and PD were the only factors associated with cancer size at diagnosis. Stratified by detection mode: the multiple adjusted OR for BMI (≥ 25 vs. < 25 kg/m2) was 1.37 (95%CI: 1.02-1.83) and 2.12 (95%CI: 1.41-3.18), for screen-detected and ICs respectively. The corresponding OR for PD (≥ 20 vs. < 20 %) was 1.72 (95%CI: 1.29-2.30) and 0.60 (95%CI: 0.40-0.90). Model-predicted probabilities for being diagnosed first after the cancer had become large, as a function of BMI, are shown in Figure 10.

Among IC, women with high BMI (solid line in Figure 11 below) had worse prognosis than women with low BMI (hazard ratio 1.70; 95% CI: 1.04-2.77). Among SDC, BMI was not associated with prognosis. PD was not associated with prognosis among IC or among SDC cases.

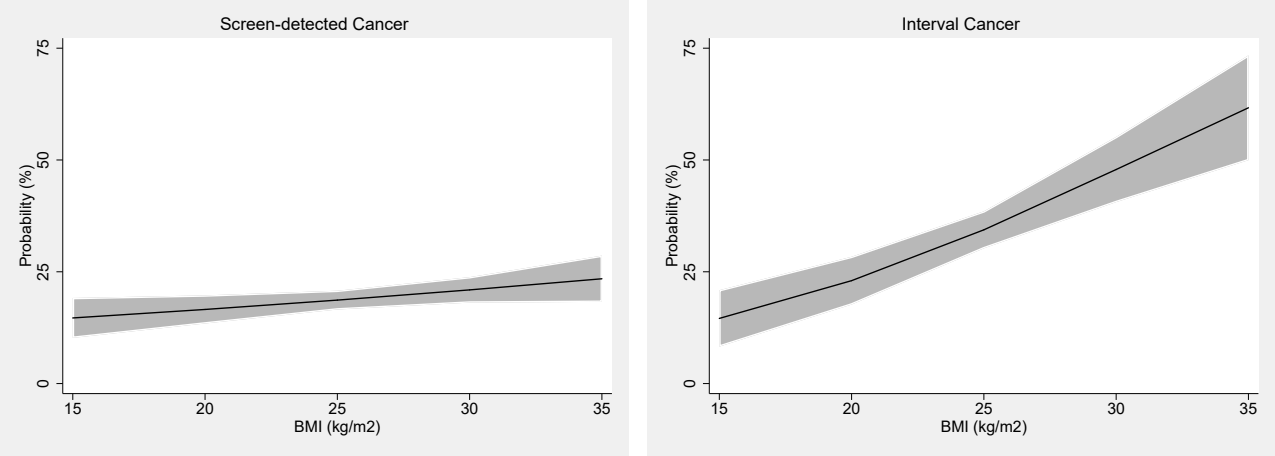

Figure 10. Model-estimated probability of being diagnosed with a large cancer, among screen-detected cancers (left) and among interval cancers (right), as a function of BMI.

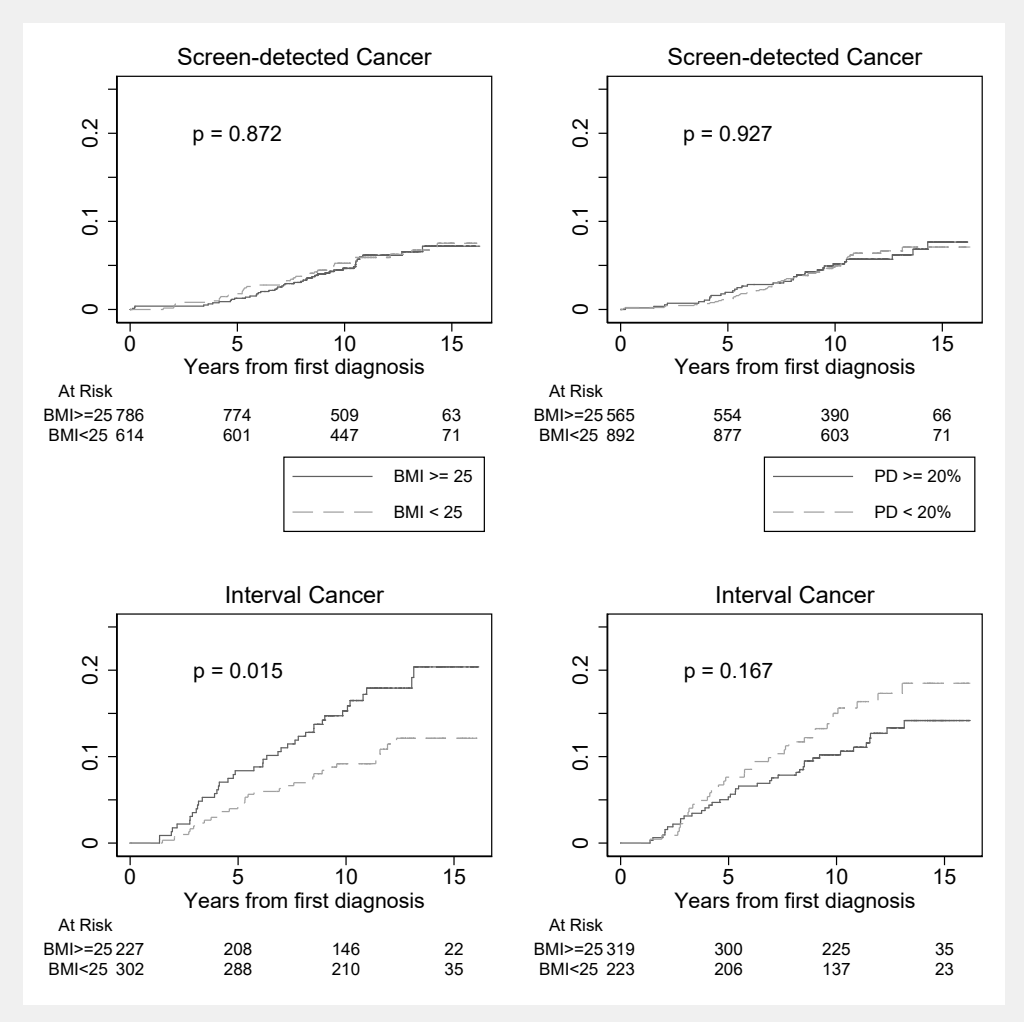

Figure 11. Cumulative Hazard function for disease progression in terms of local recurrence, distant metastasis or breast-cancer death. Screen-detected caners (top panels) and interval cancers (bottom panels) by levels of BMI (left-side) and PD (right-side).

4.4 Study IV

There were 63,130 women in the underlying study population KARMA who were within the screening age range 40 to 74 years, had information on detection mode, had at least one mammogram between study entry and before any breast cancer diagnosis, and who had no prior breast cancer. At the time of study inclusion, there were 891 women with breast cancer, and 386 for whom images were available allowing assessment of localized density. In total, there were 246 IC and 250 large cancers. There was a modest correlation between overall and localized density, r = 0.42. The probability of having concordant density categories was 0.35 (95%CI: 0.21 to 0.49), 0.60 (95%CI: 0.52 to 0.68), 0.38 (95%CI: 0.30 to 0.45) and 0.32 (95%CI: 0.18 to 0.46) for category 'A', 'B', 'C' and 'D' density respectively. Figure 12 below shows the distribution across categories of overall and localized mammographic density of the 386 cases for which local density assessment was performed.

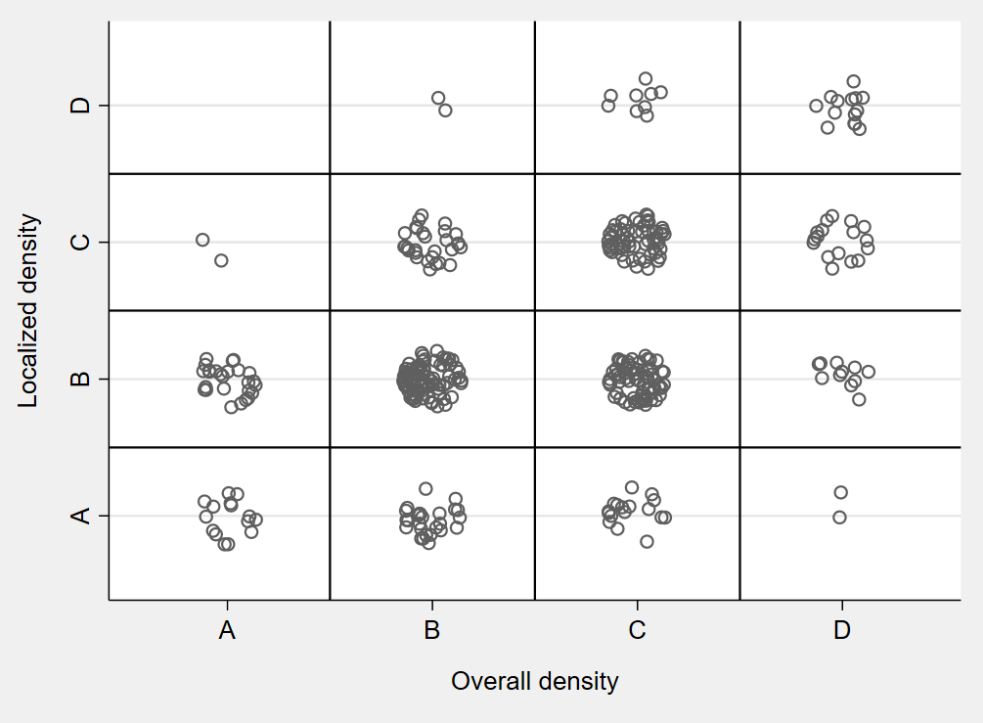

Figure 12. Distribution of the 386 cases across categories of overall and localized mammographic density. Markers in the graph were jittered for visibility.

For the outcome of IC compared to SDC: for women with localized density category 'D' compared to category 'A' the estimated OR was 3.1 (95%CI: 0.8 to 11.5) with p=0.076 for the linear trend after multiple adjustments including overall density. Among women with lymph node negative cancers the corresponding OR was 5.3 (95%CI: 1.3 to 21) with p=0.031 for the linear trend. For the outcome of large cancer compared to small cancer: the corresponding ORs were 10.3 (95%CI: 2.7 to 38.9) with p<0.001 for the linear trend overall, and 19.1 (95%CI: 4.4 to 82) with p<0.001 for the linear trend among lymph node negative cancers only.

5. Discussion

5.1 Study I

We found that large fluctuations in mammographic density were associated with an increased risk of IC compared to SDC. To what extent this association is related to inconsistencies in image acquisition, reduced mammographic detectability or a marker of growth-promoting biology remains to be elucidated in future research. In addition to different positioning, also differences in breast compression can affect the density measurement (137). Another study showed some evidence for fluctuations in mammographic density correlated with phases in the menstrual cycle (138). Future research should try to validate the findings in a screening cohort of healthy women. For such validation, it might be beneficial to use an image processing pipeline that realigns images and reduces the influence of technical variations, to more accurately appreciate true biological change (56).

5.2 Study II

We identified two novel features that were associated with the risk of IC compared to SDC, independently of PD. A high value of one feature, 'Eccentricity', was associated with decreased size at IC detection, suggesting a breast anatomy making clinical detection easier. A high value of the other feature, 'Skewness of the intensity gradient', was associated with an increased size at screen-detection, suggesting a breast anatomy making mammographic detection more difficult. Previous studies of novel image features, in relation to IC, have focused on identifying pre-malignant changes and not on features of the normal breast (139, 140).

5.3 Study III

On an overall level, we found that women with high BMI or high PD had an increased risk of being diagnosed with a large compared to a small cancer. Both findings are supported in prior research (141). Conclusions after stratification for detection mode are reported below.

High BMI was associated with having a large cancer among both SDCs and ICs, with a higher point estimate among ICs. This suggests more difficult clinical detection with high BMI, which might 'spill over' to larger sizes also at mammographic detection. The association with low BMI might to some extent be related to having smaller breasts facilitating symptomatic clinical detection (142).

High PD was associated with having a large cancer among SDCs only, and actually showed an opposite association among ICs. This suggests that high PD makes only mammographic detection, but not clinical detection, more difficult. In the studies cited above, there is ample support for high PD being associated with masking, requiring the tumor to be larger before it becomes visible. There is no prior study showing that high PD makes clinical detection easier.

In supplementary analysis, we found that that among women with IC having high BMI is associated both with larger tumors and with a larger proportion of HER2-overexpressing tumors. Not contradicting our results, a recent study by Musolino et al. showed that early stage HER2-overexpressing breast cancer had a lower survival if detected as IC compared to as SDC (143).

High BMI was associated with worse prognosis, but only if the cancer was detected as an IC. This effect of BMI among clinically detected cancers is supported by a prior Swedish study suggesting that the prognostic value of BMI was higher before, compared to after, the introduction of population-wide screening programs (144). High PD, on the other hand, was not associated with prognosis. These results agree with a recent study of density and breast-cancer specific survival by van der Waal et al. which showed no significant association, but a lower hazard ratio point estimate for SDC compared to IC (145). One can speculate that the 'clinical masking' provided by high BMI is worse, in terms of prognosis, compared to the 'mammographic masking' provided by high PD. To reduce the effect of potential clinical masking, women with high BMI should perhaps be especially encouraged to participate in screening and shorter screening intervals might be considered.

5.4 Study IV

High localized density was strongly associated with having a large rather than small cancer at diagnosis, and to some extent also with being diagnosed with IC rather than SDC. No prior studies were identified. To elucidate whether these associations were caused by mammographic masking or by biological differences affecting tumor growth was not the aim of this study. However, the stronger association among the less aggressive, node-negative, cancers suggests that masking is the more important mechanism. Before clinical implementation of localized density assessment our visual approach should be replaced by a computerized method - perhaps an adaptation of the breast-density map method used in a recent study by Holland et al. (146). As expected, high overall density was associated with both large cancers and with IC, both when comparing with healthy women and in case-case analysis.

6. Methodological Considerations

6.1 Bias and Confounding

Sampling Bias

It is important to always consider the possibility of sampling bias, i.e., that the sampling of cases and controls were not independent of the exposure of interest. The percentages of responders were 84%, 61% and 34% for the CAHRES, LIBRO-1 and KARMA cohorts respectively. CAHRES and LIBRO-1 were based on cases, and there is no obvious reason why the response rate would depend differently on any of the studied exposures when comparing women with IC to women with SDC. For the KARMA cohort which was based on all women attending mammography units, healthy women with self-perceived high risk of breast cancer might be more prone to participate in the cohort. This might attenuate the association between breast cancer and family risk, but would not affect the comparison of SDC and IC cases.

In addition, there is potential sampling bias in the LIBRO-1 cohort related to survival since study inclusion was performed between one to eight years after diagnosis. Women with the most advanced, aggressive, tumors are generally less likely to be alive at later time points, and should thus have been less likely to be included in the cohort. If this imbalance constitutes a relevant sampling bias or not depends on how survival is related to the outcomes of interest (IC vs. SDC). Since ICs are generally more aggressive than SDCs, the bias has most likely lead to an attenuation ("bias towards the null") of differences between IC and SDC in terms of characteristics related to tumor aggressiveness.

Lead Time and Length Bias

Lead time and length bias are central concepts in evaluations of screening efficacy (147). Lead time is the difference between the time when the cancer was actually detected at screening and the time when it, in the absence of screening, would have been detected clinically (148). In screening, the aim is to detect cancer as long before it becomes clinically detectable as possible, i.e., to maximize the lead time with the aim to increase survival. However, in research studies of survival time, the lead time becomes a bias in so far that it is added to the survival time of the screen-detected cancers only. The theoretical concept 'sojourn time' is closely related to lead time. Sojourn time is the time between when a tumor becomes detectable on screening and the time when it is clinically detectable. For a particular case, the sojourn time starts when a cancer becomes screen-detectable. Later, the starting point for lead time is when the cancer is actually screen-detected (if at all). Based on data from clinical trials on breast cancer early detection, it has been estimated that the average breast cancer sojourn time is around 2 years, with some dependence on age (148, 149).

The source of length bias, on the other hand, is that tumors that are intrinsically less aggressive have a longer sojourn time are more likely to become screen-detected, and less likely to be clinically detected. Length bias should be considered when comparing tumor characteristics between SDC and IC with the aim to make conclusions about survival. Obviously, the length bias is potentially stronger when the time between screenings is long and the follow-up time in the research study is short, and should disappear in a theoretical situation with daily screening and be attenuated when the follow-up time is long.

Information Bias

Information bias arises if exposure measurements are acquired in a way that makes them systematically different depending on the outcome of interest. In my studies, it would arise if exposures would be measured differently for women with IC compared to SDC or for women with large compared to small cancer. However, there is no obvious reason why the main risk factors in my studies, age, mammographic density, BMI and HRT should be affected. Related to the survival analysis in Study III, potential bias due to the timing of BMI measurements is discussed in the section on 'Study III' below.

Confounders

A confounder is a risk factor that is associated with both the outcome and the exposure of interest, and might distort the modelled association between them. There are two fundamentally different ways of deciding which potential confounders to include in an analysis. The first is a technical approach in which any parameter is considered a potential confounder, and its inclusion or exclusion is decided on statistical grounds. The other is an approach in which the parameters are selected based on prior knowledge about potential biological mechanisms. In our study, we have generally chosen the latter approach. However, for Study II forward-selection was used for novel image features since we had no a priori knowledge of connections between any image feature and biological mechanisms.

6.2 Study I

In Study I, we first examined the possibility to use standard deviation as a measure of fluctuation. However, as shown in Figure 13 below, a high standard deviation might be related to a large long-term decrease and not necessarily to fluctuations around the trend line. Therefore, we decided to use mixed effects modelling, as the first stage, to define the long-term trend line on the model-estimated PD for each woman over time.

The overall individual fluctuation measure was then calculated as the root-mean-square of all deviations between observed and model-estimated PD. As the second stage, we fitted logistic regression models to examine the association between the fluctuation and the outcome of interest, i.e., IC compared to SDC.

From a statistical standpoint, an issue with the two-stage approach in our study was that the uncertainty in estimating PD fluctuation at the first stage was not carried over to the second stage. After fitting an alternative single-stage model in which this issue would not arise, we could confirm the conclusion that PD fluctuation was higher for IC than SDC cases. A limitation of the study is that it was based on a case-only cohort, from which conclusions on screening among healthy women cannot be made.

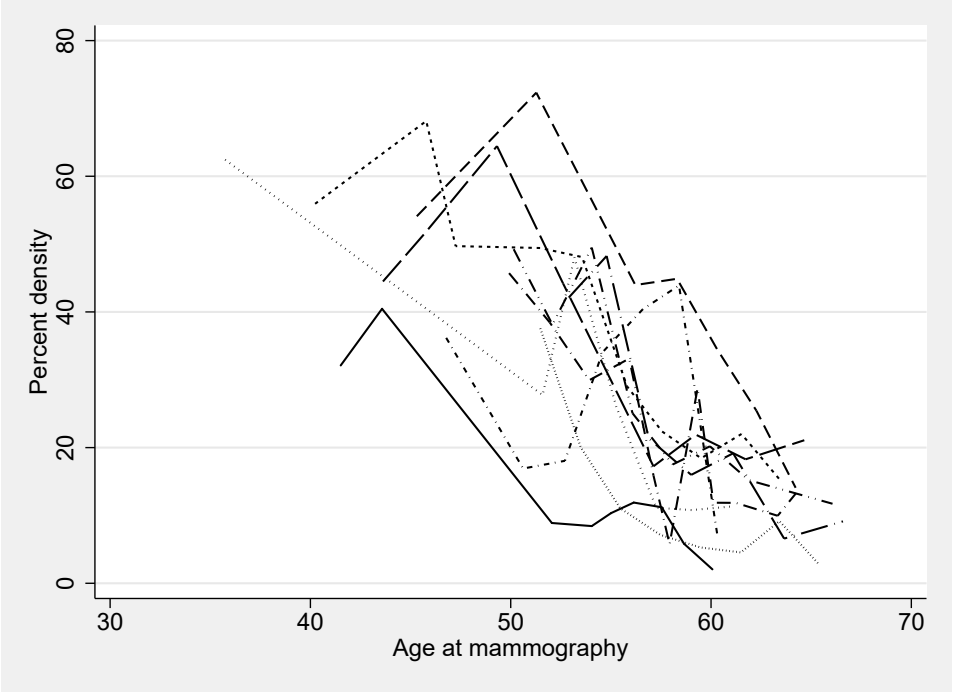

Figure 13. Observed PD values, with connecting lines, for the 10 women who had the highest standard deviation. Standard deviation did not appear to be a good measure to distinguish fluctuation from a long-term trend.

6.3 Study II

Due to the high correlation between feature values, we applied a global test of association with IC vs. SDC status before examining each feature separately. For the global test, we first carried out tests of association for each of the 32 features for each of the 3 definitions of dense area. These tests were based on fitting logistic regression models with IC vs. SDC status as the outcome and using continuous PD as an adjustment variable. Then, we performed a global test of association testing the null hypothesis that none of the features were associated with IC vs. SDC status by examining the **number** of test results that were significant at the 5 % level (global test statistic). An empirical (global) level of significance was obtained by permuting IC vs. SDC status over a large number of simulations (10,000), and calculating the fraction of (global) test statistic values based on permuted data that were larger than the test statistic value obtained for the non-permuted data set. This global test is similar to Wilkinson's test but accounts for the correlation of the features. After we had concluded that there was a global difference between the features of IC and SDC, we continued to identify individual features as described above in the section on 'Image Feature Extraction and Selection'.

A limitation of our study was that we extracted features from digitized analog mammograms and not from digital mammograms. The same feature calculation methods can be used for both types of mammograms, but our results would need to be validated among digital mammograms.

6.4 Study III

In the other studies, BMI was used as an adjustment. In Study III however, it was a key exposure of interest. Therefore, we examined the possibility that the identified associations with BMI could be biased by the fact that it was measured at time of study inclusion which could be years

after diagnosis. We had to consider the possibility of reverse causality – that having a large tumor would cause an increase in BMI. However, this did not seem biologically plausible. To examine the influence of the time difference between diagnosis and measurement, we conducted a sensitivity analysis based on three separate regression models, with tumor size as the outcome: less than 3 years, from 3 and less than 6 years, 6 years and more. The analysis showed similar effect sizes of the association between BMI and tumor size across all three time-delay categories. Furthermore, meta-analysis had shown that higher BMI was associated with worse outcome regardless of the time when it was ascertained (150); before or after diagnosis, more or less than 12 months after diagnosis.

Even though we thought that potential survival bias would confer a 'bias towards the null', we carried out a sensitivity analysis by introducing a term for the time between diagnosis and study. An additional sensitivity analysis was performed by limiting the survival analysis to the 20% of cases diagnosed closest to study entry. The hazard ratio point estimate was somewhat higher in this group than in the entire study sample. The results of these analyses showed no evidence for survival bias affecting the conclusions of our study.

Finally, having a categorical cut-off at a certain tumor size is to some extent arbitrary, even if our definition has been used in several other publications and systems. Therefore, it was reassuring that linear regression modelling confirmed the corresponding associations with BMI and PD.

6.5 Study IV

In Study IV, there was potential uncertainty in the assessment of localized density. The radiologist had to localize the tumor in the current mammogram and then identify the corresponding location in the prior mammogram. The degree of uncertainty depended on the tumor appearance and on the similarity between the two mammograms from different time points. A potential bias in Study IV was that high localized density actually was an early manifestation of cancer. It was reassuring that limiting the analysis to cases that did not show cancer signs at retrospective review did not markedly affect the estimated effect. Even if this potential bias could not have explained our findings, it cannot be ruled out that certain increased densities corresponded to subtle cancer.

7. Ethical Considerations

Research ethics at Karolinska was recently brought into the spotlight of public debate and it was evident that ethics must be an important and integral part of conducting research. Even if these incidents concerned questionable surgical procedures, ethics is naturally relevant for all research whether it is observational or interventional. As researchers, we must protect study participants against the risk of physical injury and violation of integrity. They must have the right to be fully informed of any potential risks before actively choosing through 'informed consent'. The European Council has developed a 'Convention on Human Rights and Biomedicine', which requires legal regulation of the ethical process and how the ethical review boards should operate. The Ethical Review Act (2003:460) addresses the handling of sensitive personal data, the collection of biological samples, and informed consent:

http://www.epn.se/media/45159/the_etical_review_act.pdf

For the CAHRES cohort, women with incident breast cancer were invited shortly after diagnosis as well as a similar number of healthy control women selected from the population register. For the LIBRO-1 cohort, participants with incident breast cancer during 2001 to 2008 were asked to participate. For the KARMA cohort, all women attending mammography were asked to participate in a prospective cohort. All three cohorts required biological samples, access to hospital records and images, as well as questionnaire data.

Ethical approval was granted by the regional ethical review board in Stockholm for each of the studies included in this thesis:

- CAHRES Dnr 155/93 with extensions Dnr 99-338 and Dnr 2005/360-32
- LIBRO-1 Dnr 2009/254-31/4 with extension Dnr 2011/2010-32
- KARMA Dnr 2010/958-31/1

Use of register-based data is one of the few areas in medical research, which is sometimes exempted from the requirement of informed consent by the decision of the regional ethics review board. However, for the three cohorts of my studies written informed consent was obtained from each participant. Participants were informed that they could withdraw at any time. They gave explicit informed consent to allow for matching to registers such as the Cancer Register, the In-and Outpatient Register and Cause of Death Register. The personal numbers of study participants were exchanged for an anonymous study participant id number. Only a small number of study personnel had access to the linking key.

8. Concluding Remarks

The work in this thesis has been carried out with the main aim to contribute to future trials of risk-stratified screening. The focus has not been on breast cancer risk per se, but on the risk of delayed detection in terms of interval cancer and large cancer at diagnosis. In LIBRO-1, the study population mostly used in my studies, the proportion of those two outcomes together was around 40% (table below).

	Screen-detected n=1,636	Interval Cancer n=734
Small <= 2cm	1,187 (59%)	**360 (18%)**
Large > 2 cm	**279 (14%)**	**186 (9%)**

For risk-stratified screening to be successful, women at elevated risk of delayed detection must be identified as precisely as possible. Women who are predicted to have a low risk but actually have a high risk, would be wrongfully assigned a less intense screening program potentially leading to delayed detection and increased mortality. Conversely, women who are predicted to have a high risk but actually have a low risk, would need to go through an unnecessary amount of examinations and procedures. When starting my research studies, it was already known that mammographic density was associated with interval cancer and large cancer. The results from each of my studies has contributed understanding of further determinants and independent risk factors:

- In Study I, we estimated the individual magnitude of density fluctuation over time and found that high fluctuation was a determinant of IC compared to SDC.

- In Study II, we extracted novel image features and identified two features that were determinants of IC compared to SDC.

- In Study III, we identified high PD and high BMI as determinants of having a large cancer, and we showed that these associations differed between detection modes. In addition, we showed that for women who were diagnosed with IC, having a high BMI was a determinant of worse prognosis.

- In Study IV, we found that the localized density in the prior negative mammogram at the location of the subsequent cancer development was a determinant of future tumor size and also of being diagnosed with IC among the less aggressive node-negative cancers.

How our understanding of these determinants can be applied, and further validated, will be outlined in the next section 'Future Perspectives'.

9. Future Perspectives

Taking the next step in improving breast cancer screening through risk-stratified screening assigning individual time intervals between screenings and individual choice of radiological method require precise risk assessments (101). As a next step in the scientific exploration, I would suggest to set up a trial based on a combination of a general breast cancer risk model (102) and a delayed detection model. The latter model could benefit from the results of my studies.

Determinants for interval cancer could be used to identify women that might benefit from shorter time intervals between screenings or from a more sensitive radiological method. Determinants for large cancers are to some extent different depending on mode of detection. In Study III, BMI was found to be associated with larger tumors and worse prognosis among interval cancers, suggesting that women with high BMI should be offered shorter time intervals between screenings. In Study III, overall density was associated with large cancer among screen-detected cases, suggesting that women with high density should be offered a more sensitive radiological modality. Future research should study the effect of these different criteria for the individual choice of time interval and of radiological method. I am personally aiming to contribute by studying the effects of offering MRI-based screening for a selected group of women at elevated risk of breast cancer and of delayed, or reduced, mammographic detection.

Overall density has previously been shown to be associated with reduced mammographic detection. In Study IV, we identified localized density as a further determinant of large cancer and interval cancer independently of overall density. Since visual assessment of localized density is not feasible in a screening setting, a suggestion is to develop computer-calculated 'minimal detectable tumor size' maps and summary measure for each mammogram.

In all of my studies the algorithms for calculating image characteristics, including density, were specified by regular computer programming. The current rise of machine learning and specifically deep learning which enables classification directly from image pixel data shows great promise for breast imaging (151-153). I plan to contribute to this discovery process by examining how deep learning can be applied to improve different aspects of breast cancer screening including early detection and a more efficient radiological workflow.

In summary, based on the findings of novel determinants and increased understanding of interval cancers and large cancers, there are several promising avenues for future research that I hope will lead to further improvement of breast cancer screening and reduced breast cancer mortality.

Svensk sammanfattning (abstract in Swedish)

Bröstcancer är den vanligaste cancerformen för kvinnor, både i Sverige och globalt. I många höginkomstländer har mammografiscreening i ett par decennier framgångsrikt bidragit till minskad dödlighet. På senare tid har intresset ökat för att prova att ersätta dagens åldersbaserade system med ett riskbaserat system. I ett riskbaserat system avgörs, för varje individ, ett lämpligt tidsintervall mellan screening-tillfällen och en lämplig radiologisk undersökningsmetod. För att göra detta krävs en bra modell för att individuellt avgöra risken för bröstcancer samt risken för fördröjd upptäckt. Mina studier i denna avhandling har varit inriktade på att bidra till det sistnämnda. Mammografisk täthet är ett mått på hur mycket tät bröstvävnad som finns, och är en känd riskfaktor för fördröjd upptäckt. Mitt mål har varit att identifiera och förstå ytterligare riskfaktorer genom att fokusera på kvinnor med intervallcancer samt stora cancrar och jämföra dessa med screeningupptäckta respektive små cancrar. Intervallcancer är cancer som upptäcks efter en negativ screening i tidsintervallet fram till nästa screening. Mammografierna som datoranalyserades var de negativa, som alltså saknade påvisad tumör. Studie I till III baserades på retrospektiva fall, medan studie IV baserades på en prospektiv kohort.

I **Studie I** utvecklade vi ett mått på fluktuationer i mammografisk täthet över tid. Vi kunde konstatera att de kvinnor som diagnostiserades med intervallcancer hade högre täthets-fluktuationer än de vars cancer upptäcktes vid screening.

I **Studie II** undersökte vi 32 datorberäknade bildegenskaper i det täta området i mammografin, och fann att två av dessa skilde sig åt mellan intervallcancrar och screeningupptäckt cancer. Den ena bildegenskapen var relaterad till formen på det täta området. När detta var mer platt än runt ökade risken för intervallcancer, möjligen på grund av att klinisk upptäckt underlättades. Den andra bildegenskapen föreföll vara relaterad till om tätheten var koncentrerad eller om den var uppbruten av mindre täta stråk. När värdet indikerade mer koncentrerad täthet var risk för intervallcancer högre, möjligen på grund av att mammografisk upptäckt försvårades.

I **Studie III** identifierade vi riskfaktorer för att cancern hunnit bli större än 2 cm innan diagnos. Det var redan känt att hög mammografisk täthet liksom högt BMI ökade risken för detta. Vi undersökte om riskfaktorerna skilde sig åt beroende på upptäcktsmetod – kvinnor med screeningupptäckt cancer respektive kvinnor med kliniskt upptäckt intervallcancer. Högt BMI var associerat med större tumörstorlek oberoende av upptäcktssätt, medan hög täthet bidrog till stora cancrar endast bland de screeningupptäckta. En långtidsuppföljning visade att bland kvinnor med intervallcancer var högt BMI relaterat till ökad risk för lokalrecidiv, metastas och bröstcancerspecifik död.

I **Studie IV** fann vi att den lokaliserade tätheten på platsen för framtida cancer ofta skilde sig åt jämfört med den generella tätheten i bröstet som är det mått som brukar användas. Vi undersökte effekten av den lokaliserade tätheten, och upptäckte att den var starkt kopplad till risken att ha en stor cancer vid diagnos. Den var också kopplad till risken för intervallcancer bland de mindre aggressiva bröstcancrar som var lymfkörtelnegativa. Dessa associationer existerande även efter att vi tagit hänsyn till skillnader i generell täthet.

Sammanfattningsvis har vi identifierat nya riskfaktorer för fördröjd upptäckt av bröstcancer, vilka bör valideras i framtida studier där riskstratifierad screening testas.

Acknowledgements

Funding for the research reported in this thesis has been received from: the Swedish Research Council, FORTE, the Swedish Cancer Society, the Stockholm County Council, the Cancer Society in Stockholm, and the Karolinska University Hospital.

I am honoured by the numerous people that have generously contributed to my work on this thesis. My sincerest gratitude goes to all of you. To mention a few:

Kamila Czene, my main supervisor. For always encouraging me to go that extra mile, and for sharing her vast research experience. I feel truly lucky to have had such a wonderful supervisor.

Edward Azavedo, my co-supervisor and breast radiology guru. For inspiring me to start out in the clinical field of breast radiology, and for always having a smile near at hand.

Keith Humphreys, my co-supervisor. For being a dedicated statistical expert always reading my manuscripts very thoroughly and sharing his valuable knowledge.

Per Hall, my co-supervisor and the one who introduced me to the wonderful research culture at MEB. For always giving great feedback from the clinical perspective on my, sometimes, odd ideas.

John Shepherd, my unofficial mentor and Californian liaison. For his contagious enthusiasm, for always making me feel welcome in San Francisco and for great research discussions.

Mikael Eriksson, fellow PhD student and density estimation wizard. For not getting (too) tired of my plentiful requests and questions on density data, and for nice lunches.

Johanna Holm, fellow former PhD student. For always making time for discussing the many aspects of our research, and for her carefulness in preparing cohort data.

Emilio Ugalde, fellow PhD student. For being such an all-around great guy, engaged in scientific and philosophical conversations during many lunches and evening events.

Haomin Yang, fellow PhD student. For his thoughtful research perspectives and nice dinner conversations.

David Nordemar, great radiologist and unfailing friend. For his unremitting engagement in hours-long conversations on radiological and non-radiological topics alike.

Lasse Wilhelmsson, most dedicated general practictioner. For always offering valuable practical feedback and for decades of the best friendship.

Fabian Arnberg, **Fredrik Jäderling**, and **Peter Lindholm**, for being great friends and clinical colleagues as well as deeply interested in research and always having time for encouraging talks. Fellow radiologists at the breast imaging department, not yet mentioned: **Gunilla Svane**, **Malin Bygdeson**, **Gabriela Iliescu**, **Zlatan Alagic**, **Malin Laurell Lövefors**, **and newer colleagues,** for their dedication to our patients and to making our department such a great place to work. And to the hard-working **Magnus Tengvar** and **Anders Karlsson** for introducing me to the wonderful world of Breast MRI. **Current and former bosses at the Radiology department** for supporting my absence from clinical work to perform research.

Sophia Zackrisson, **Hanna Sartor** and **Olle Bruér** for their hospitality and willingness to facilitate my data collection work in Malmö.

The leadership of the **Swedish Society of Breast Radiology** that has given me opportunities to share research findings with colleagues from around Sweden: **Karin Leifland**, **Edward Azavedo**, **Astrid Rocchi**, **Shahin Abdsaleh**, and others not mentioned but not forgotten.

Co-authors not yet mentioned, for your highly valuable contributions: **Jingmei Li**, **Therese ML Andersson**, **Sven Törnberg**, **Abbas Cheddad** and **Roxanna Hellgren**.

Maya Alsheh Ali for being so helpful in contributing her image analysis expertise. **Shadi Azam**, **Wei He**, **Marike Gabrielson** and **Mattias Hammarström** for good company and plentiful KARMA discussions. **José Tapia** and **Aki Tuuliainen**, and **all data collectors** and **prior researchers** involved in the CAHRES, LIBRO-1 and KARMA studies. My many colleagues and co-workers at the department of Medical Epidemiology and Biostatistics, and not least **Camilla Ahlqvist** our reliable educational administrator.

All smart, funny and friendly **class mates in the Forskarskolan i Epidemiologi för Kliniker**.

My half-time committee, for valuable feedback: **Theodoros Foukakis**, **Pär Sparén** and **Marie Reilly**.

The Stockholm patient association **Amazonas** for the opportunity to share my research perspectives, and for being so encouraging for me to keep on striving.

My dedicated research colleagues at KTH, **Kevin Smith**, **Hossein Azizpour** and **Yue Liu** with whom I will continue to explore the realms of deep learning.

**

My children, **Teodor** and **Ingrid**, who are the true inspiration of my life. And **Karin**, for taking wonderful care of them, and for being a discerning speaking partner at all times.

**

References

1. Javed A, Lteif A. Development of the human breast. Seminars in plastic surgery. 2013;27(1):5-12.
2. Harris JR, Lippman ME, Osborne CK, Morrow M. Diseases of the Breast: Lippincott Williams & Wilkins; 2012.
3. Shore AN, Rosen JM. Regulation of mammary epithelial cell homeostasis by lncRNAs(). The international journal of biochemistry & cell biology. 2014;54:318-30.
4. Wellings SR, Jensen HM. On the origin and progression of ductal carcinoma in the human breast. Journal of the National Cancer Institute. 1973;50(5):1111-8.
5. Ferguson DJ, Anderson TJ. Morphological evaluation of cell turnover in relation to the menstrual cycle in the "resting" human breast. British journal of cancer. 1981;44(2):177-81.
6. Ferguson JE, Schor AM, Howell A, Ferguson MW. Changes in the extracellular matrix of the normal human breast during the menstrual cycle. Cell and tissue research. 1992;268(1):167-77.
7. Fowler PA, Casey CE, Cameron GG, Foster MA, Knight CH. Cyclic changes in composition and volume of the breast during the menstrual cycle, measured by magnetic resonance imaging. BJOG: An International Journal of Obstetrics & Gynaecology. 1990;97(7):595-602.
8. Russo J, Rivera R, Russo IH. Influence of age and parity on the development of the human breast. Breast cancer research and treatment. 1992;23(3):211-8.
9. Pasqualini JR, Chetrite G, Blacker C, Feinstein MC, Delalonde L, Talbi M, et al. Concentrations of estrone, estradiol, and estrone sulfate and evaluation of sulfatase and aromatase activities in pre- and postmenopausal breast cancer patients. The Journal of clinical endocrinology and metabolism. 1996;81(4):1460-4.
10. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486(7403):400-4.
11. Nik-Zainal S, Van Loo P, Wedge David C, Alexandrov Ludmil B, Greenman Christopher D, Lau King W, et al. The Life History of 21 Breast Cancers. Cell. 2012;149(5):994-1007.
12. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
13. Conklin MW, Keely PJ. Why the stroma matters in breast cancer: insights into breast cancer patient outcomes through the examination of stromal biomarkers. Cell adhesion & migration. 2012;6(3):249-60.
14. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol. 2006;24(14):2137-50.
15. Statistics on Cancer Incidence 2016: The National Board of Health and Welfare; 2017 [Available from: http://www.socialstyrelsen.se/publikationer2017/2017-12-31.
16. Plevritis SK, Munoz D, Kurian AW, Stout NK, Alagoz O, Near AM, et al. Association of Screening and Treatment With Breast Cancer Mortality by Molecular Subtype in US Women, 2000-2012. Jama. 2018;319(2):154-64.

17. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. The New England journal of medicine. 2000;343(2):78-85.
18. Buell P. Changing incidence of breast cancer in Japanese-American women. Journal of the National Cancer Institute. 1973;51(5):1479-83.
19. Schneider R. Comparison of age, sex, and incidence rates in human and canine breast cancer. Cancer. 1970;26(2):419-26.
20. Clemons M, Goss P. Estrogen and the risk of breast cancer. New England Journal of Medicine. 2001;344(4):276-85.
21. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiologic reviews. 1993;15(1):36-47.
22. Kahlenborn C, Modugno F, Potter DM, Severs WB. Oral contraceptive use as a risk factor for premenopausal breast cancer: a meta-analysis. Mayo Clinic proceedings. 2006;81(10):1290-302.
23. Narod SA. Hormone replacement therapy and the risk of breast cancer. Nature reviews Clinical oncology. 2011;8(11):669-76.
24. Longnecker MP. Alcoholic beverage consumption in relation to risk of breast cancer: meta-analysis and review. Cancer causes & control : CCC. 1994;5(1):73-82.
25. Thune I, Brenn T, Lund E, Gaard M. Physical activity and the risk of breast cancer. The New England journal of medicine. 1997;336(18):1269-75.
26. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. The New England journal of medicine. 1985;312(3):146-51.
27. Cleary MP, Maihle NJ. The role of body mass index in the relative risk of developing premenopausal versus postmenopausal breast cancer. Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine (New York, NY). 1997;216(1):28-43.
28. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer. 1976;37(5):2486-92.
29. Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. Journal of the National Cancer Institute. 1995;87(9):670-5.
30. 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. American journal of human genetics. 2003;72(5):1117-30.
31. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nature genetics. 2015;47(4):373-80.
32. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. Nature. 2017;551(7678):92-4.
33. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nature genetics. 2013;45(4):353-61, 61e1-2.

34. Saadatmand S, Bretveld R, Siesling S, Tilanus-Linthorst MM. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients. BMJ (Clinical research ed). 2015;351:h4901.
35. Sobin LH, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours: John Wiley & Sons; 2011.
36. Martinez V, Azzopardi JG. Invasive lobular carcinoma of the breast: incidence and variants. Histopathology. 1979;3(6):467-88.
37. Berg WA, Gutierrez L, NessAiver MS, Carter WB, Bhargavan M, Lewis RS, et al. Diagnostic Accuracy of Mammography, Clinical Examination, US, and MR Imaging in Preoperative Assessment of Breast Cancer. Radiology. 2004;233(3):830-49.
38. Silverstein MJ, Lewinsky BS, Waisman JR, Gierson ED, Colburn WJ, Senofsky GM, et al. Infiltrating lobular carcinoma. Is it different from infiltrating duct carcinoma? Cancer. 1994;73(6):1673-7.
39. Sørlie 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. Proceedings of the National Academy of Sciences. 2001;98(19):10869-74.
40. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. Journal of the National Cancer Institute. 2009;101(10):736-50.
41. Bröstcancer - Nationellt Vårdprogram. 2018.
42. Yaffe MJ. Basic Physics of Digital Mammography. In: Bick U, Diekmann F, editors. Digital Mammography. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010. p. 1-11.
43. Johns PC, Yaffe MJ. X-ray characterisation of normal and neoplastic breast tissues. Physics in medicine and biology. 1987;32(6):675-95.
44. Feig SA. Breast masses. Mammographic and sonographic evaluation. Radiologic clinics of North America. 1992;30(1):67-92.
45. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic Performance of Digital versus Film Mammography for Breast-Cancer Screening. New England Journal of Medicine. 2005;353(17):1773-83.
46. McCormack V, dos Santos S. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006;15:1159 - 69.
47. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. The New England journal of medicine. 2007;356(3):227-36.
48. Ciatto S, Visioli C, Paci E, Zappa M. Breast density as a determinant of interval cancer at mammographic screening. British journal of cancer. 2004;90(2):393-6.
49. Weigel S, Heindel W, Heidrich J, Hense HW, Heidinger O. Digital mammography screening: sensitivity of the programme dependent on breast density. European radiology. 2017;27(7):2744-51.
50. Rice MS, Tamimi RM, Bertrand KA, Scott CG, Jensen MR, Norman AD, et al. Does mammographic density mediate risk factor associations with breast cancer? An analysis by tumor characteristics. Breast cancer research and treatment. 2018.

51. Byrne C, Ursin G, Martin CF, Peck JD, Cole EB, Zeng D, et al. Mammographic Density Change With Estrogen and Progestin Therapy and Breast Cancer Risk. Journal of the National Cancer Institute. 2017;109(9).
52. Tehranifar P, Rodriguez CB, April-Sanders A, Desperito E, Schmitt KM. Migration history, language acculturation and mammographic breast density. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2018.
53. Tabar L, Dean PB. Mammographic parenchymal patterns. Risk indicator for breast cancer? Jama. 1982;247(2):185-9.
54. 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 research : BCR. 2012;14(4):R114.
55. Sovio U, Li J, Aitken Z, Humphreys K, Czene K, Moss S, et al. Comparison of fully and semi-automated area-based methods for measuring mammographic density and predicting breast cancer risk. British journal of cancer. 2014;110(7):1908-16.
56. Eriksson M, Li J, Leifland K, Czene K, Hall P. A comprehensive tool for measuring mammographic density changes over time. Breast cancer research and treatment. 2018.
57. Highnam RP, Brady JM, Shepstone BJ. Estimation of compressed breast thickness during mammography. Br J Radiol. 1998;71(846):646-53.
58. Acr. 2013 ACR BI-RADS Atlas: Breast Imaging Reporting and Data System: American College of Radiology; 2014.
59. 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 epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2005;14(2):343-9.
60. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. AJR Am J Roentgenol. 2012;198(3):W292-5.
61. Eng A, Gallant Z, Shepherd J, McCormack V, Li J, Dowsett M, et al. Digital mammographic density and breast cancer risk: a case-control study of six alternative density assessment methods. Breast cancer research : BCR. 2014;16(5):439.
62. Hellmann SS, Lynge E, Schwartz W, Vejborg I, Njor SH. Mammographic density in birth cohorts of Danish women: a longitudinal study. BMC cancer. 2013;13:409.
63. Lokate M, Stellato RK, Veldhuis WB, Peeters PH, van Gils CH. Age-related changes in mammographic density and breast cancer risk. American journal of epidemiology. 2013;178(1):101-9.
64. Krishnan K, Baglietto L, Stone J, Simpson JA, Severi G, Evans CF, et al. Longitudinal Study of Mammographic Density Measures That Predict Breast Cancer Risk. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2017;26(4):651-60.
65. Boyd N, Martin L, Stone J, Little L, Minkin S, Yaffe M. A longitudinal study of the effects of menopause on mammographic features. Cancer epidemiology, biomarkers &

prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2002;11(10 Pt 1):1048-53.
66. Hutson SW, Cowen PN, Bird CC. Morphometric studies of age related changes in normal human breast and their significance for evolution of mammary cancer. Journal of clinical pathology. 1985;38(3):281-7.
67. Hovhannisyan G, Chow L, Schlosser A, Yaffe MJ, Boyd NF, Martin LJ. Differences in measured mammographic density in the menstrual cycle. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2009;18(7):1993-9.
68. Ursin G, Parisky YR, Pike MC, Spicer DV. Mammographic density changes during the menstrual cycle. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2001;10(2):141-2.
69. Walker K, Fletcher O, Johnson N, Coupland B, McCormack VA, Folkerd E, et al. Premenopausal mammographic density in relation to cyclic variations in endogenous sex hormone levels, prolactin, and insulin-like growth factors. Cancer Res. 2009;69(16):6490-9.
70. Astley SM, Harkness EF, Sergeant JC, Warwick J, Stavrinos P, Warren R, et al. A comparison of five methods of measuring mammographic density: a case-control study. Breast cancer research : BCR. 2018;20(1):10.
71. Kerlikowske K, Ma L, Scott CG, Mahmoudzadeh AP, Jensen MR, Sprague BL, et al. Combining quantitative and qualitative breast density measures to assess breast cancer risk. Breast cancer research : BCR. 2017;19(1):97.
72. Gastounioti A, Conant EF, Kontos D. Beyond breast density: a review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. Breast cancer research : BCR. 2016;18(1):91.
73. Wang C, Brentnall AR, Cuzick J, Harkness EF, Evans DG, Astley S. A novel and fully automated mammographic texture analysis for risk prediction: results from two case-control studies. Breast cancer research : BCR. 2017;19(1):114.
74. Zheng Y, Keller BM, Ray S, Wang Y, Conant EF, Gee JC, et al. Parenchymal texture analysis in digital mammography: A fully automated pipeline for breast cancer risk assessment. Med Phys. 2015;42(7):4149-60.
75. Winkel RR, von Euler-Chelpin M, Nielsen M, Petersen K, Lillholm M, Nielsen MB, et al. Mammographic density and structural features can individually and jointly contribute to breast cancer risk assessment in mammography screening: a case-control study. BMC cancer. 2016;16:414.
76. Malkov S, Shepherd JA, Scott CG, Tamimi RM, Ma L, Bertrand KA, et al. Mammographic texture and risk of breast cancer by tumor type and estrogen receptor status. Breast cancer research : BCR. 2016;18(1):122.
77. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med. 2003;138(3):168-75.
78. Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH, et al. Digital tomosynthesis in breast imaging. Radiology. 1997;205(2):399-406.

79. Skaane P, Sebuodegard S, Bandos AI, Gur D, Osteras BH, Gullien R, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. Breast cancer research and treatment. 2018.
80. Lang K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmo Breast Tomosynthesis Screening Trial, a population-based study. European radiology. 2016;26(1):184-90.
81. Conant EF, Beaber EF, Sprague BL, Herschorn SD, Weaver DL, Onega T, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. Breast cancer research and treatment. 2016;156(1):109-16.
82. Conant EF, Keller BM, Pantalone L, Gastounioti A, McDonald ES, Kontos D. Agreement between Breast Percentage Density Estimations from Standard-Dose versus Synthetic Digital Mammograms: Results from a Large Screening Cohort Using Automated Measures. Radiology. 2017;283(3):673-80.
83. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. Jama. 2012;307(13):1394-404.
84. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women with Average Risk of Breast Cancer. Radiology. 2017;283(2):361-70.
85. Grady I, Chanisheva N, Vasquez T. The Addition of Automated Breast Ultrasound to Mammography in Breast Cancer Screening Decreases Stage at Diagnosis. Acad Radiol. 2017;24(12):1570-4.
86. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. The Cochrane database of systematic reviews. 2013;6:Cd001877.
87. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ. Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. Eur J Cancer. 2013;49(16):3517-25.
88. Independent UKPoBCS. The benefits and harms of breast cancer screening: an independent review. Lancet (London, England). 2012;380(9855):1778-86.
89. Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer. 1995;75(10):2507-17.
90. Duffy SW, Day NE, Tabar L, Chen HH, Smith TC. Markov models of breast tumor progression: some age-specific results. Journal of the National Cancer Institute Monographs. 1997(22):93-7.
91. Otten JD, van Schoor G, Peer PG, den Heeten GJ, Holland R, Broeders MJ, et al. Growth rate of invasive ductal carcinomas from a screened 50-74-year-old population. Journal of medical screening. 2018;25(1):40-6.
92. Tabar L, Vitak B, Chen HH, Prevost TC, Duffy SW. Update of the Swedish Two-County Trial of breast cancer screening: histologic grade-specific and age-specific results. Swiss surgery = Schweizer Chirurgie = Chirurgie suisse = Chirurgia svizzera. 1999;5(5):199-204.

93. 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. Journal of medical screening. 2000;7(1):14-8.
94. Lind H, Svane G, Kemetli L, Tornberg S. Breast Cancer Screening Program in Stockholm County, Sweden - Aspects of Organization and Quality Assurance. Breast care (Basel, Switzerland). 2010;5(5):353-7.
95. Sickles EA, Wolverton DE, Dee KE. Performance parameters for screening and diagnostic mammography: specialist and general radiologists. Radiology. 2002;224(3):861-9.
96. Roelofs AA, Karssemeijer N, Wedekind N, Beck C, van Woudenberg S, Snoeren PR, et al. Importance of comparison of current and prior mammograms in breast cancer screening. Radiology. 2007;242(1):70-7.
97. Thurfjell MG, Vitak B, Azavedo E, Svane G, Thurfjell E. Effect on sensitivity and specificity of mammography screening with or without comparison of old mammograms. Acta radiologica (Stockholm, Sweden : 1987). 2000;41(1):52-6.
98. Taylor-Phillips S, Jenkinson D, Stinton C, Wallis MG, Dunn J, Clarke A. Double Reading in Breast Cancer Screening: Cohort Evaluation in the CO-OPS Trial. Radiology. 2018:171010.
99. Kerlikowske K. Evidence-based breast cancer prevention: the importance of individual risk. Annals of internal medicine. 2009;151(10):750-2.
100. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. Jama. 2009;302(15):1685-92.
101. Shieh Y, Eklund M, Madlensky L, Sawyer S, Thompson C, Stover FA, et al. Breast Cancer Screening in the Precision Medicine Era: Risk-Based Screening in a Population-Based Trial. Journal of the National Cancer Institute. 2017;109(5).
102. Eriksson M, Czene K, Pawitan Y, Leifland K, Darabi H, Hall P. A clinical model for identifying the short-term risk of breast cancer. Breast cancer research : BCR. 2017;19(1):29.
103. Kerlikowske K, Zhu W, Tosteson AN, Sprague BL, Tice JA, Lehman CD, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. Ann Intern Med. 2015;162(10):673-81.
104. Onega T, Beaber EF, Sprague BL, Barlow WE, Haas JS, Tosteson AN, et al. Breast cancer screening in an era of personalized regimens: a conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. Cancer. 2014;120(19):2955-64.
105. Health ECD-Gf, Protection C, Perry N, Puthaar E. European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis: Office for Official Publications of the European Communities; 2006.
106. Tornberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Seradour B, et al. A pooled analysis of interval cancer rates in six European countries. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2010;19(2):87-93.
107. Houssami N, Hunter K. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. NPJ Breast Cancer. 2017;3:12.
108. Barton MB, Elmore JG, Fletcher SW. Breast symptoms among women enrolled in a health maintenance organization: frequency, evaluation, and outcome. Ann Intern Med. 1999;130(8):651-7.

109. Hofvind S, Geller B, Skaane P. Mammographic features and histopathological findings of interval breast cancers. Acta radiologica (Stockholm, Sweden : 1987). 2008;49(9):975-81.
110. Duncan AA, Wallis MG. Classifying interval cancers. Clin Radiol. 1995;50(11):774-7.
111. Burnside ES, Vulkan D, Blanks RG, Duffy SW. Association between Screening Mammography Recall Rate and Interval Cancers in the UK Breast Cancer Service Screening Program: A Cohort Study. Radiology. 2018:171539.
112. van Gils CH, Otten JD, Verbeek AL, Hendriks JH. Mammographic breast density and risk of breast cancer: masking bias or causality? European journal of epidemiology. 1998;14(4):315-20.
113. Boyd N, Guo H, Martin L, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. The New England journal of medicine. 2007;356:227 - 36.
114. Pollán M, Ascunce N, Ederra M, Murillo A, Erdozáin N, Alés-Martínez JE, et al. Mammographic density and risk of breast cancer according to tumor characteristics and mode of detection: a Spanish population-based case-control study. Breast cancer research : BCR. 2013;15(1):R9.
115. Lowery JT, Byers T, Hokanson JE, Kittelson J, Lewin J, Risendal B, et al. Complementary approaches to assessing risk factors for interval breast cancer. Cancer Causes & Control. 2011;22(1):23-31.
116. Blanch J, Sala M, Ibáñez J, Domingo L, Fernandez B, Otegi A, et al. Impact of Risk Factors on Different Interval Cancer Subtypes in a Population-Based Breast Cancer Screening Programme. PloS one. 2014;9(10):e110207.
117. Holm J, Humphreys K, Li J, Ploner A, Cheddad A, Eriksson M, et al. Risk factors and tumor characteristics of interval cancers by mammographic density. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(9):1030-7.
118. Li J, Holm J, Bergh J, Eriksson M, Darabi H, Lindstrom LS, et al. Breast cancer genetic risk profile is differentially associated with interval and screen-detected breast cancers. Annals of Oncology. 2015;26(3):517-22.
119. Hofvind S, Sagstad S, Sebuodegard S, Chen Y, Roman M, Lee CI. Interval Breast Cancer Rates and Histopathologic Tumor Characteristics after False-Positive Findings at Mammography in a Population-based Screening Program. Radiology. 2018;287(1):58-67.
120. Kalager M, Tamimi RM, Bretthauer M, Adami HO. Prognosis in women with interval breast cancer: population based observational cohort study. Br Med J. 2012;345:10.
121. Eriksson L, Czene K, Rosenberg LU, Tornberg S, Humphreys K, Hall P. Mammographic density and survival in interval breast cancers. Breast Cancer Research. 2013;15(3):7.
122. Tot T, Gere M. Radiological-pathological correlation in diagnosing breast carcinoma: the role of pathology in the multimodality era. Pathology oncology research : POR. 2008;14(2):173-8.
123. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Annals of surgical oncology. 2010;17(6):1471-4.
124. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness. The New England journal of medicine. 2016;375(15):1438-47.

125. Eriksson L, Czene K, Rosenberg L, Humphreys K, Hall P. The influence of mammographic density on breast tumor characteristics. Breast cancer research and treatment. 2012;134(2):859-66.
126. Robinson B, Currie M, Phillips E, Dachs G, Strother M, Morrin H, et al. Body mass index (BMI): association with clinicopathological factors and outcome of women with newly diagnosed breast cancer in New Zealand. The New Zealand medical journal. 2017;130(1451):46-56.
127. Carmichael AR, Bates T. Obesity and breast cancer: a review of the literature. Breast (Edinburgh, Scotland). 2004;13(2):85-92.
128. 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.
129. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, et al. Predicting the survival of patients with breast carcinoma using tumor size. Cancer. 2002;95(4):713-23.
130. Magnusson C, Baron J, Persson I, Wolk A, Bergstrom R, Trichopoulos D, et al. Body size in different periods of life and breast cancer risk in post-menopausal women. International journal of cancer Journal international du cancer. 1998;76(1):29-34.
131. Gabrielson M, Eriksson M, Hammarström M, Borgquist S, Leifland K, Czene K, et al. Cohort profile: The Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA). International journal of epidemiology. 2017:dyw357-dyw.
132. Mattsson B, Wallgren A. Completeness of the Swedish cancer register non-notified cancer cases recorded on death certificates in 1978. Acta Radiologica: Oncology. 1984;23(5):305-13.
133. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. Physics in medicine and biology. 1994;39(10):1629-38.
134. Cheddad A, Czene K, Eriksson M, Li J, Easton D, Hall P, et al. Area and Volumetric Density Estimation in Processed Full-Field Digital Mammograms for Risk Assessment of Breast Cancer. PloS one. 2014;9(10):e110690.
135. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Annals of Oncology. 2013;24(9):2206-23.
136. Porta M. A Dictionary of Epidemiology.
137. Moshina N, Roman M, Waade GG, Sebuodegard S, Ursin G, Hofvind S. Breast compression parameters and mammographic density in the Norwegian Breast Cancer Screening Programme. European radiology. 2018;28(4):1662-72.
138. Morrow M, Chatterton RT, Jr., Rademaker AW, Hou N, Jordan VC, Hendrick RE, et al. A prospective study of variability in mammographic density during the menstrual cycle. Breast cancer research and treatment. 2010;121(3):565-74.
139. Banik S, Rangayyan RM, Desautels JE. Detection of architectural distortion in prior mammograms of interval-cancer cases with neural networks. Conf Proc IEEE Eng Med Biol Soc. 2009;2009:6667-70.
140. Tan M, Zheng B, Ramalingam P, Gur D. Prediction of near-term breast cancer risk based on bilateral mammographic feature asymmetry. Acad Radiol. 2013;20(12):1542-50.

141. Maehle BO, Tretli S, Skjaerven R, Thorsen T. Premorbid body weight and its relations to primary tumour diameter in breast cancer patients; its dependence on estrogen and progesteron receptor status. Breast cancer research and treatment. 2001;68(2):159-69.
142. Reeves MJ, Newcomb PA, Remington PL, Marcus PM, MacKenzie WR. Body mass and breast cancer. Relationship between method of detection and stage of disease. Cancer. 1996;77(2):301-7.
143. Musolino A, Falcini F, Sikokis A, Boggiani D, Rimanti A, Pellegrino B, et al. Prognostic impact of interval breast cancer detection in women with pT1a N0M0 breast cancer with HER2-positive status: Results from a multicentre population-based cancer registry study. European journal of cancer (Oxford, England : 1990). 2018;88:10-20.
144. Olsson A, Garne JP, Tengrup I, Zackrisson S, Manjer J. Body mass index and breast cancer survival in relation to the introduction of mammographic screening. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology. 2009;35(12):1261-7.
145. van der Waal D, Verbeek ALM, Broeders MJM. Breast density and breast cancer-specific survival by detection mode. BMC cancer. 2018;18(1):386.
146. Holland K, van Gils CH, Mann RM, Karssemeijer N. Quantification of masking risk in screening mammography with volumetric breast density maps. Breast cancer research and treatment. 2017;162(3):541-8.
147. Duffy SW, Nagtegaal ID, Wallis M, Cafferty FH, Houssami N, Warwick J, et al. Correcting for Lead Time and Length Bias in Estimating the Effect of Screen Detection on Cancer Survival. American journal of epidemiology. 2008;168(1):98-104.
148. Walter SD, Day NE. Estimation of the duration of a pre-clinical disease state using screening data. American journal of epidemiology. 1983;118(6):865-86.
149. Shen Y, Zelen M. Screening Sensitivity and Sojourn Time From Breast Cancer Early Detection Clinical Trials: Mammograms and Physical Examinations. J Clin Oncol. 2001;19(15):3490-9.
150. Chan DS, Vieira AR, Aune D, Bandera EV, Greenwood DC, McTiernan A, et al. Body mass index and survival in women with breast cancer-systematic literature review and meta-analysis of 82 follow-up studies. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2014;25(10):1901-14.
151. Giger ML. Machine Learning in Medical Imaging. Journal of the American College of Radiology : JACR. 2018;15(3 Pt B):512-20.
152. Gastounioti A, Oustimov A, Hsieh MK, Pantalone L, Conant EF, Kontos D. Using Convolutional Neural Networks for Enhanced Capture of Breast Parenchymal Complexity Patterns Associated with Breast Cancer Risk. Acad Radiol. 2018.
153. Houssami N, Lee CI, Buist DSM, Tao D. Artificial intelligence for breast cancer screening: Opportunity or hype? Breast (Edinburgh, Scotland). 2017;36:31-3.