Thesis for doctoral degree (Ph.D.)
2021

Risk assessment and prevention of breast cancer

Mikael Eriksson

Karolinska Institutet

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

Risk assessment and prevention of breast cancer

Mikael Eriksson

Karolinska Institutet

Stockholm 2021

In this thesis I investigate dual ideas for improving mammography screening and prevention of breast cancer and how they better can complement each other to improve the health of women.

Image: Elmer Laahne

Published by Karolinska Institutet.
Printed by Karolinska Institutet.

ISBN 987-91-8016-039-1

RISK ASSESSMENT AND PREVENTION OF BREAST CANCER
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Mikael Eriksson

Principal Supervisor:
Prof. Per Hall
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Co-supervisor(s):
Prof. Kamila Czene
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Opponent:
Prof. Jeff Tice
University of California, San Francisco
Department of School of Medicine

Examination Board:
Prof. Johan Askling
Karolinska Insitutet
Department of Medicine Solna

Assoc. prof. Nicola Orsini
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Prof. Laszlo Tabár
Uppsala University
Department of Medicine

to N&M

ABSTRACT

One woman in eight develops breast cancer during her lifetime in the Western world. Measures are warranted to reduce mortality and to prevent breast cancer. Mammography screening reduces mortality by early detection. However, approximately one fourth of the women who develop breast cancer are diagnosed within two years after a negative screen. There is a need to identify the short-term risk of these women to better guide clinical follow-up. Another drawback of mammography screening is that it focuses on early detection only and not on breast cancer prevention. Today, it is known that women attending screening can be stratified into high and low risk of breast cancer. Women at high risk could be offered preventive measures such as low-dose tamoxifen to reduce breast cancer incidence. Women at low risk do not benefit from screening and could be offered less frequent screening.

In study I, we developed and validated the mammographic density measurement tool STRATUS to enable mammogram resources at hospitals for large scale epidemiological studies on risk, masking, and therapy response in relation to breast cancer. STRATUS showed similar measurement results on different types of mammograms at different hospitals. Longitudinal studies on mammographic density could also be analysed more accurate with less non-biological variability.

In study II, we developed and validated a short-term risk model based on mammographic features (mammographic density, microcalcifications, masses) and differences in occurrences of mammographic features between left and right breasts. The model could optionally be expanded with lifestyle factors, family history of breast cancer, and genetic determinants. Based on the results, we showed that among women with a negative mammography screen, the short-term risk tool was suitable to identify women that developed breast cancer before or at next screening. We also showed that traditional long-term risk models were less suitable to identify the women who in a short time-period after risk assessment were diagnosed with breast cancer.

In study III, we performed a phase II trial to identify the lowest dose of tamoxifen that could reduce mammographic density, an early marker for reduced breast cancer risk, to the same extent as standard 20 mg dose but cause less side-effects. We identified 2.5 mg tamoxifen to be non-inferior for reducing mammographic density. The women who used 2.5 mg tamoxifen also reported approximately 50% less severe vasomotor side-effects.

In study IV, we investigated the use of low-dose tamoxifen for an additional clinical use case to increase screening sensitivity through its effect on reducing mammographic density. It was shown that 24% of the interval cancers have a potential to be detected at prior screen.

In conclusion, tools were developed for assessing mammographic density and breast cancer risk. In addition, two low-dose tamoxifen concepts were developed for breast cancer prevention and improved screening sensitivity. Clinical prospective validation is further needed for the risk assessment tool and the low-dose tamoxifen concepts for the use in breast cancer prevention and for reducing breast cancer mortality.

LIST OF SCIENTIFIC PAPERS

I. **Mikael Eriksson**, Jingmei Li, Karin Leifland, Kamila Czene, Per Hall
A comprehensive tool for measuring mammographic density changes over time
Breast Cancer Research and Treatment 2018, doi:10.1007/s10549-018-4690-5

II. **Mikael Eriksson**, Kamila Czene, Fredrik Strand, Sophia Zackrisson, Peter Lindholm, Kristina Lång, Daniel Förnvik, Hanna Sartor, Nasim Mavaddat, Doug Easton, Per Hall
Identification of women at high risk of breast cancer who need supplemental screening
Radiology 2020, doi:0.1148/radiol.2020201620

III. **Mikael Eriksson**, Martin Eklund, Signe Borgquist, Roxanna Hellgren, Sara Margolin, Linda Thoren, Ann Rosendahl, Kristina Lång, José Tapia, Magnus Bäcklund, Andrea Discacciati, Alessio Crippa, Marike Gabrielson, Mattias Hammarström, Yvonne Wengström, Kamila Czene, Per Hall
Low dose tamoxifen for breast cancer prevention and mammographic density reduction – a randomized controlled trial
Submitted for publication, Journal of Clinical Oncology

IV. **Mikael Eriksson**, Kamila Czene, Emily Conant, Per Hall
Use of low-dose tamoxifen to increase screening sensitivity in mammography of premenopausal women
Manuscript

RELATED PUBLICATIONS

I. **Mikael Eriksson**, Kamila Czene, Yudi Pawitan, Karin Leifland, Hatef Darabi and Per Hall
A clinical model for identifying the short-term risk of breast cancer
Breast Cancer Research 2017, doi:10.1186/s13058-017-0820-y

II. Natalie Holowko*, **Mikael Eriksson***, Ralf Kuja-Halkola, Shadi Azam, Wei He, Per Hall, Kamila Czene
Heritability of Mammographic Breast Density, Density Change, Microcalcifications, and Masses
Cancer Res. 2020, doi:10.1158/0008-5472.CAN-19-2455

III. Shadi Azam*, **Mikael Eriksson***, Arvid Sjölander, Roxanna Hellgren, Marike Gabrielson, Kamila Czene, Per Hall
Mammographic Density Change and Risk of Breast Cancer
J Natl Cancer Inst. 2020, doi:10.1093/jnci/djz149

IV. Shadi Azam*, **Mikael Eriksson***, Arvid Sjölander, Marike Gabrielson, Roxanna Hellgren, Kamila Czene, Per Hall
Predictors of mammographic microcalcifications
Int J Cancer. 2020, doi:10.1002/ijc.33302

* Equal contributions

CONTENTS

LIST OF ABBREVIATIONS

95%CI	95 percent confidence interval
AUC	Area Under the receiver operating characteristic Curve
BC	Breast Cancer
BCAC	The Breast Cancer Association Consortium
BI-RADS	A four-category visual classification of breast composition issued by the density American College of Radiology
BMI	Body Mass Index
BRCA1/2	BReast CAncer susceptibility gene 1 or 2
CAD	Computer Aided Detection
CAHRES	Cancer and Hormone Replacement Study
cBIRADS	Computer-generated score mimicking the BI-RADS density classification
CIF	Cumulative incidence function
DNA	Deoxyribonucleic acid
Dnr	Reference number in public archives, 'Diarienummer'
EMT	Epithelial-mesenchymal transition
ER	Estrogen Receptor
FISH	Fluorescence in situ hybridization
GWAS	Genome-Wide Association Study
HER2	Human Epidermal growth factor Receptor 2
HR	Hazard Ratio
HRT	Hormone Replacement Therapy used mainly for menopausal
IC	Interval Cancer
ICD	International statistical Classification of Diseases and related health problems
IHC	Immunohistochemistry
KARMA	Karolinska mammography screening cohort for risk prediction
LIBRO1	Linné-Bröst study 1 breast cancer cohort 2001-2008 in Stockholm-Gotland region
NIH	National cancer institute, US
OR	Odds Ratio

p	The probability to obtain a result that is at least as extreme as
PD	Percent mammographic density, i.e. the percentage of radio dense tissue of total breast tissue area
PNR	Unique personal identification number
PR	Progesterone Receptor
PRS	Polygenic Risk Score
SC	Screen-detected Cancer
SEER	The Surveillance, Epidemiology, and End Results program
SNP	Single Nucleotide Polymorphisms
TNM	Tumor, affected Nodes, Metastasis classification of malignant tumors

"It's better to be approximately right than exactly wrong"

Carveth Read

1 INTRODUCTION

Thirteen percent of all women in the Western world develop breast cancer during their lifetime. This makes breast cancer the most common cancer among women, which accounts for approximately thirty percent of all female cancers [1]. Globally there are approximately 1.5 million women diagnosed with breast cancer every year and 500,000 women die from the disease. The incidence has increased over the last thirty years, while breast cancer mortality has decreased over the same period. The reasons for the increase are not well understood, but the mortality decrease is estimated to be due to mammography screening by 20% and due to improved cancer therapies by 60% [2].

In this thesis, my aim is to a) show feasibility of reducing the mortality by more than 20% and to b) show feasibility of increasing the uptake of preventive therapies in the population to reduce breast cancer incidence.

Mammography screening invites women based on their age every one-to-three years to identify cancers that are rare in the population [3, 4]. However, approximately 25% of the women develop breast cancer in between screening visits [5]. These women are a symptom that the age-based mammography screening is suboptimal. The screening could be improved by individualizing the invitations and the clinical follow-up of the women, based on the risk of breast cancer. I develop a risk tool that potentially can be used in a risk-based screening setting and I suggest how clinical follow-up could be performed.

Tamoxifen reduces breast cancer incidence by approximately 30% [6], but the uptake in the population is low and is challenged by severe side-effects [7]. In this thesis, I investigate if a low dose of tamoxifen is as efficient as the standard tamoxifen dose to reduce mammographic density but have less severe side-effects. Mammographic density reduction is a known proxy for a reduction of breast cancer incidence and, could be used early in the treatment to judge which women benefit from the therapy. A low dose of tamoxifen with less side-effects could increase the uptake in the population and therefore reduce breast cancer incidence.

In addition, I investigate if low-dose tamoxifen could be used to improve the sensitivity of a mammogram by the effect from fibro-glandular tissue reduction. Today, mammography screening is challenged by the sensitivity of screening modalities that are used to distinguish tumors from the radio dense healthy tissue. Approximately fifty percent of the breast cancers are missed in screening in the group of women with extremely dense breasts. Low-dose tamoxifen lowers mammographic density and could potentially increase the sensitivity of a mammogram. Therefore, low-dose tamoxifen could improve early detection of interval cancers. Interval cancers are known to be more aggressive and a reduction of interval cancers has the potential to reduce breast cancer mortality further.

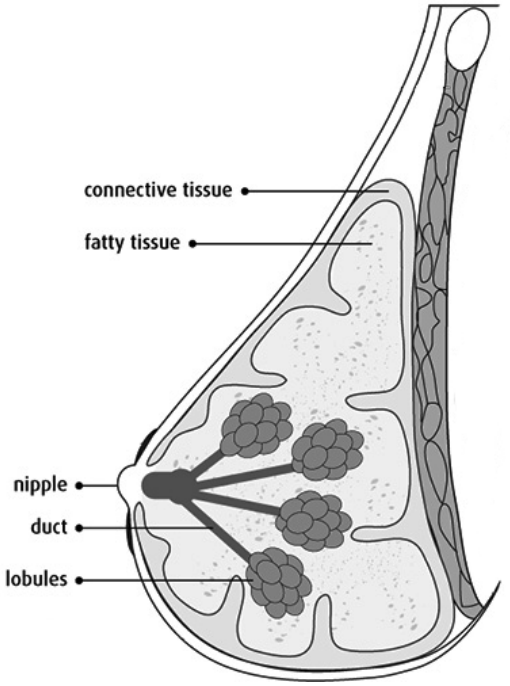
2 BACKGROUND

2.1 BREAST CANCER

2.1.1 Breast anatomy and development

The female breast is composed of lobular units that are responsible for producing milk, milk ducts for draining milk to the nipple, connective tissue (stroma), and adipose tissue [8]. The lobular terminal units are supported by the connective tissue, which gives the breast its shape.

Figure 1. Anatomic picture of the breast. American Cancer Society.



The female breast starts to develop during puberty in its first reproductive phase. The ducts elongate and is branching under the influence of oestrogen and, the lobular units develops into cellular structures including epithelial and myoepithelial cells [9].The epithelial cells are positioned in the lobular units and in the inner lining of the milk ducts. The second development phase occur during pregnancy and breast feeding [10]. The lobular units develop from no cell differentiation (type 1) to complete cell differentiation (type 4) at the end of pregnancy, which is a type that can secret milk. The final breast development phase occurs during menopause [11]. The breast involutes into mainly fatty tissue by shrinkage of the glandular tissue.

2.1.2 Rare and common genetic mutations

Breast cancer is a genetic disease that origin from the epithelial cells for almost all breast cancers. The best-known breast cancer susceptibility gene is the BRCA1 gene. BRCA1 is involved in the deoxyribonucleic acid (DNA) repair mechanism, which is effective for other genetic abnormalities such as double strand breaks that occur due to external stimuli or during DNA replication [12]. Deleterious mutations in the BRCA1 gene impairs the repair mechanism, which could lead to further carcinogenic processes at a later time during the woman's lifetime [13]. Mutations are commonly categorised by prevalence and inferred risk. High-penetrant rare variants are deleterious mutations in genes such as BRCA1, BRCA2, TP53, and PTEN genes. Medium-penetrant rare variants are mutations in CHECK2, ATM, PALB2, and BRIP1. Single nucleotide polymorphisms (SNP) belong to the third category of low-penetrance common variants with >1% frequency in the population. A SNP is a mutation where one base-pair in the DNA double-helix has been replaced with an alternate base-pair in the same double-helix position.

A polygenic risk score (PRS) is a weighted multiplicative model construct that consists of several SNPs that show an association with breast cancer outcome [14]. PRSs have been developed in the Breast Cancer Association Consortia (BCAC) consortia over the last 10 years aiming at identifying women at increased risk of breast cancer to be used in clinical practice [15]. Several PRSs have been published and described over the last years and it has been shown that women with a high PRS score more commonly have ER-positive tumors, i.e. less aggressive tumors [16].

BCAC recently extended the PRS to include 313 SNPs [17]. 305 SNPs are used in the overall breast cancer PRS, 311 SNPs for the ER-positive breast cancer score, and 196 SNPs for the ER-negative score. The consortia also developed an alternative PRS score including 3,820 SNPs that has a slight performance improvement. The PRS was developed from case-control data that origin from multiple countries and included over 100,000 cases and controls. Penalized regression was used in the later models to improve the generalizability of the PRS discrimination performance of external cohorts.

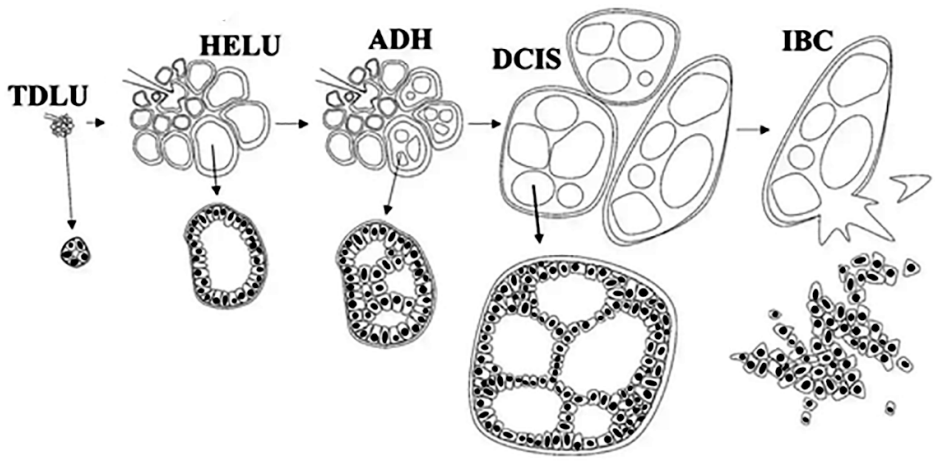
There is an interaction between family history of breast cancer and PRS. Women with a family history of breast cancer show a lower PRS. The effect of ER-positive and ER-negative PRSs are attenuated with approximately 21% and 12% respectively in women with a family history of breast cancer compared to women with no family history of breast cancer [17]. Risk models therefore use different estimates for women with and without a family history of breast cancer.

2.1.3 Cancer development

Breast cancer is believed to be initiated by exposure to various agents such as ionising radiation, virus, hormones, and spontaneous mutations. Underlying germline alterations influence the susceptibility for a cancer [18, 19]. Breast cancer carcinogenesis is a multistep process where normal cells develop to invasive cancers. In a carcinogenic progression epithelial cells initially enlarge in the terminal ducts lobular units (TDLU) in the lobes into hyperplastic enlarged lobular

units (HELU). The enlarged lobular units may progress further to atypical hyperplasia and, through further proliferation into a carcinoma in situ lesion (DCIS) [20].

Figure 2. Wellings-Jensen model of invasive breast cancer development.



Non-invasive carcinoma in situ consists of nearly 15% of newly diagnosed cancers [21]. Ductal cancer in-situ (DCIS) is a common precursor for breast carcinoma [22]. Invasion occurs when abnormal cells break through the cell barrier and spread to the surrounding. The lymph system and blood vessels could further transport the cancerous cells to form metastases in the skeleton, lungs, liver, and brain. The tumor growth is fuelled by oestrogen, progesterone, and HER2 [23]. The adenocarcinomas (epithelial based cancers) are responsible for 99% of all breast cancers [24].

2.1.4 Risk factors

Breast cancer is a complex disease with a genetic origin and with lifestyle factors that affect the progression of a genetic abnormality into a cancer. A Swedish study showed that the heritability of breast cancer is approximately 25% [25]. The lifestyle effect of the increased cancer development rates has been studied among domestic Asian populations in comparison to Asian populations living in US [26]. A study showed that the lifestyle component could induce three times increased breast cancer incidence in populations, who live in developed countries compared to women living in non-developed countries [27]. The lifestyle factors are associated with increased oestrogen and progesterone female hormones, which in turn are growth factors for developing cancers. Prior history of in-situ cancer and benign diseases also increases the risk for developing invasive breast cancer later in life [28], due to a common heritable cause [29].

Table 1. Risk factors are summaries below.

Relative risk	Risk factor
>4	Age of woman BRCA1/2 mutation carrier ship (TP53, PTEN) High mammographic density
4.0-2.0	Microcalcifications Benign breast disorders and in-situ cancer Family history of breast cancer High polygenic risk score (combined SNPs) CHECK2, ATM, PALB2, BRIP1 gene mutation carrier ship Recent and long-term use of hormone replacement therapy Nulliparity and no breastfeeding
2.0->1	Late age at first full-term pregnancy Early menarche Late menopause Postmenopausal body mass index Recent use of oral contraceptives Tallness Alcohol and tobacco consumption Physical activity

A short description of the risk factors is given below. Mammographic features are described in more detail in the separate 'Breast Imaging' section.

Age and sex

Age in females is the strongest risk factor for developing breast cancer [30]. Women above age 60 is 5 times more likely to develop breast cancer compared to women below age 60 [31]. Women with early cancer onset are more likely to have ER-negative tumors. Lifestyle factor exposures become more important for cancer onset at a later age where ER-positive tumors are also more common. The cancer incidence increases non-linearly with age and peaks at age 60 to 70. This may be partly caused by the hormonal milieu [32].

BMI

BMI affects breast cancer risk in pre- and postmenopausal women, but studies show inconsistent results in the direction of the association. A study showed increased risk in women with high BMI in both pre- and postmenopausal women [33]. Another study showed decreased risk in obese premenopausal women, but increased risk in obese postmenopausal women [34]. Large childhood body size has been shown to infer a reduced breast cancer risk in both pre- and postmenopausal women [35].

Age at menarche

Earlier age at menarche increases the risk for breast cancer later in life [36]. The mechanism is probably a prolonged exposure to female sex hormone [37]. The risk is higher for developing ER-positive cancers and BRCA1 mutated tumors compared to ER-negative and BRCA2 mutated cancers. The elevated risk for specific subtypes could be caused by a breast type differentiation earlier in life [38, 39].

Oral contraceptives

Current use of oral contraceptives and earlier onset increases the breast cancer risk [40]. Later oral contraceptive leads to a lower risk due to lower hormone doses.

Parity

Number of children decreases the risk of ER-positive cancers [41]. HER2-positive and triple negative cancers are not associated with parity [39]. Neither are BRCA1/2 mutated cancers [42].

Age at first childbirth

Older age at first childbirth increases the risk for ER-positive breast cancer [37]. HER2-cancers and triple negative cancers are not associated with the woman's age at first childbirth [41]. Studies show that BRCA1-mutated cancer are less common in women with later age at first childbirth [42].

Breastfeeding

Women who have a child and not breastfeed the child or have short breast-feeding periods have increased risk for breast cancer compared to women who breastfeed [43]. Breast feeding is protective for ER-positive and triple-negative cancers. HER2-positive cancers are not associated with breast feeding [38].

Hormonal replacement therapy

Women using oestrogen-progesterone based hormonal replacement therapy (HRT) or oestrogen-only HRT have an increased risk for breast cancer up to two years following HRT treatment [44, 45]. Alternative HRT treatment including phytoestrogens is not associated with breast cancer risk [46]. HRT increases mammographic density, but it is unclear whether phytoestrogens affect density.

Menopause

Menopause is defined as the time in time when menstrual periods has stopped for the last 12 months. Menopause often occurs close to age 50 [47] by a reduction of oestrogen and progesterone production in the ovaries [48]. Women who have an earlier menopause have a decreased risk of breast cancer [37]. Breast cancer risk is also lower in women who had hysterectomy or oophorectomy prior to natural menopause [49].

Alcohol

Alcohol increases the risk of breast cancer with 40-50% in heavy drinkers compared to non-drinkers [50, 51]. Studies show that alcohol causes ER-positive cancers to a higher extent compared to ER-negative cancers [52]. Alcohol also increases the level of mammographic density in the breast [53].

Tobacco

Cigarette smoking is measured as intake of number of cigarettes per day over one year [54]. One pack-year is 20 cigarettes per day in 1 year. Current smoking increases the risk of breast cancer with approximately 12% compared to non-smokers [54]. Smoking have been associated with both ER-positive and ER-negative cancers [55]. Studies suggest that smoking could also have an anti-oestrogen effect by impairing the ovarian functioning [56]. Smoking could alter oestrogen metabolism [57], and lower body fat [58]. Mammographic density could also be affected by smoking, but studies are non-conclusive [59].

Physical activity

Physical activity means any kind of bodily movement leading to energy expenditure [60]. Physical activity is measured in metabolic equivalent of task (MET) [61]. By sitting on a chair for 1 hour is equivalent with 1 MET hour. Physical activity is further categorized into sedentary activity, light intensity, moderate activity, and vigorous activity. Studies show that physically active pre- and postmenopausal women have lower breast cancer risk compared to less active women [62]. Physical activity reduces the absolute level of glandular tissue in the breast measured as absolute dense mammographic area [63, 64].

Family history of breast cancer

A family history of breast cancer in a 1st degree relative doubles the breast cancer risk for the woman herself. Women who develop breast cancer before age 50 are more likely to have a BRCA1/2 mutation [65] and an aggressive tumor. An inherited risk is also captured in polygenic risk scores (PRS) which combine the risks from multiple low-susceptibility SNPs [66]. Currently, the PRSs predict risk for ER-positive, ER-negative, and overall breast cancer risk. The largest proportion of the inherited breast cancers are still not explained by the known genetic variants [67]. A study shows that 25% of the cancers can be explained by a heritable pathway [25] in a Scandinavian population.

BRCA 1 and 2 mutation

Genetic mutations are constantly occurring during deoxyribonucleic acid (DNA) replication in the cells and, by external stimuli such as ionizing radiation, tar, virus, and alcohol that cause DNA damage [68]. The cell has mechanisms to repair such abnormalities and the most famous repair mechanism is related to the BRCA 1 protein that is transcribed and translated based on the DNA-region with the same name. It repairs DNA damages where both strands of the double-helix are broken. A mutation in the BRCA genes could cause the DNA repair mechanism to malfunction. Women with specific deleterious mutations in the BRCA genes are therefore more likely to develop breast cancer [69]. Women with malicious BRCA 1 mutation have an

approximately 70% probability to develop breast cancer at some point in time during lifetime, while BRCA 2 mutations inflict a lifetime risk of approximately 30% to develop breast cancer [70].

2.1.5 Tumor characteristics

Tumor size is initially assessed in a clinical examination or by investigating a mammogram, but is most commonly reported in registers based on the pathology report. The size is reported as the widest diameter of the tumor [71]. Tumor size is also categorized into T0 (not palpable), Tis (ductal carcinoma in-situ), T1 (<=20 mm), T2 (21-50 mm), T3 (>50 mm). The additional category T4 refers to a tumor that is attached to the chest wall or is breaking through the skin. Tumor size is one of three prognostic factors that defines the TNM classification.

Lymph nodes are positioned in the axilla area and are clumps of immune cells that act as filters in the lymphatic system [69]. A tumor in the breast most commonly spreads through the lymphatic system to the lymph nodes. Affected lymph nodes means that the cancer spread to one or more lymph nodes. Lymph node status is categorized into N0 (no regional lymph nodes metastasis), N1 (moving lymph node metastases in the axilla), and N3 (fixed lymph node metastases in the axilla). Lymph node status is one of the prognostic factors defining the TNM classification.

Metastasis refers to the distant spread of a cancer, most commonly to the brain, lungs or skeleton. M0 means that there is no known metastasis and M1 means that a metastasis has been discovered [69]. Metastasis is a highly prognostic factor and is part of defining the TNM classification.

Grade is a microcopy judgement of the abnormality of the tumor cells [72]. Grade 1 are well-differentiated tumor cells where most cells are slow-growing, Grade 3 are poor-differentiated cells where most cells are fast-growing. Grade 2 refers to that most cells are moderate differentiated, that is between grade 1 and 3.

Estrogen receptor (ER) status refers to the immunohistochemistry (IHC) classification of the percentage of cells that express estrogen receptors. Oestrogen is an important growth factor for tumor cells. ER is positive if 10% or more of the cells are positive. Progesterone receptor (PR) status refers to an IHC classification of the percentage of cells that express progesterone receptor. PR is positive if 10% or more cells are positive.

A human epidermal growth factor receptor 2 (HER2) positive tumor refers to tumor cells that have several copies of the HER2 gene, with the result of an over-expression of HER2 protein [73]. Increased levels of the HER2 protein promotes tumor cell growth. IHC staining is used as a screening technique for HER2 and, fluorescence in situ hybridization (FISH) analysis is used in addition to confirm HER2 gene amplification. HER2 is positive if at least 10% of the cells are positive and confirmed by FISH.

Ki-67 is a protein marker for cell proliferation, an antigen protein encoded by the MKI67 gene. IHC staining is used to classify Ki-67 status [74]. Ki-67 is positive if 20% or more cells are positive.

2.1.6 Diagnosis

Diagnosis of breast cancer is performed using a triple-diagnostic method [75]. The method consists of a clinical examination of the breast, imaging (e.g. digital mammography and ultrasound), and fine-needle biopsy for cytopathology diagnosis. If at least one of these examinations indicates a malignancy, the finding is treated as malignant.

BI-RADS codes

Radiologists classify their radiological findings on a seven-grade scale called BI-RADS [76, 77]. Women who receive code 3 or higher are routinely examined in further work-up. The proportion of women with code 3 or higher is in Europe approximately 3-7% and in US more than 10% [78]. Approximately 2% of the women with code 3 are diagnosed with breast cancer. Women with code 4 and 5 are 30% and 95% likely to be diagnosed with breast cancer, respectively.

Table 2. BI-RADS malignancy coding.

Code	Description
Code 0 – assessment is incomplete	The assessment was not complete, and the woman could be recommended additional work-up, with further examinations.
Code 1 – negative	No suspicious finding was found, i.e. no microcalcifications, no suspicious mass, and no asymmetrical glandular structure.
Code 2 – benign finding	An abnormal lesion was found, but it was a definitive non-malign finding.
Code 3 – probably benign finding	An abnormal lesion was found but is probably a non-malignant finding and no palpable lesion was found.
Code 4 – suspicious finding	An abnormal lump is present, but initial judgement did not indicate malignant morphological characteristics.
Code 5 – highly suspicion of malignant finding	An abnormal finding was found with a very high suspicion of malignancy. An immediate biopsy will be performed.
Code 6 – known cancer finding	A cancer is proven by biopsy. This category applies to women that has follow-up mammograms after proven cancer.

Clinical examination

In a clinical examination of the breast the breast is palpated to examine the solidity and size of the lump with a potential malignancy. In the screening setting, the clinical examination is performed after the mammogram is taken. For women who themselves detect a suspicious lump in the breast have a triple-diagnostic procedure referred to as a clinical detection outside the screening program [2].

Histopathology

Radiologists perform biopsies on suspicious cases and send the specimen to pathologists for microscopy analysis [77]. The specimen is examined for morphological characteristics and is categorized into tumor size, histological grade, oestrogene receptor positivity, progesterone receptor positivity, HER2 over amplification, a marker of cell proliferation Ki67, and lymph node status [79-81]. Approximately 85% of the specimens are found to be ductal carcinoma, 15% are lobular carcinoma [82]. Stage and grade are defined based on these characteristics [71].

Staging

Stage is the most important classification of breast cancer due to its importance in prognostication [83]. Stage is defined based on the TNM classification, where T refers to the tumor size of the primary tumor, N is number of affected lymph nodes and marks the regional spread, and M is distant metastasis [73]. T1 is defined as a tumor with a maximum diameter of 2 cm or less, T2 is a tumor larger than 2 cm but no more than 5 cm. More than 90% of the tumors have size T1 or higher, while only 30% of the women have affected lymph nodes. Few women have distant metastasis, M1.

Molecular subtypes

Molecular subtyping is a recent addition to tumor subtyping, where gene expression analysis [84] is used to categorize subtypes into the five intrinsic molecular subtypes Luminal A, Luminal B, HER2 enriched, basal-like, and normal-like tumors [85]. Molecular subtyping has improved decisions for assigning the appropriate oncological treatment to improve survival. Luminal A are ER and PR positive, but HER2 negative cancers. Luminal A benefit from hormone therapy and may also benefit from chemotherapy. Luminal B are ER positive, PR negative and HER2 positive tumors. The luminal B breast cancers benefit from chemotherapy and may benefit from hormone therapy and treatment targeted to HER2. HER2 tumors are negative for ER and PRS, but positive for HER2. HER2 breast cancers benefit from chemotherapy and treatment targeted to HER2. The triple-negative tumors are negative for ER, PR, HER2 negative. Basal-like breast cancers benefit from chemotherapy.

2.1.7 Prognosis

The five-year and ten-year survival from breast cancer is approximately 90% and 85%, respectively [86]. However, the breast cancer survival is differential dependent on tumor size, affected lymph nodes, and distant metastasis. Approximately 15% of the women have in-situ cancers and have a 5-year survival of more than 99%. Approximately 60% of the women have an

invasive cancer not spread to the lymph nodes and a 5-year survival of 99%. Approximately 10% of the women have regional spread of the cancer to lymph nodes resulting in 85% 5-year survival. Cancers with distant spread are found in approximately 1% of the women who in consequence have a 5-year survival of 30% [1, 87].

The female hormones oestrogen and progesterone (HR) and the human epidermal protein (HER2) are growth factors for tumors and are commonly used in addition to staging to characterize tumor subtypes. Cells with abundant receptors of these hormones could lead to increased tumor growth. Approximately 10% of the women have triple negative cancers with a 75% 5-year survival [87].

2.2 BREAST IMAGING

2.2.1 Mammograms

Ionizing radiation is used to x-ray the breast in digital mammography [88]. A radiographer positions the breast between two plates, one compression plate that is transparent to x-rays and a larger plate that contains the detector. An image sensor registers the x-ray that is transmitted through the breast. The fibro-glandular tissue attenuates the amount of radiation that reaches the sensor, while the radiation transmitted through the fatty tissue easily reaches the image sensor. Prior to presenting the image it is inverted so the radio-dense tissue appears white on the image and the fatty tissue appears dark. During mammography, images of the left and right breasts are taken from the craniocaudal (CC) view from above the breast, and in addition from the medio-lateral oblique view diagonal from the outer side of the breast. In the case of a suspicious finding, additional views could be taken such as magnification views or views from the side of the breast.

Prior to approximately year 2000 x-rays of breasts were developed on analogue films. The films were narrow in dynamic range and, after development of the film the image contrast was fixed. Nowadays, digital mammography uses a semiconductor detector that has a large dynamic range, which results in images with high contrast and, images can be further manipulated in post-processing.

2.2.2 Mammographic density and density change over time

Mammographic density is the x-ray attenuated image depicturing the fibro-glandular tissue from the breast. The bright part of the image represents the radio dense fibro-glandular tissue, while the dark part of the image depictures the fatty breast tissue. Mammographic density is largely composed of collagen (30%), but also by glandular structures [89], while less than 5% consists of epithelial cells [90]. Breast cancer is an epithelial based cancer but a cancer could develop in the near milieu of stromal and connective tissue [91].

Wolfe was the first to classify different levels of mammographic density into four categories and he also described an association with breast cancer risk [92]. Tabár later presented an alternative classification of mammographic density [93]. Boyd defined the concept of percent mammographic density in relation to breast size using a semi-automated method called Cumulus [94]. The American colleagues of radiology has also presented the BI-RADS breast composition

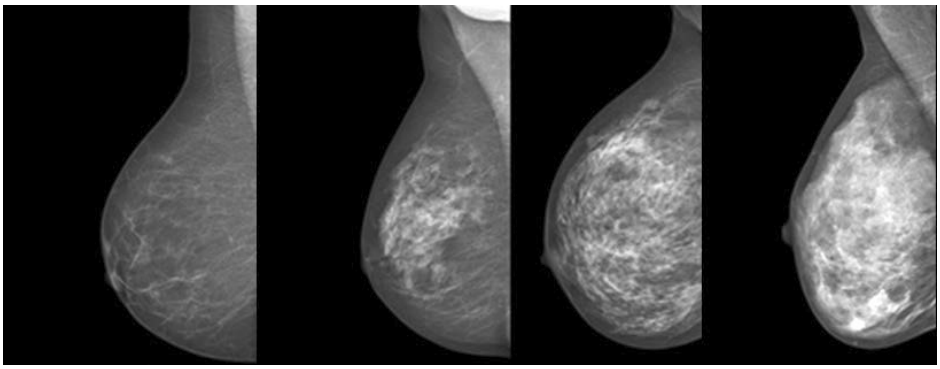
classification for assessing the probability of masking of a cancer by mammographic density [95]. The BI-RADS 5th edition is the most commonly uwed density categorization and is widely used by radiologists today [76].

Table 3. BI-RADS breast composition coding.

BI-RADS breast composition category	Description
A	Almost entirely fatty breasts
B	Scattered areas of fibro-glandular tissue present
C	Heterogeneously dense breasts that could obscure small masses
D	Extremely dense breasts that lowers screening sensitivity

In a screening population of age 50-70 approximately 10% of the women are found in category A, 40% in category B and C each, and 10% in category D.

Figure 3. Mammograms of four breasts with breast compositions BI-RADS A, B, C, D from left to right.



Fully automated software for percent mammographic density assessment have been developed and they measure mammographic density as either the area percent density of the total breast area [96] or as the volumetric percent density of the total breast volume [97]. Several software were then developed over the years for either area or volumetric assessment of mammographic

density [98-101]. Computerized scores which mimic the BI-RADS A, B, C, D categories have also been developed based on percent density cut-offs.

Percent density (PD) is mainly affected by age and BMI [102] and the PD decrease is largest during menopause [103]. Women in the highest density category have 4-6 times higher risk of breast cancer compared to women in the lowest category [104]. At the same time, mammographic density lowers the sensitivity of a mammogram, i.e. the probability for a radiologist to find a cancer. An on-going study in KARMA shows that the sensitivity varies from 88% in BI-RADS A women to 51% in BI-RADS D women.

Mammographic density change

The bulk of mammographic density research literature is based on mammograms from cross-sectional studies. A broad understanding has been reached on how mammographic density is associated with risk of breast cancer, associated with other risk factors, masking of breast cancer, cyclic menstrual changes, and natural involution [102, 103, 105, 106].

Women with a high mammographic density have 4-6 higher risk for developing breast cancer compared to women with low mammographic density [107]. Masking reduces the detection of breast cancer by up to fifty percent [76]. It is known that mammographic density is reduced 5 days prior to menses and is increased during the second half of the menstrual cycle [106]. Natural involution reduces mammographic density mainly during menopausal transition, and is on average 1% per year in premenopausal women and 0.5% in postmenopausal women [108]. In addition, studies have been performed on how density change over time is associated with breast cancer [109-111], and how mammographic density could be used to predict response to risk reducing therapies [112]. Differential breast involution over time has not been shown to be associated with breast cancer. However, mammographic density reduction has shown to be an early marker of women who respond to tamoxifen therapy and experience a reduction in breast cancer incidence [113].

A mammographic density change over time is a good proxy for women that respond to endocrine treatment and show a reduction of recurrence and initial development of breast cancer [112, 114]. By visually inspecting mammograms in a time series, it is obvious that different parts of the breast are captured by the radiographers in the images. This problem needs to be addressed. Imaging registration techniques are generally available [115], but they are not currently used for correcting the technical differences prior to measuring mammographic density. In this thesis we describe how an alignment protocol was developed to address this issue.

Radiographers are challenged everyday with requirement of consistent positioning and compression of the breast during mammography. Below image illustrates the problem (A) and shows how this could be handled (B) by aligning the images prior to measuring mammographic density. The global rigid registration technique was used to correct the images.

Figure 4. Two mammograms of the same breast were taken within minutes apart by the same radiographer. In panel A, the mammograms were superimposed to show the difference in breast placement in the mammography machine. In panel B, the two images were digitally aligned to the image showing the smallest breast size outlined with red in panel A prior to density measurement.

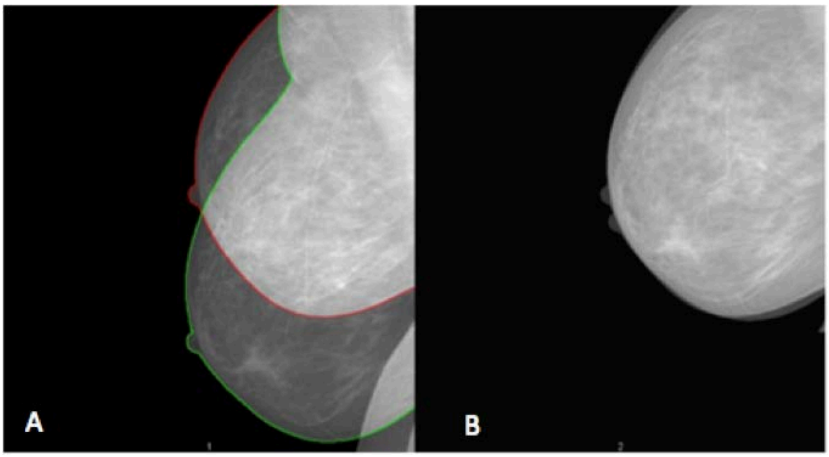
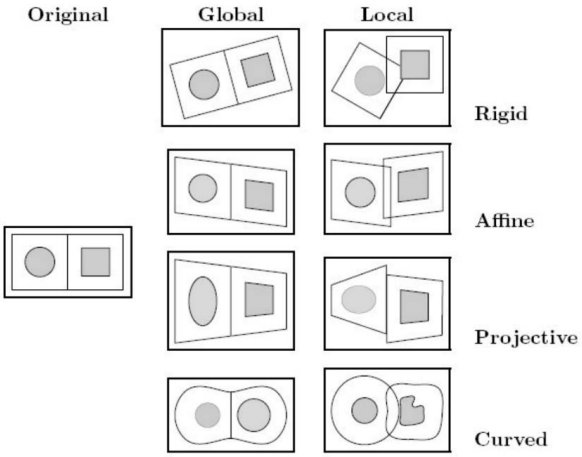


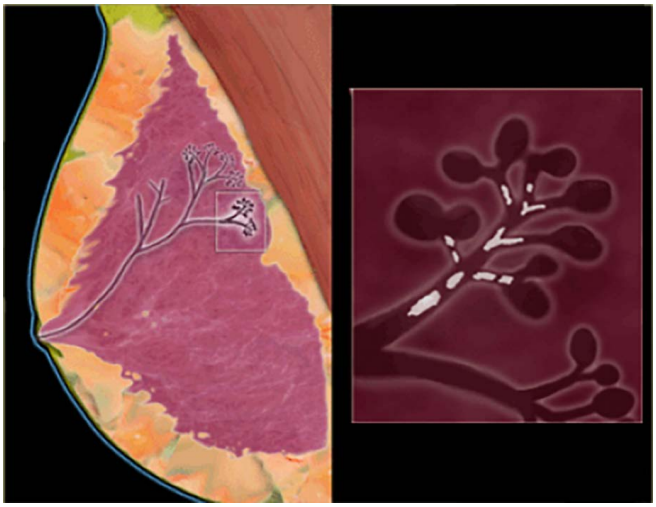
Figure 5. Image registration techniques. Maintz and Viergever 1998.



2.2.3 Microcalcifications

Microcalcifications are deposits of calcium smaller than 1 mm and are commonly located in the terminal duct lobular units and in the ducts. Microcalcifications appear as white dots on the mammogram [116].

Figure 6. Microcalcifications with a typical pattern inside the terminal ducts. American Cancer Society.



Microcalcifications are found in 90% of the ductal carcinoma in-situ tumors [117]. Ductal carcinoma in-situ is a common precursor for breast cancer [118] with a 40-100% increased risk for invasive cancer [119]. Microcalcifications are BI-RADS classified according to their mammographic morphology and distribution [120]. Type I are calcium oxolate microcalcifications that form pyramidial structures in a planar surface. Type II are calcium phosphate (hydroxyapatite) microcalcifications with diffuse shapes and irregular surfaces. The morphology of the microcalcifications determines whether the microcalcifications are potentially malignant or is a risk factor for breast cancer [121].

Table 4. BI-RADS malignancy classification of microcalcifications [76].

Code	Microcalcification description
Code 2 – benign finding	a) Round opacities or scattered macrocalcifications, typically calcified fibroadenoma or cyst b) Vascular calcifications
Code 3 – probably benign finding	Clusters of smaller calcifications of round or oval shape.
Code 4 – suspicious finding	a) Microcalcifications that appear amorphous or indistinct in a cluster b) Heterogeneous and pleomorphic microcalcifications
Code 5 – highly suspicion of malignant finding	a) Linear branching pattern of microcalcifications, segmental distribution b) Microcalcification cluster with segmental or galactophorous distribution c) Microcalcifications in architectural distortions

One of the Hanahan & Weinberg 10 Hallmarks of cancer is the activation of invasion and metastasis [19]. Epithelial-mesenchymal transition (EMT) is a phenomenon where epithelial cells lose their characteristic traits and gain mobile mesenchymal traits. This phenomenon is part of intravasation when malignant cells start to gain mobile mesenchymal characteristics to migrate from the extracellular matrix toward the blood vessels to metastasize [122, 123]. It has been hypothesized that microcalcifications could result from a mineralization process that is sustained by EMT [124] similar to bone osterogenesis. Other explanations for the development of microcalcifications has also been studied, including cell necrosis [125]. Microcalcifications have been shown to predict breast cancer lymph node status [126]. A study also showed that microcalcifications predicts HER2 and Luminal A molecular subtypes in the pre-operative setting [126]. It is not known at what earliest point in time microcalcifications are predictive for a breast cancer.

2.2.4 Masses

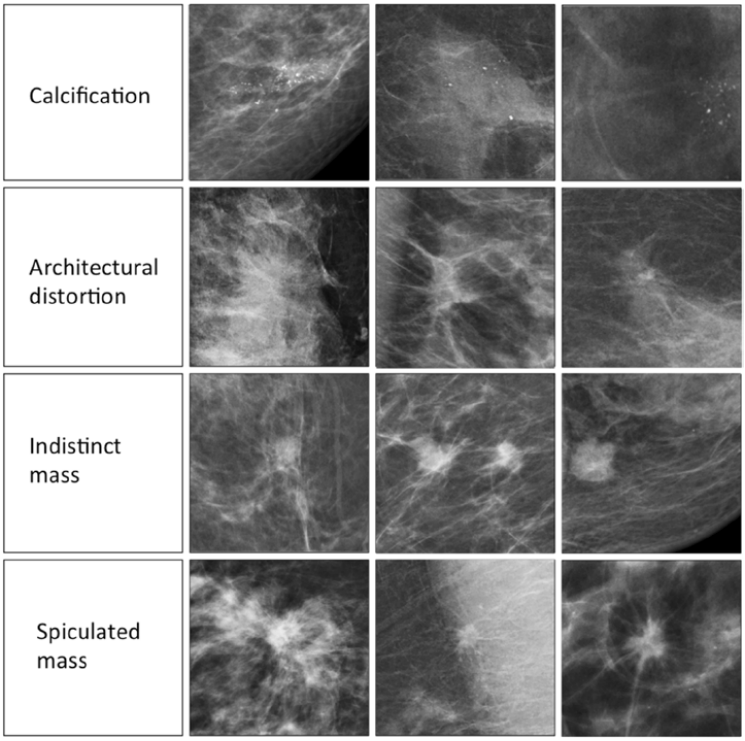
A mass in the breast is a benign or a malignant breast lesion. A benign lesion could lead to a proliferative lesion, hyperplasia or atypical hyperplasia with an increased risk for developing into a malignant tumor [127, 128]. Fibroadenoma is a common benign disease that through epithelial elements in their nodules of fibrous tissue could develop into a breast cancer [129]. A study showed that fibroadenomas share the same genetic, reproductive, and lifestyle factor risks as

malignant tumors [29]. In the diagnostic setting, a Computer Aided Detection (CAD) software is used to indicate lesions that have a high probability for malignancy. In this thesis we study the risk of breast cancer based on a software that uses a lower probability for malignancy to identify women that are likely to be diagnosed with breast cancer.

2.2.5 Bilateral breast asymmetry of mammographic features

The asymmetric distribution of density in a single breast is routinely examined by radiologists [76]. However, bilateral differences of mammographic features (mammographic density, microcalcifications, masses, distortions) between left and right breast is not regulated in the radiologists' examination procedures. In this thesis we describe the first effort to use bilateral breast asymmetry of mammographic features for breast cancer risk assessment. A recent study paid interest to this and further studied bilateral breast asymmetry of mammographic features [130]. The potential value to study bilateral asymmetry of mammographic features is based on the fact that the vast majority of breast cancers are developed in one breast only. For this reason, the breast tissue could be investigated for risk factors of breast cancer, where pre-diagnostic images are examined for differences in mammographic features. One breast is considered diseased and the other breast is a paired control. The paired comparison is by design adjusted for the woman's germline, personal disease history, and lifestyle factors.

Figure 7. X-ray image of microcalcifications, masses, and architectural distortions. American Cancer Society.



2.2.6 Mediation of risk factors through mammographic features

Breast cancer is a genetic disease, but several factors contribute to the development of a tumor [19, 131]. Mammographic features are measures from the imaged breast tissue. The most studied mammographic features are mammographic density, microcalcifications, masses, and tissue distortions. Several hormonal risk factors (age at menarche, parity, age at first childbirth, prior breast biopsy, HRT use) are influencing a change in the breast tissue and mediates their risk association with breast cancer through mammographic density [132, 133]. Studies also suggest that familial history of breast cancer is partly mediated through mammographic density.

An overview of how risk factors for breast cancer incidence are mediated through mammographic density and microcalcifications is seen in below table. The mediation analyses were based on the KARMA cohort using a Cox regression method developed by Nevo et al. [134]. The models were adjusted for potential confounders of the associations between a) risk factors and breast cancer, b) mammographic features and breast cancer, and c) risk factors and mammographic features. In addition, the models were adjusted for d) potential confounders of the association between risk factors and confounders for the association between mammographic features and breast cancer (i.e. mediator-outcome confounders). The potential confounders were age, BMI, parity, hormone replacement therapy, prior biopsy, and family history of breast cancer. The mediation property of mammographic density and microcalcifications are of special interest for the prediction model that is developed in this thesis, because the model uses mammographic features as the main component.

Table 5. Mediation of breast cancer risk factors through mammographic features.

Risk factor	Mediation through mammographic density (%)	Mediation through microcalcifications (%)
Parity	40	Not significant
Age at first child	17	Not significant
Current HRT use	25	52
Current alcohol use	25	Not significant
Family history of BC	6	7
Benign breast disease	20	23
Prior biopsy	24	41
PRS score	6	14

HRT – hormone replacement therapy
PRS – polygenic risk score including 313 SNPs

Body mass index, age at menarche, and current smoking were not significantly mediated through mammographic density nor through microcalcifications.

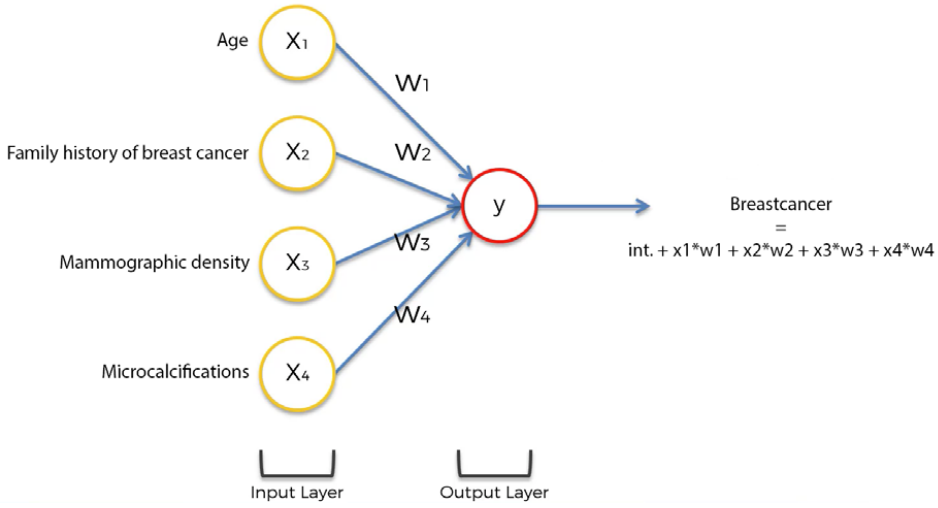
2.3 ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is today in extensive use in many areas, and especially in the area of classification or prediction based on image information. Today, mammography screening units make use of AI based detection tools to improve their ability to identify cancers. In this thesis, we use AI for assessing breast cancer risk based on mammograms to improve the accuracy of risk assessment.

2.3.1 General principle

Neutral networks date back to the 40's and was initially constructed as a threshold logic method to mimic human brain intelligence [135]. A second milestone in the 60's was the development of a method called backpropagation that is a method to fit network models to input data [136]. The simplest form of a neural network can be illustrated using logistic regression.

Figure 8. Neural network of logistic regression. (w=beta, int.=intercept).

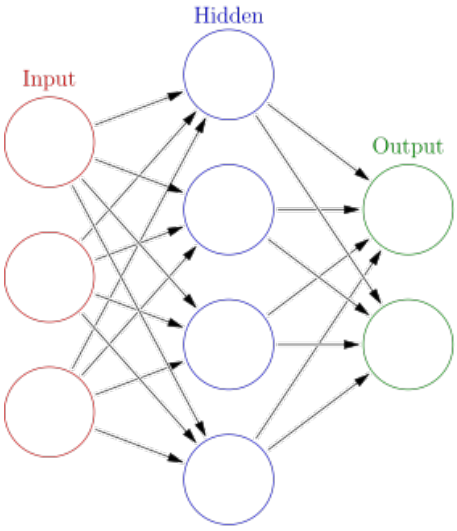


Risk factors are used in the input layer. Each node has data values corresponding to each risk factor. Each risk factor data value is multiplied with a unique weight. An activation function (here a sigmoid logistic regression function) calculates the probability that each woman is positive or negative for breast cancer.

In general, a neural network is constructed by neurons that are structured in several layers starting with the input data layer, then hidden layers, and ends with the output layer. Each node

belongs to one layer and is connected (weighted) to the other notes in the adjacent layer. The input nodes are bits of raw data that is used by the network, e.g. millions of pixel data points from one mammogram. Each node in the first hidden layer receives data from each input node. The value for each node is multiplied with a unique weight between 0 and 1. The first hidden layer sums up the weighted values from all input nodes together with a bias and calculates the output value using an activation function. The output value is sent to the nodes in the next layer. The last output layer calculates the probability of each output value, e.g. positive and negative breast cancer status.

Figure 9. Three-layer neural network.



Supervised neural networks are trained by knowing the output data values (e.g. breast cancer status). After the raw data has been input from several individuals (e.g. mammograms of women), initial weights are applied to each node. The backpropagation procedure includes a gradient descent algorithm that finds the best weights for the hidden layer(s) to minimize a loss-function. The loss-function minimizes the probability of making an error when classifying the breast cancer case status from each of the input mammograms [137]. Neural networks that analyze images commonly interprets the data as two-dimensional objects and is referred to as convolutional neural networks.

2.3.2 Computer aided detection

Computer Aided Detection (CAD) is a complementary device for helping radiologists to identify a cancer on a mammogram [138]. Artificial intelligence (AI) is used in recent developments as a decision support tools. The performance of an AI based tool for detection of cancer in a digital mammogram is now on par with a radiologist performance with a sensitivity above 70% and a

specificity above 95% in the screening setting. CAD systems are classified as medical devices and are regulated in US by the Food and Drug Administration (FDA).

2.3.3 Detection vs short-term risk

By the definition from FDA, detection is the identification of a malignant lesion in the breast. Based on that, I defined short-term risk based on mammographic features in a distinct different manner as the identification of a breast with a malignant predisposition, but without identifying a specific lesion or region in the breast. I also set a time constraint of up to five years risk projection to be considered a short-term risk.

2.4 RISK ASSESSMENT

2.4.1 General concepts

Prediction versus explanation

Epidemiology showed great success using explanatory statistics in areas such as lung cancer to explain lung cancer outcome from smoking [139]. An ideal epidemiological scenario is to estimate known necessary and sufficient causal factors to explain the outcome of interest. The causal relationship could in addition be supported by a theory describing an underlying biological mechanism. However, in many health quests a complete explanation cannot be reached. Familial risk factors and germline genetic abnormalities explain approximately 25% of the breast cancers [25]. Most breast cancers occur in women without a family history of breast cancer and are caused by somatic mutations in the genome [68]. In contrast to explanatory modelling, predictive modelling could be defined as the development of models that estimates outcomes in new data based on factors in that data [140]. The aim is to optimize the accuracy of estimating the outcome in the new data by reducing the prediction error. The prediction error is measured by a loss-function. The statistical approach for prediction is fundamentally different from explanatory statistics. Predictive modelling predicts the outcome based on predictive factors, using statistics to minimize a loss-function; while explanatory modelling estimates causal associations between exposures and outcome. However, both statistical approaches make use of the same basic scientific principle of replication to warrant the accuracy of the models. In this respect, the two approaches could be compared through their abilities to replicate results in new data.

Sensitivity and specificity

A group of women with breast cancers is referred to as true positives. In mammography screening radiologists will identify a proportion of the true positives, referred to as the radiologists' sensitivity. In general terms, the sensitivity is the proportion of individuals who tested positive among all true positives, that is the probability of testing positive using a medical test in the group where all are diseased individuals [141]. Specificity is the probability of testing negative using a medical test in the group where all individuals are healthy.

Confusion matrix

A risk model predicts the probability for an individual to be a breast cancer case. For any practical use of the risk model a cut-off is needed to classify at what probability level an individual is considered to be a breast cancer case. If the cut-off for being considered a breast cancer case is set at zero percent, then all individuals will be considered by the model to be breast cancer cases. This means that the sensitivity of the model will be 100%, but the specificity of the model will be 0%. If on the other hand the cut-off for being considered a breast cancer case is set to hundred percent, then no individual will be considered by the model to be a breast cancer case; the sensitivity of the model will be 0% and the specificity will be 100%. A two-by-two table can be used to present how the medical test predicts disease status in relation to the true disease status. A confusion matrix is created by counting the number of individuals in each cell.

Table 6. Confusion matrix with 0% cut-off probability for classifying a case as positive. Sensitivity 100%, specificity 0%.

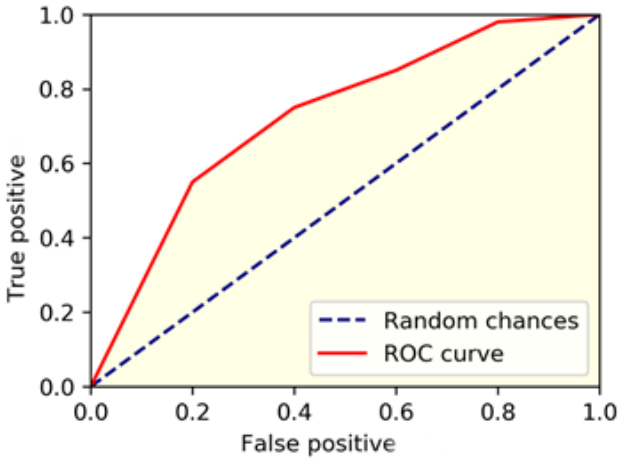
	True disease status	
Test result	**Breast cancer case**	**Breast cancer free**
Positive	100 true positive cases	0 false positive cases
Negative	0 false negative cases	0 true negative cases

Multiple tables are calculated for different probability cut-offs. Then a receiver operating characteristics curve (ROC) is created by plotting the sensitivity and specificity, for each of the probability cut-offs, on a two-dimensional plot where sensitivity (true positives) is on the Y-axis and 1-specificity (false positives) is on the X-axis.

Discrimination performance

The discrimination performance of a model is calculated as the area under the ROC curve (AUC) as is illustrated in the below figure [142]. An AUC of 0.5 corresponds to the diagonal dotted line and means that regardless of which probability cut-off is used to classify a woman as a positive case there will not be a greater chance than 50% that the positive case is truly positive.

Figure 10. ROC curve and random chances diagonal.



AUC can be calculated based on the c-statistic (concordance statistic) using logistic regression. The c-statistic is the probability that the individual who truly has the outcome have a higher predicted probability by the test than the individual who truly does not have the outcome. AUC is a theoretical concept that not necessarily give a practical understanding of how well the risk model can distinguish true cases from true healthy individuals in a clinical setting. In a clinical setting it will be required that a risk model shall operate at a certain sensitivity or specificity. The ROC can tell which specificity will be reached given a certain sensitivity or vice-verse.

Calibration

A risk model predicts the probabilities for individuals to have the disease. This results in a distribution of risk probabilities that commonly is stratified into deciles for an estimation of calibration [143]. Calibration compares the observed probabilities for having the disease with the expected probabilities, as predicted by the model, for having the disease in each of the deciles. A statistic called the Hosmer-Lemeshow test estimates how well the observed risks compares with the expected risks.

Risk stratification

The clinical use of a risk model is the model's ability to distinguish individuals with a high and a low probability for developing the disease, respectively. The risk classification in breast cancer is defined by clinical guidelines [144, 145]. The most common guideline in Europe is the National Institute for Health and Care Excellence (NICE) guidelines [144]. NICE recommends different types of clinical follow-up of women dependent of their levels of risk. Women in the high-risk category are recommended more frequent screening or a more sensitive screening modality from age 30 and above. The guideline is described in more detailed under Prevention.

Validation

Validation is a technique that critically tests a risk model using new data that was not used during the training of the model [146]. The preferred form of validation is external validation, where the new data origin from another population than was used in the training. The external population can either be women that attend screening under similar circumstances, e.g. at another hospital in the same country. The external population can also be women from another screening setting. Examples of screening settings are that different screening modalities, screening intervals, personal screening history, and ethnicities are included. The generalizability of a risk model is less challenged by predicting new data in a screening setting similar to the training setting and is challenged more by predicting new data in new screening settings.

Common validation outcome measures are sensitivity, specificity, AUC, risk stratification, and clinical usability.

2.4.2 Risk assessment (long term)

Over the last 40 years, attempts have been made to identify women that will develop breast cancer. The Gail risk model was introduced in 1989 and was based on approximately 2,852 cases and 3,142 controls retrieved from a large screening cohort [147]. The model identified age, age at menarche, number of previous taken biopsies, age at first childbirth, and number of relative with breast cancer as risk factors. Gail constructed the model to estimate 5-year absolute risk of breast cancer, calibrated to the general female population, based on i) estimating the relative risks for each risk factor adjusted for the others, ii) estimate the absolute risks of the women based on their profile of risk factor exposures, while accounting for competing mortality due to other causes. A logistic regression model was used to estimate the relative risks and a Fine and Gray regression model was used to estimate the absolute risks accounting for the competing risks [148, 149]. The discrimination performance has been reported in ranges from AUC 0.52 to 0.7 in cohorts with different criteria for selecting cases and controls [150]. The model was validated in several populations.

A second landmark in the risk model development was seen with the Tyrer-Cuzick risk model that estimates 10-year and lifetime risks [151]. By this time, more risk factors had been identified. The Tyrer-Cuzick model include age, BMI, age at menarche, age at first childbirth, use of HRT, menopausal status, benign breast disorders (atypical hyperplasia, lobular cancer in situ), first and second order family history of breast and ovarian cancer, Ashkenazi origin, and BRCA-gene mutation. Cuzick also introduced the "low susceptible" gene which he meant should be prevalent in the population but have a lower risk association with breast cancer. A later update to the Tyrer-Cuzick risk model also includes an 18 PRS score and mammographic density [152].

A third landmark in the risk model development was done with the BOADICEA model which estimates lifetime risk for developing breast cancer based on the genetic risk [67]. BOADICEA was developed to assess the probability for a woman to carry a BRCA1/2 mutation given her family history of breast cancer. The family history covers up to 3rd degree relatives, known BRCA

mutations in the family, Ashkenazi origin, bilateral cancer status, and ovarian cancer. The model was further developed to include a PRS score. The model has been validated in 22 populations. An on-going development will also include classical lifestyle risk factors and mammographic density.

Many models have been developed over the decades that have similar setups of risk factors as Gail, Tyrer-Cuzick, and BOADICEA [150]. For instance, the BCSC model developed as an extension to the Gail model. The prediction accuracies are low to moderate and the models may not be cost-effective for the use in risk screening of the general female population.

Today, a breast cancer risk model is more or less synonymous with the concept of predicting lifetime risk or at least ten-year risk [150]. The aim is to identify women that could be prevented from breast cancer. This concept has great value for women with an extensive familial risk of breast cancer [67]. However, most cancers occur in women without a family history of breast cancer. A recent study questioned the use of assessing lifetime risk as is commonly requested by clinical guidelines [153]. Risk models may show lower accuracy in long-term risk assessment compared with shorter term risk assessment.

2.4.3 Short-term risk assessment

A challenge with traditional risk models is that the predictive accuracy is low to moderate and that they are not designed to improve mammography screening. In paper II I constructed a prediction model that is designed to circumvent these problems. The model uses mammograms as the main component and could add lifestyle factors and a polygenic risk score to further increase the accuracy. The model is a two-year risk for the purpose to be useful in screening programs with biennial screening. The model's ability to stratify women into high-to-low risk is essential for clinical use. The risk model fits with clinical guidelines that have been developed for the general population, where more intense screening is recommended for women at high risk of breast cancer [144, 145]. More intense screening will lead to more detected cancers. This means that the intervention will lead to earlier detection of breast cancer, rather than primary prevention of breast cancer. This means that the clinical aim for using the risk model in this setting is to improve the screening efficiency for these women. The Envision consortium recently recognized this as the second aim for using a risk model [154]. A recent systematic review observed that a risk model could benefit from a short-term prediction to increase the accuracy of identifying women that are at high risk of breast cancer [155].

2.5 MAMMOGRAPHY SCREENING

2.5.1 Age based screening

Breast cancer screening was designed to detect cancer early and to reduce breast cancer mortality. In Sweden mammography screening was implemented between 1976 and 1997 in different counties [3]. Landmark papers have been shown that tumors nowadays are found at earlier stages [156, 157] and that screening reduces mortality from breast cancer by approximately 20% compared to women not attending a screening program [158, 159]. The screening age varies between countries. In Sweden, the screening age range is 40 to 74. The current screening

intervals are 18 and 24 months. Some Swedish counties screen women in the ages 40 to 54 in 18 months intervals and older women in 24 months intervals. In European screening programs approximately 3% of the women are recalled after each mammography screen. After further work up approximately 0.5% of the women are diagnosed with a cancer. Approximately 75% of eligible women attend the screening [160] and among these women approximately 75% of the cancers are detected by the screening program. Women who develop interval cancer, approximately 25% of the women in Sweden, do not benefit from the mammography screening program. In this thesis we develop a prediction model that could lead to detecting the interval cancers earlier at their prior regular mammography screening on average 1 year before the current interval diagnosis date.

2.5.2 Supplemental imaging

In addition to digital mammography, ultrasound or magnetic resonance imaging could be used for women with x-ray radio dense breasts [161]. Magnetic resonance imaging is more commonly used to monitor women with a familial or genetic high risk for developing cancer over time [106]. The advantage with these modalities is that they have a higher sensitivity, but on the other hand have a lower specificity than a mammogram. Contrast-enhanced mammography is a late development that has a potential to be a large-scale screening modality for women in need of supplemental screening with a similar performance as MRI but with a shorter protocol [162].

In this thesis, we identify high-risk women who also have a high density and therefore potentially could have a value from supplemental imaging.

2.5.3 Screening frequency

Mammography screening intervals varies between one and three years depending on age-group and screening program [163]. The proportion of cancers that will be detected in screening or as interval cancers is related to screening frequency and the age of the woman [164]. A study found that in-situ cancer more likely develop into invasive cancers in screening programs with three-year intervals compared to screening programs with shorter screening intervals [165].

In this thesis, we investigate the potential value for recommending more frequent screening intervals for women at high risk of breast cancer who in addition have low mammographic density, that is with an increased probability to have a fast-growing cancer that develop between mammography screening intervals. Women with low risk of breast cancer could potentially be recommended less frequent screening.

2.6 BREAST CANCER PREVENTION

Risk reducing strategies has been thoroughly investigated and, risk factors and preventive measures have been identified [166]. Risk reducing mastectomy is offered to women with a high genetic risk for breast cancer [167]. A less drastic preventive therapy is tamoxifen that reduce recurrence of oestrogen-positive breast cancer but can also be used to reduce the risk in healthy women from developing breast cancer [113]. Studies have shown that approximately 30% of the oestrogen-positive cancer can be primary prevented using tamoxifen, raloxifene or aromatase

inhibitors [168, 169]. Side-effects are venous thromboembolic events and endometrial cancers, which increase by a two-fold, with an increased 5-year probability of 0.5% and 0.25%, respectively [170]. Endometrial cancers and endometrial thickness could develop in consequence of the pro-oestrogen agonist effect of tamoxifen in a low oestradiol environment. Menopausal symptoms are prevalent side-effects which are the main reason for not continuing taking the medication [171]. Vasomotor symptoms are most common with a prevalence of approximately 35%. Vasomotor symptoms are also associated with increased cardiovascular morbidity and mortality [172]. A study of 5 mg low-dose tamoxifen indicated that breast cancer events reduced by half and, rare serious adverse events were not significantly higher in the treatment arm compared with the placebo group; but menopausal similar symptoms were slightly increased in the treatment arm (2.1 vs. 1.5 hot flashes per day) [173]. The number needed to be helpful to reduce breast cancer events was ten times better than the number needed to be harmful of serious adverse events, including vein thrombosis and endometrial cancer events.

Tamoxifen is an anti-oestrogen that inhibits oestrogen to bind to oestrogen receptors in the cell nucleus [174]. It is not well-known how the selective oestrogen receptor modulator reduces the risk for developing breast cancer, but the modulator is considered to reduce the rate of cell division and proliferation and to induce apoptosis. Tamoxifen is a pro-drug that is metabolized into metabolites, where endoxifen is considered the main metabolite that affects breast cancer risk. The ability to metabolize tamoxifen into endoxifen is inherited and is affected by polymorphisms in the CYP2D6 gene [175]. Up to 10% of the women have poor or ultrarapid CYP2D6 gene activity, which makes tamoxifen less useful for these women to reduce their breast cancer risk [176]. Poor metabolizers experience too low therapeutic levels, and ultrarapid metabolizers experience too high therapeutic levels which leads to severe side-effects and discontinuation of the therapy.

Tamoxifen also reduces mammographic density and has shown to be an effective early marker for which women respond to tamoxifen [112, 114]. One hypothesis is that cells with reduced proliferative activity are less radiolucent on a mammogram which leads to lower mammographic density. A study on premenopausal breast cancer patients showed that mammographic density was more likely to remain low two years after discontinuing medication, if the woman initially showed a density response to the therapy [177]. The sustained effect was seen in sixty percent of the women who had an initial density response to tamoxifen.

The focus of paper III was to address the question whether the tamoxifen dose could be lowered with non-inferior reduction of mammographic density as a proxy for reduced breast cancer risk, while lowering the side-effects for the woman.

Guidelines have been developed to assign clinical preventive action points to women at different levels of breast cancer risk. The NICE guideline (UK National Institute of Health and Care Excellence) has guidelines for women with familial history of breast cancer [144]. The guideline recommends increased surveillance, use of tamoxifen, and possibly prophylactic mastectomy in the high-risk group dependent on the genetic carrier ship. Women with moderate risk are recommended increased surveillance and lifestyle changes.

Lifestyle changes is another potential means for primary prevention of breast cancer [178]. BMI, alcohol, physical activity, smoking, and hormone replacement therapy use are factors that could change the breast cancer risk.

3 AIMS AND HYPOTHESES

The overall aim of the thesis is two-fold - to improve mammography screening and to reduce breast cancer incidence. My work focused on developing and evaluating tools for measurement of mammographic density and risk assessment of breast cancer, and in addition to identify the lowest dose of tamoxifen that could be used to reduce breast cancer incidence and to improve the sensitivity of mammograms. The work was laid out in the following four studies.

I. To develop and evaluate a tool for measuring mammographic density in large-scale mammography screening cohorts and in clinical trials.

   The hypothesis was that processed mammograms, that are generated as part of the mamography screening workflow and are stored at the hospitals, could be used to measure mammographic density regardless of the type of mammogram and vendor of the mammography machine.

II. To develop and evaluate a short-term risk assessment tool for the general female population to identify women who are sent home with a negative screen, but are diagnosed with breast cancer before or at next mammography screen.

   The hypotehsis was that a short-term risk model based on mammograms could be developed that identifies women who will be diagnosed with breast cancer before or at next mammography visit and is more accurate than established risk models.

III. To test if lower doses of tamoxifen could be as efficient in reducing mammographic density compared to the standard dose of tamoxiofen but cause less side-effects.

   The hypothesis was that a lower dose of tamoxifen is non-inferior to reduce mammographic density, a proxy for therapy respons to tamoxifen, but still cause fewer side-effects.

IV. To model the feasibility of using low-dose tamoxifen to improve the sensitivity of mammograms due to the low-dose tamoxifen effect on reducing mammographic density.

   The hypoteshis was that screening sensitivity could be improved due to the effect that low-dose tamoxifen has on reducing mammographic density.

4 PATIENTS AND METHODS

4.1 STUDY POPULATIONS

The description below refers to the original populations which formed the basis of my study populations in the thesis. Selection criteria are stated in the 'Results' section for each of the included studies.

4.1.1 KARMA

In study I, II, and IV, the main study population was based on the KARMA (Karolinska mammography project for risk prediction of breast cancer) cohort [179]. KARMA is a prospective mammography screening cohort consisting of 70,877 women recruited between 2011 and 2013. KARMA includes approximately 35% of the women in the south Stockholm area and southern of Sweden that were invited to mammography screening under the recruitment period. KARMA is a prospective cohort, which means that women were followed from baseline and onwards. At baseline, information on exposures that possibly are related to breast cancer were collected. Exposures were collected using surveys, mammograms from the screening units, register data, and medical records. Women also donated blood. Whole blood and plasma were stored, and DNA was extracted. The information is continuously updated through collection of mammograms, register data, surveys, and additional blood drawing.

Mammograms were collected from 2011 and onwards for all women. In addition, a retrospective collection of available mammograms was performed for all breast cancer cases in the cohort. The collected mammograms were raw and processed full-field digital mammograms and analogue images. The analogue mammograms were digitized at the department with 50-micron spatial resolution and 12-bit greyscale dynamic range. Till current date almost three million mammograms have been collected.

4.1.2 LIBRO1

In study I, the LIBRO1 (Linné-Bröst1) population was included. The LIBRO1 population consists of 9,348 incident breast cancer cases, registered in the Swedish National Breast Cancer register, in the Stockholm-Gotland area between 2001 and 2008 in the age range 40 to 74. The study population was invited to donate blood and answer a survey in 2009. Sixty-one percent of the women (N=5,715) gave informed consent to participate in the study. Women responded to a survey and donated blood. The survey collected information on exposures that possibly could be associated with breast cancer. Whole blood was stored, and DNA was extracted. The median between cancer diagnosis in the full LIBRO1 population and the sub-population was 4.8 years.

Mammograms were collected from 2001 and onwards for all women, and a retrospective collection was performed for all available mammograms in digital archives and analogue films. The collected mammograms were mainly processed full-field digital mammograms and analogue images. The analogue mammograms were digitized at the department with 50-micron spatial resolution and 12-bit greyscale dynamic range.

4.1.3 CAHRES

In study I, the CAHRES (Cancer And Hormone Replacement Therapy Study), also referred to as SASBAC (Singaporean And Swedish Breast Cancer Study), population was included. The CAHRES study identified all breast cancer cases, registered in the Swedish National Breast Cancer Register, between 1993 and 1995 in ages 50 to 74. Women were invited to participate in the study and 84% signed an informed consent and responded to a survey. Healthy controls were selected in an incidence density sampling scheme, based on information from the Swedish Cancer Register, and were selected from the same time-period and were matched on age. In all, 2,818 breast cancer cases and 3,111 controls were included in the study. Exposures were collected based on survey data, mammograms from screening units, register data, and medical records. A sub-group of women later donated blood which formed the SASBAC sub-cohort. DNA was extracted for 1,534 breast cancer cases and 1,504 controls.

Analogue mammograms were collected from hospitals before and at the time of diagnosis. Mammograms were available for approximately 75% of the study population. Mammograms with poor quality were excluded. The analogue mammograms were digitized at the department with 50-micron spatial resolution 12-bit greyscale dynamic range.

4.1.4 KARISMA

In study III, the study population was KARISMA (KARMA Intervention Study). KARISMA is a six-month double-blind randomized six-armed and placebo-controlled non-inferiority dose-determination phase II trial. Women participating in the Swedish national mammography screening program at Södersjukhuset in Stockholm and at Unilabs mammography screening unit in Lund were invited to the study. In all, 159,207 women were invited, and 2,314 women volunteered and were investigated for eligibility to participate. Main exclusions were women with almost entirely fatty breasts and women with a history of cardiovascular disorder. The predefined number of women (N=1,440) were included. Women signed informed consent and information was collected based on survey data, mammograms, and in addition women donated blood. Plasma endoxifen levels and CYP2D6 gene activity were analysed. The women were allocated to placebo, 1, 2.5, 5, 10, or 20 mg of tamoxifen at study start for a six month administration of the medication.

Digital full-field raw and processed mammograms were collected at baseline and at study exit, that is at the scheduled six months visit or at time of discontinuation.

4.1.5 CSAW

In study II, the CSAW (Cohort of Screen-Age Women) was included as an external validation cohort. The women were recruited from the Karolinska Hospital mammography screening program between 2008 and 2015. The available 613 breast cancer cases with mammogram data were included and a random sample of 10,000 women with a negative mammogram were selected. After excluding controls outside the range of mammography years of the breast cancer cases, 8,489 controls remained. The mean age was 53 and the average follow-up time 5.2 years.

4.1.6 MBTST

In study II, the MBTST (Malmö Breast Tomosynthesis Trial) population was included as an external validation cohort. The trial included a subgroup of women, who volunteered to have tomosynthesis examinations in addition to digital mammography, among all women participating in the national screening program in Malmö, Sweden. The women were recruited to the trial between 2010 and 2015 and included 104 incidence breast cancer cases and 9,745 women with a negative mammogram [180]. The mean age in the cohort was 57 and the follow-up time was on average 3.7 years.

4.2 DATA

4.2.1 Research platform

Prior to my thesis work, I developed the KARMA research platform which was the first epidemiological platform to hold an extensive research dataset in one single database that was readily available through the web. At that time, skilled principal investigators not often knew in detail what data they were possessing, which caused a slow process in designing new studies. In the new system, principal investigators got information at their fingertips to perform typical tasks such as investigating inclusion and exclusion criteria for what studies were possible to perform. More extensive on-line analyses were also provided through the system [181]. The platform delivered the individual research data to the researcher after ethical approval by the research project principal investigator.

Figure 11. Schematics over the KARMA research platform data sources.

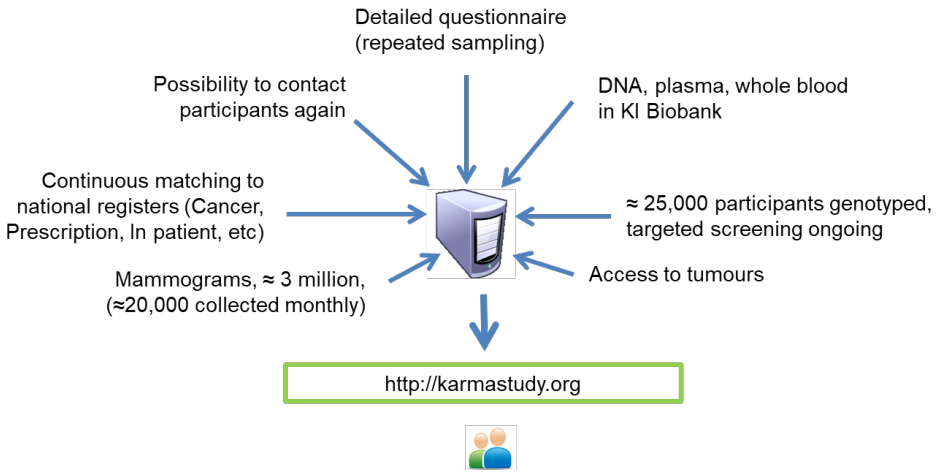
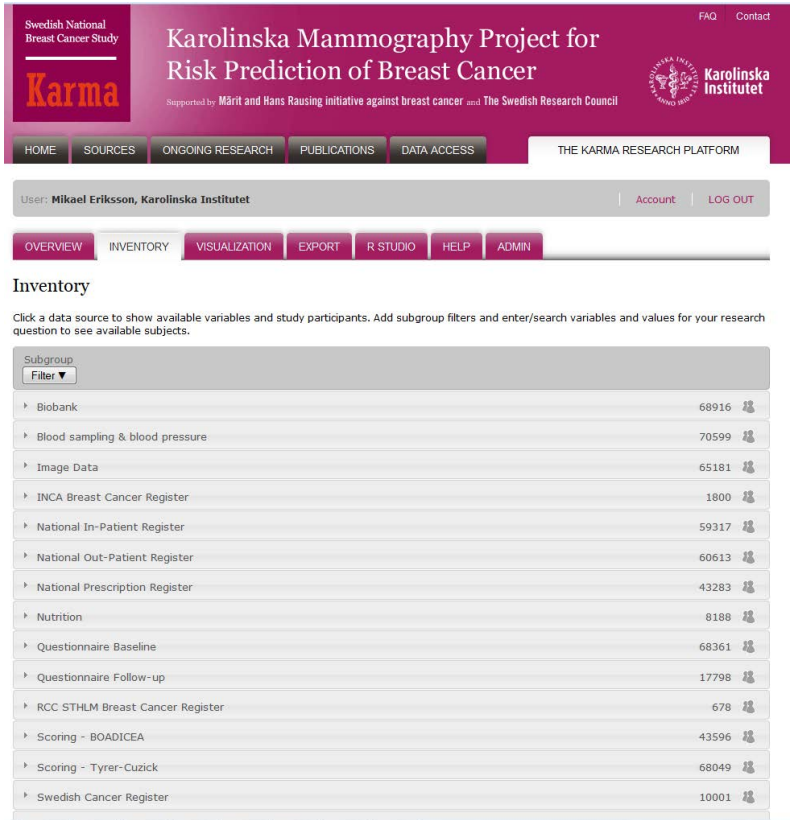


Figure 12. Web-view of the KARMA research platform.



In addition, I created the extensive KARMA web questionnaire and the vast majority of the finalized research datasets by quality checking, recoding, and derivations of variables based on the collected data [182]. The KARMA research platform is since its creation the basis for research in the breast cancer research group at the department of Medical Epidemiology and Biostatistics, Karolinska Institutet. The research platform was also promoted in 2015 as the raw model for the National Cancer Institute (NIH) epidemiological future projects [183].

Figure 13. The KARMA web questionnaire. Icons represents themes of questions.

4.2.2 Register data

In Sweden, population-based registers have a centuries-old tradition. The personal identifier PNR has been used since 1947. The personal number is given to each Swedish citizen at the time of birth. PNR makes it possible to link the register data to the women individually and in addition to the other individual information that the women contributed to the study. The following registers were used in this thesis:

- The Swedish Cancer Register containing information on type of cancer, date of diagnosis, invasiveness, TNM stage, and histological type. The register has a high coverage (98%) of all breast cancer diagnoses that were reported [184].

- The Breast Cancer Quality Register containing additional data on tumor size, stage, tumor receptor status, histological grade, and more [185].

- The Cause of Death Register started in 1952 containing data on the cause of death for each individual [186].

- The Screening Register at Regional Cancer Centre Stockholm-Gotland containing data on mammography screening status and recall status of the individuals in the Stockholm-Gotland area [187].

Register data were used in all studies.

4.2.3 Survey based data

Survey based data was used in studies I, II, and III. The questionnaires in KARMA and LIBRO1 were web based and the questionnaire in CAHRES was paper based. The women in LIBRO1 could request a paper-based questionnaire to replace the web-based questionnaire. The baseline questionnaires were filled in at time of enrolment. KARMA also included follow-up questionnaires. The KARMA questionnaire was the most extensive and included questions on background, reproductive health, use of medication, use of alcohol and tobacco, previous and current diseases and treatments, family history of breast and ovarian breast cancer, quality of life, physical activity, and diet. All cohorts used questionnaires that included questions on the essential breast cancer risk factors age, BMI, family history of breast cancer, age at menarche, parity, age at first child, contraceptives, menopausal status, benign breast disease, and use of hormone replacement therapy.

4.2.4 Mammograms

Mammograms were used in all studies. Mammograms from left and right breasts from medio-lateral oblique and craniocaudal views were collected. Mammograms from the KARMA cohort were used in all studies. Mammograms from KARISMA were used in study III and IV. Mammograms in the KARMA and KARISMA studies were collected prospectively from hospitals in the middle part and the southern part of Sweden. Digital full field processed and raw mammograms were collected integrated with the screening workflow. This made it possible to also include the raw images which otherwise are deleted automatically within a short timeframe in the screening workflow. The images were regularly transferred to the Department of Medical Epidemiology and Biostatistics (MEB), Karolinska Institutet, from the hospitals. Mammograms from the LIBRO1 study were retrospectively collected from hospitals in the Stockholm-Gotland region. Digital processed mammograms and analogue mammograms were available for the LIBRO1 women. Mammograms in the CAHRES study were collected from multiple hospitals in Sweden. All mammograms were analogue. The analogue mammograms were digitized at MEB using an Array 2905HD Laser Film Digitizer (Array Corp, Tokyo, Japan).

Analyses of mammographic features were performed on the mammograms using STRATUS and iCAD algorithms [188, 189].

4.2.5 Mammographic density and density change over time

Mammographic density was assessed on mammograms using the STRATUS tool developed and validated in study I. Mammographic density was used in all studies. In short, mammographic density assesses the radio dense representation of fibro-glandular tissue in the breast. The total breast area (cm^2) and the radio dense tissue is measured. Percent mammographic density was calculated as the radio dense area divided by the total breast area. Percent density was categorized into four groups referred to as cBIRADS to mimic the BI-RADS fifth edition breast composition definition [76], where BI-RADS A refers to breasts that are almost entirely fatty and BI-RADS D refers to breast that are extremely dense and lowers

screening sensitivity. Mammographic density change was studied as relative and absolute change over time of mammograms taken within minutes from each other and in mammograms taken years apart.

4.2.6 Microcalcifications and masses

In study II, microcalcifications were assessed based on the iCAD algorithm [188, 189]. In short, the iCAD algorithm is based on a deep convolutional neural network trained on radiologists' expert annotated microcalcifications and soft-tissue lesions. Microcalcification malignancy scores were trained on amorphous, coarse heterogeneous, fine pleomorphic, fine linear and fine-linear branching microcalcifications. Masses malignancy scores were trained on masses, architectural distortions, and asymmetries. iCAD uses malignancy score cut-offs to identify cancers. In study II, these cut-offs were re-trained to identify at-risk lesions on prior images to discriminate the risk of breast cancer compared to women who did not develop breast cancer. The risk scores were validated in three external datasets.

4.2.7 Differences of mammographic features between left and right breasts

Bilateral asymmetry of the occurrence of mammographic features between left and right breasts were investigated in study II. The x-ray representation of the breast tissue was investigated for risk factors of breast cancer. Pre-diagnostic images were examined for absolute differences in mammographic features in a paired breast analysis. Bilateral asymmetry between left and right breasts were investigated by region of breast tissue. Each region of breast tissue in left and right breasts were compared with each other. This approach has a statistical interesting property as both breasts have been exposed to the same personal and familial history including germline genetics, lifestyle factors, and family history of breast cancer. The remaining factor that stands out that differs between the two breasts is the disease.

4.2.8 Polygenic risk score

In study II, a polygenic risk score was used [17]. The polygenic risk score (PRS) was developed by the Breast Cancer Association Concortia (BCAC). The PRS includes 313 single nucleotide-polymorphisms (SNPs) selected based on 94,075 breast cancer cases and 75,107 controls from 69 studies in Europe. The score was developed using logistic ridge regression. The PRS was validated in an independent test set from 10 prospective studies. No evidence was found for any statistically significant interactions between the SNPs. The polygenic score predicts the probability of developing breast cancer during a lifetime.

4.3 EPIDEMIOLOGICAL STUDY DESIGN

Epidemiologists' study health and disease, and applications for health promotion and disease prevention [190]. Descriptive distributions are studied, and determinants are analyzed. In addition, experimental research investigates interventions for changing health and disease outcomes. Two main and important steps in performing a study is to define the research question and to design the study.

Epidemiological study design is defined based on two concepts: outcome and sampling of the outcome from the underlying population [191]. Common types of outcome are incidence or prevalence. Incident data is data that is collected of new exposures and disease statuses at a given time interval. Prevalent data is data that is collected of all available exposures or disease statuses in a population at a given point in time. As an example, prevalent data sampling that is based on the outcome is referred to as prevalence case-control study design. Incident data sampling not based on the outcome, but typically following the exposure, is named incidence study design.

The ideal research study is to have the full study population of interest readily available in the study and have the study individuals exposed to the factor of interest and, at the same time not exposed to the same factor. Such circumstances could ideally calculate how the exposure affects the outcome. Any other factor that could affect the association between exposure and outcome would affect both the exposed and unexposed study population equally. The association between the exposure and outcome would be causal and applicable to the whole population. This scenario is of course not possible to achieve, but epidemiological study design aims to achieve the best possible estimation of a real situation. Various epidemiological study designs have been constructed to estimate the association between exposure and outcome under different scenarios of available data. In the sections below, I briefly describe and discuss the study designs that I use in my studies.

4.3.1 Randomized controlled trial

The golden study design for performing a medical study is a randomized controlled trial due to its ability to infer causality between exposure and outcome [192]. The study design assures, at a sufficient sample size, that the baseline characteristics of the individuals in the study are evenly distributed in each of the study arms. This means that there is no bias in how any pre-exposure can affect the association between exposure and outcome at the time of baseline in the study. Statistically, this means that the randomization procedure breaks any confounding pathway between each of the pre-exposures and the association between exposure and outcome in the study.

4.3.2 Cohort study

A cohort study follows a group of individuals over time, where exposures are collected at baseline and outcomes are collected during follow-up as they occur in the study group [193]. A cohort study can also include repeated measures of exposures during the follow-up. The association between exposure and outcome is either estimated per person or per person-year. Person-years are defined as the number of years the person is at risk in the study and contributes with time in the cohort. A person that is included after study baseline is referred to as contributing with left-truncated person-time. A person that exits the study before the end of follow-up contributes with right-truncated person-time. The point in time when a person is not monitored any longer in the study is referred to as censoring.

Effects are estimated by comparing the outcome between the exposed and unexposed groups. Incidence rate (cases per person-years at risk) is possible to estimate in a cohort study and it is possible to infer causality, if any confounders are known.

4.3.3 Case-cohort study

A case-cohort study is a special case of a cohort study where all cases in the cohort that are known at a point in time during follow-up are selected [194]. In addition, a random sample of controls are selected from the cohort at study baseline. Study individuals in the random sample keep their control status throughout the study.

The case-cohort has the same statistical properties as the cohort study, but with the sample size efficiency of a nested case-control study. In the analysis, the random sample of controls needs to be up weighted to the cohort sample size. This will lead to underestimation of the sample variance, which can be adjusted using e.g. robust standard error estimation [195].

4.3.4 Case-control study

A case-control study starts by selecting cases for the study and then samples controls that are at risk of the outcome from the same underlying population [196]. Prior exposures of cases and controls are collected. The selection of controls, whether an individual is selected or not, should not be associated with the exposure that is studied. Controls are often matched to cases to increase the efficiency to estimate the association. Common matching schemes are matching by year at study inclusion (nested case-control) or e.g. by age (age matching). Case-control studies are cost-efficient for diseases with rare outcome, because all cases that would occur in a cohort could be sampled together with a smaller number of controls (commonly 5 times the controls). This is sufficient for studying the difference of exposures between cases and controls with a sufficiently small confidence intervals [197].

Effects are estimated by comparing exposed and unexposed individuals between the cases and controls. Person-years are not possible to define in case-control studies and therefore incidence rate cannot be estimated in a case-control study. However, it is possible to infer causality if the confounders are known and collected in the study.

4.4 STATISTICAL METHODS

4.4.1 Linear regression (study I, IV)

The association between exposure and outcome could be analyzed by assuming that individuals who have higher levels of exposures, e.g. age at menarche, also have higher levels of the outcome, e.g. mammographic density, compared to women with lower levels of exposures [198]. The statistical analysis of the association assumes that there will be a linear relationship between the exposure and the mean value of the outcome. Secondly, the analysis also assumes that the residual variance, i.e. difference between the actual outcome values and the estimated mean values, are the same for any value of the exposure (homoscedasticity). Thirdly and fourthly, the individuals need to be independent of each other, and the actual outcome values in relation to

the model's estimated mean values of the outcome (the error term) need to be normally distributed. There is also a rule of thumb that the number of independent factors that can be added to the model is limited by the number of observations in the model population.

Linear regression estimates a beta value, which is the estimated mean change in the outcome from one-unit change of the exposure. A beta-value with confidence intervals that includes the number zero means that there is a non-significant association between exposure and outcome.

4.4.2 Logistic regression (study I, II)

Logistic regression estimates the association between the exposure and the probability that a binary outcome occurs [199]. The association is estimated on the log scale to achieve a proportional scale of positive and negative associations around the no-association zero. Logistic regression assumes that the association between exposure and the probability of the outcome is linear. The regression also assumes that the individuals are independent, that there is no multicollinearity between exposures, and that there is no strong influence from outliers. There is also an assumption that there shall be approximately 10 to 20 events per exposure in the model.

Logistic regression estimates a change in the probability that the outcome occurs from one-unit change of the exposure. A log-odds beta-value with confidence intervals that includes zero means that there is a non-significant association between exposure and outcome. Odds-ratios are estimated by comparing the probability that the outcome occurs in exposed individuals versus unexposed individuals.

4.4.3 Penalized regression (study I)

Penalized regression models (also known as shrinkage or regularization models) have the advantage that the number of independent factors that can be added to the model is not limited by the number of individuals that are in the model population [200]. The model fitting technique multiplies the penalty term lambda to the slope and to the regression coefficients of the model. Ridge regression adds lambda to the square of the slope and regression coefficients, while lasso regression adds lambda to the slope and regression coefficients as they are. Lasso regression can use different lambdas for the slope and each regression coefficient. The slope and coefficients are not squared, and the lambdas can result in zero beta coefficient estimates. Elastic net regression adds both the ridge regression lambda and the lasso regression lambdas to the model. In consequence, lambda adds a bias to the model fit. Lambda is estimated using cross-validation. This modelling technique is potentially valuable for improving model fit on new data.

The penalized regression technique can be applied to several regression models such as linear and logistic regression.

4.4.4 Log-binomial regression (study III, IV)

Risk ratios and prevalence ratios (relative risks) could be estimated using log-binomial regression by estimating the association between exposure and the probability that the binary outcome occurs [201]. Odds-ratio approximates the relative risk only when the event is rare in the study

population. Log-binomial regression is a more general method for estimating relative risks. Log-linear regression assumes that the association between exposure and the probability of the outcome is linear. The regression interprets the outcome as the probability of success in a series of independent Bernoulli trials.

Log-binomial regression estimates prevalence ratios for prevalent associations and risk ratios for incident associations.

4.4.5 Cox regression and competing risk analysis (study II)

Cox regression estimates the association between person-years of exposure and the probability that the binary outcome occurs over the follow-up period [202]. The time-to-event survival analysis estimates hazards between exposed and outcome events, and between unexposed and outcome events, in infinite small time slice dataset over the follow-up time. Hazard ratios of the outcome compares the exposed and unexposed hazards. Cox regression assumes that the hazard ratio is the same over time, independent of time, which is referred to as the proportionality assumption. Further, Cox regression also assumes a linear association between continuous exposures and events, such that e.g. a two-times higher exposure level results in a two-times higher beta estimate of the event. In addition, Cox regression assumes that the individuals are independent. These assumptions are tested based on Schoenfeld and Martingale residuals.

Cox regression assumes that only one type of event, e.g. breast cancer, is occurring over study follow-up for a woman. In the actual situation, women could experience several events including a death event. A competing risk of breast cancer means that the woman dies from another cause than breast cancer before the woman could develop breast cancer. Competing events could be included in a model by using a cumulative incidence function to estimate the marginal probability for the competing events [203]. A marginal probability refers to the probability that a woman develops breast cancer regardless of any competing event or censoring occurring. The marginal probability does not assume any independence of the competing events. Fine and Gray developed a model using a hazard function that is based on a sub distribution function, analogous to the Cox model, but can also account the competing events [149].

Cox regression estimates a change in the probability that an event occurs from one-unit change of person-time exposure. The Fine and Gray regression, in addition accounts for competing events.

4.4.6 Model generalization (study II)

There are efficient ways to optimize a model to improve its generalization performance. One approach is as follows. A subset of women is set aside for testing the model by estimating the prediction error. The model is fitted in a second subset and validated for prediction error in a third subset, where the lowest average square error in the third training subset determines the model selection in the second dataset. This is done in an iterative process [204]. A similar approach is to perform nested cross-validation [205]. In nested cross-validation, the dataset is randomly split into e.g. ten subsets (outer loop). One of the datasets is used as the test dataset. The training of the model is done in the remaining nine subsets combined. The combined

training subset is further split into e.g. five subsets (inner loop). The model is trained in four of the subsets combined and it is evaluated it in the fifth fold. This procedure is repeated by rotating the inner loop subsets (folds). In each iteration, the average model score metrics is calculated, and eventually the model with the best hyperparameter setting is chosen. This model is then trained on the main training dataset with the nine subsets (outer loop) and the model is evaluated on the main test dataset. This procedure is repeated by rotating the test dataset in the main dataset (outer loop).

4.4.7 Non-inferiority analysis (study III)

Non-inferiority analysis estimates the difference in the proportion of individuals that show an effect from an experimental intervention compared to the proportion of individuals that show an effect from the standard intervention [206]. The estimated proportion of responders in the experimental arm is compared to a non-inferiority margin. If the point estimate confidence intervals include the non-inferiority margin, then the experimental intervention is not considered non-inferior. The validity of the non-inferiority analysis relies on the constancy assumption which means that the effect of the standard treatment that is reported in the current trial is consistent with the effect that has been observed in previous trials. In study III, the proportion of mammographic density responders was 50%, similar to what was reported in previous trials. The validity of the non-inferiority analysis also relies on that the difference between the standard dose and tested lower dose are not compromised by study design or procedure.

4.4.8 Potential outcome analysis (study IV)

Potential outcome analysis estimates the association between exposure and an outcome that follows if the individual would have had the exposure [207]. The exposure is not actually occurring but is counterfactual. Potential outcome analysis is commonly used to study causation between exposure and outcome but could also be used to study counterfactual associations in general, e.g. to study feasibility of a planned study.

Risk difference, risk ratio, and odds ratio can be estimated in potential outcome analysis using the same statistical methods that are used for estimating associations between exposure and outcome based on factual data.

5 RESULTS

5.1 STUDY I

In all, 45,417 women from the KARMA, LIBRO1, and SASBAC cohorts were included in the study. After development of the mammographic density tool in healthy women the relative risk of breast cancer was estimated in a case-control setting using three datasets. The risk association was estimated to 1.6 (95% confidence interval CI 1.3-1.8) per standard deviation averaged over the three studies after adjustments for lifestyle and familial breast cancer factors. The case-control discrimination was AUC 0.62 (CI 0.60-0.64). The type of image did not influence the risk association. The alignment protocol that was developed and evaluated decreased the non-biological variability observed in density change and did re-estimate the yearly overall percent density decrease seen in aging from 1.5 to 0.9 percent, $p<0.001$.

Table 7. Odds-ratios and 95% confidence-intervals of breast cancer per standard deviation from density measures in processed, raw, and analogue mammograms.

Case – control study sample	Model 1[1]	Model 2[2]	Model 3[3]
1a. KARMA (processed)	1.6 (1.5-1.7)	1.7 (1.6-1.8)	1.7 (1.6-1.8)
1b. KARMA (raw)	1.6 (1.5-1.7)	1.7 (1.6-1.8)	1.7 (1.6-1.8)
2. LIBRO1 / KARMA (processed/analogue)	1.5 (1.4-1.6)	1.6 (1.4-1.8)	1.6 (1.4-1.8)
3. LIBRO1 / SASBAC (analogue)	1.5 (1.3-1.7)	1.5 (1.3-1.8)	1.5 (1.3-1.7)
Study samples combined	1.5 (1.3-1.6)	1.6 (1.3-1.8)	1.6 (1.3-1.8)

[1]Model 1 - percent density and age.

[3]Model 2 - percent density, age, and BMI.

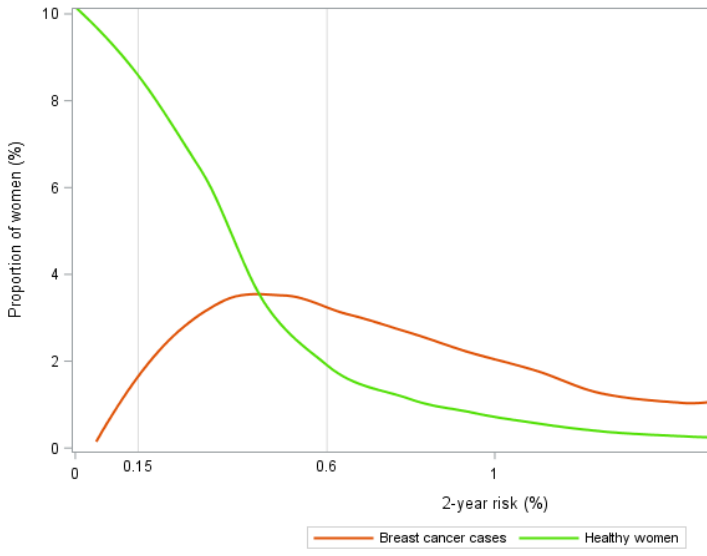
[3]Model 3 - percent density, age, BMI, ever use of HRT, menopause status, and family history of breast cancer.

The risk tool has the potential to be used for assessing risk, masking, and mammographic density change over time.

5.2 STUDY II

The KARMA case-cohort included 10,350 women sampled from the KARMA screening cohort that includes 70,877 women recruited between 2011 and 2013. The case-cohort consisted of the available 974 incident breast cancers and a random sample of 9,376 healthy controls from the KARMA screening cohort. The discrimination performances of model 1, model 2, and model 3 was 0.73, 0.74, and 0.77, respectively. The Hosmer-Lemeshow model fit statistics was non-significant for each of the models. The AUCs of the three external validation cohorts CSAW, MBTST, and independent KARMA validation cohort were 0.73, 0.71, and 0.73, respectively. In comparison, the established risk models Tyrer-Cuzick, Gail, and PRS showed AUCs of 0.62, 0.61, and 0.64, respectively.

Figure 14. Frequency distribution of 2-year absolute risks for developing breast cancer in cases and healthy women in the KARMA case-cohort using Model 3.



Risk groups[1]	Percent women at risk	Absolute 2-year risk (%)	Relative risk[2]
0-0.15 (low)	26.7	0.09	0.3
0.15-<0.6 (general)	48.2	0.29	1.0 (reference)
0.6-<1.6 (moderate)	17.3	0.87	3.0
≥1.6 (high)	7.8	2.70	9.4

Microcalcifications and masses showed the strongest risk associations, followed by polygenic risk score and mammographic density. The left-right breasts differences of microcalcifications and masses added to the risk association. The risk association in screen detected and interval detected cancers were 8-fold in the high-risk group compared with the general risk group. The model predicted more often cancers with large tumors ≥20mm) and stage II tumors. The distribution of two-year absolute risks for developing breast cancer in cases and healthy women in the KARMA case-cohort using Model 3 is presented in the figure.

The short-term risk model has the potential to identify a group of women that currently are sent home from screening but are more likely to come back with an interval cancer or a cancer at the next mammography screen.

5.3 STUDY III

Women from the Swedish mammography screening program were invited to participate in the KARISMA randomized controlled trial. In all, 2,314 volunteering women were investigated for eligibility to the study. There were 566 premenopausal women and 873 postmenopausal women included in the study. The premenopausal women showed non-inferior reduction in mammographic density after exposure to 2.5, 5, and 10 mg tamoxifen compared with the standard dose of 20 mg. Postmenopausal women showed no reduction in mammographic density. Severe vasomotor symptoms were reduced by approximately 50% in the 2.5, 5, and 10 mg groups compared to the standard dose. The 2.5 mg group were therefore identified as the lowest dose that showed non-inferior reduction of mammographic density reduction with less severe side-effects. The lack of the mammographic density reduction effect in the postmenopausal women was not due to the fact that postmenopausal women had lower initial level of mammographic density at baseline. Below figure presents the non-inferiority analysis of the proportions of responders in the intention to treat population.

Figure 15. Non-inferiority analysis of proportion of responders after six months of tamoxifen in the intention to treat population.

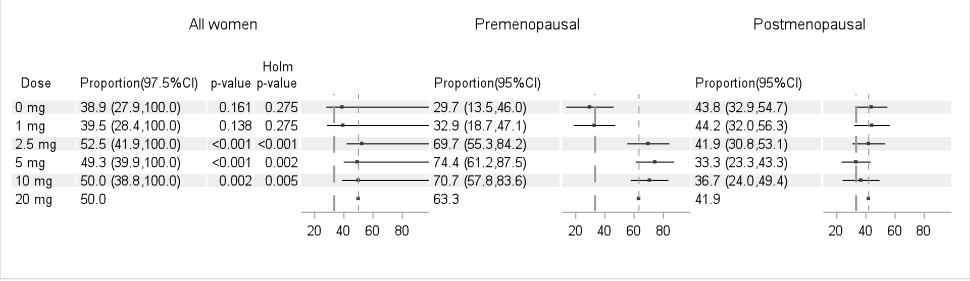
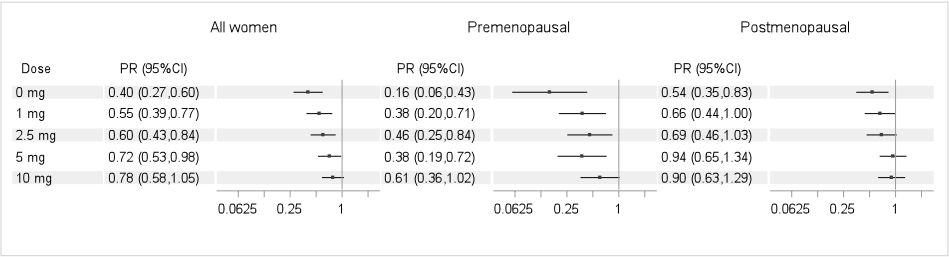
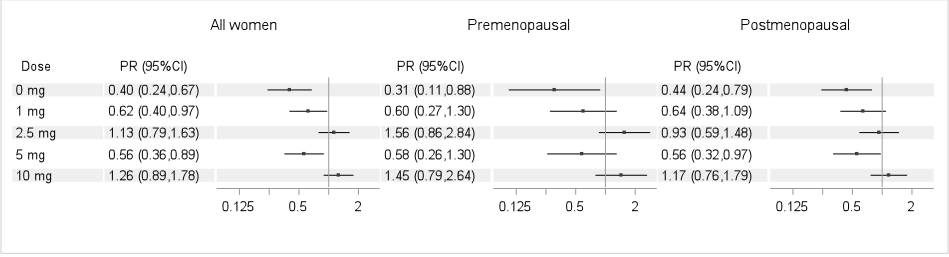


Figure 16. Prevalence ratios of severe vasomotor symptoms after six-months of tamoxifen in the intention to treat population.

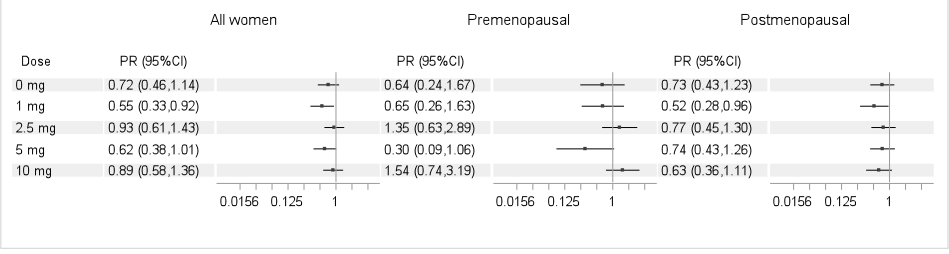
A.Vasomotor severe events



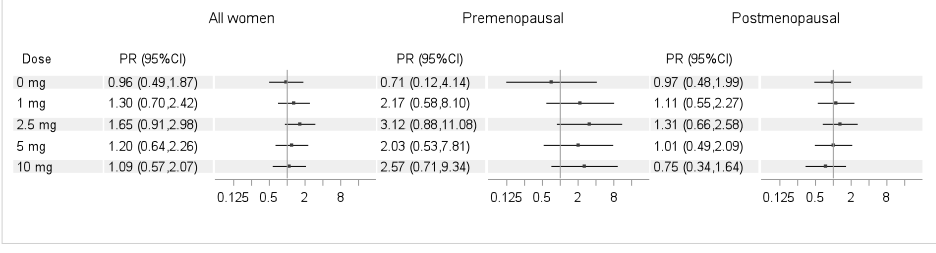
B.Gynecological severe events



C. Sexual severe events



D.Musculoskeletal severe events

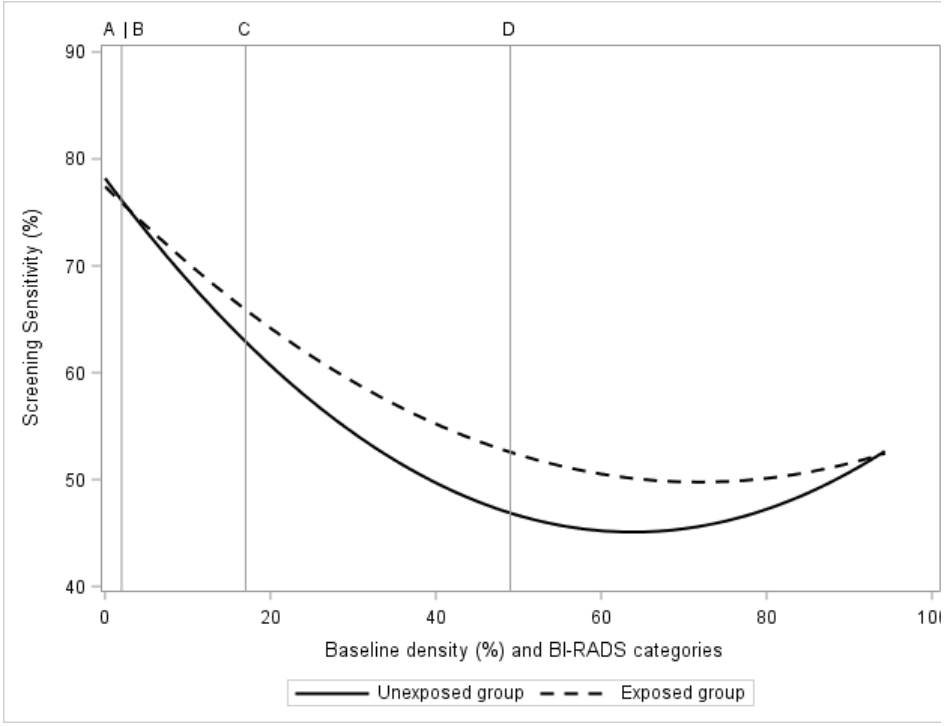


Low-dose tamoxifen has the potential to reduce breast cancer incidence in currently healthy women.

5.4 STUDY IV

Premenopausal women in the KARMA prospective mammography screening cohort were included in this pilot study. In all, 28,282 women had mammograms with two-year follow-up. There were 287 screen-detected and 230 interval cancers in the study population, with available screening mammograms at the time of diagnosis of a screen detected cancer or with an available screening mammogram prior to diagnosis of an interval cancer. The screening sensitivities in the KARMA cohort were 76%, 69%, 53%, 46% for BI-RADS density categories A, B, C and D, respectively. After the potential exposure to tamoxifen, the modelled screening sensitivities increased by 0% (p=0.35), 2% (p=<0.01), 5% (p<0.01), and 5% (p<0.01), in the four density categories. A potential relative density decrease by ≥20% could reduce tumour sizes >2cm at the time of detection by 4% (p<0.01). The table presents the number of interval cancers per 100,000 age standardized screening premenopausal women together with the change in numbers of interval cancers by percentage mammographic density decrease. The unexposed group is included as the reference. In the exposed group, more than fifty percent of the women experienced a ≥20% relative density reduction. This group included 24% of the interval cancers that potentially could be identified on the earlier screening mammogram.

Figure 17. Screening sensitivity in the unexposed and exposed groups by mammographic breast density at baseline.



Screening sensitivity,	BI-RADS density category				Low	High	Low+
category mean (%)	A	B	C	D	A+B	C+D	High
Unexposed group	76	69	53	46	70	51	56
Exposed group	76	71	58	51	72	55	60
Difference	0	2	5	5	2	4	4
p-value	0.35	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01

In the KARMA exposed group, the figure presents screening sensitivity by mammographic density at baseline. Baseline mammographic density is presented as a regression plot and as categories of computer-generated BI-RADS density categories A, B, C, D. Low density is defined as categories A+B and high density as C+D.

6 DISCUSSION

6.1 STUDY I

We developed an automated measurement tool that could measure mammographic density on different types of mammograms from different vendors. At the time the project was initiated, processed images were considered unusable for image analysis, due to the processing that vendors performed on the raw mammograms. The processing was done to better show potential tumors, thereby reducing mammographic density in the images. The study completely changed that usability of the mammograms since the tool was the first that measured mammographic density on processed images.

A measurement at single points in time is usable to assess risk and masking of breast cancer, while longitudinal series of mammograms could be used to assess mammographic density as a marker for therapy response. Recent developments of other tools also use this approach [100].

Mammographic density is a well-established risk factor for breast cancer. The risk association could be explored at a large scale through the use of the automated tool. Masking is a main challenge in mammography screening and delays detection of cancers. The density tool enables studies where mammographic density better can be used to guide clinical decision for follow-up of women that are at high risk of breast cancer and in addition have high mammographic density.

Mammographic density has proven to be an excellent early marker of response to tamoxifen therapy and breast cancer incidence reduction. Given the challenges that has been observed for radiographers to take mammograms, the density alignment tool could help to reduce the non-biological variability that otherwise obscures the true mammographic response to therapy.

The study was limited by the lack of a publication for validating the density tool in non-Swedish populations. The density tool is currently being tested in other populations. Screening modalities are constantly developing and tomosynthesis become more commonly used. In addition, contrast enhanced mammography is more frequently used as an alternative to magnetic resonance imaging.

Further development of the density tool is needed to assess mammographic density on modalities other than digital mammography and analogue films.

6.2 STUDY II

We developed a short-term risk model based on mammographic features with the advantage that it could be used to automatically assess risk in the general population at mammography screening units. At the time the project was initiated, risk assessment was done mainly in the research setting where lifetime, ten year, or five-year projection were estimated. In the clinical setting risk models were not used to assess risk in the general population and no risk model was designed to improve screening. Mammography screening is a well-established examination procedure for early detection of breast cancer. Age-based screening is efficient for reducing

breast cancer mortality but is suboptimal for women that develop breast cancer that is not identified by the screening. Suggestions have been made to move towards a more efficient screening scheme [208]. The image-based model could be further expanded with lifestyle factors including BMI which brings a public health message. Family history of breast cancer is an important factor, not the least for a risk tool to be approved by regulatory authorities where clinics can be reimbursed for assessing risk. The model could further be expanded with a polygenic risk score, which improves the model substantially to well-define the low-risk group. This positive effect of using a polygenic risk score to well-define the low-risk group has been reported recently [209]. One potential value from using the full risk model including genetic determinants is for women in the 40's. The women found with a low risk could be recommended to come back in 10 years' time to participate in regular biennial screening.

Risk models based on lifestyle factors or polygenic risk scores has been shown to identify the less aggressive oestrogen-receptor positive cancers to an increasing extent with a higher risk score. Our model is the first to target interval cancers and large cancers. Women that were identified in our model had both oestrogen-receptor positive and oestrogen-negative tumors, larger tumor sizes, and were more often stage II cancers. More work is needed in this field to improve the ability to identify more aggressive cancers. A main challenge is that the more aggressive cancers are only a few ten percent of all cancers and therefore naturally contributes with the smallest amount of information in the training step of the risk model.

In a post hoc analysis, a cancer detection tool was used to assess the short-term risk to complement the risk model comparison that was published including the Tyrer-Cuzick, Gail, and PRS risk models. The iCAD detection tool showed an AUC of 0.69. This makes sense because a proportion of the cancers that are identified in the short time interval of two years are missed cancers. This addressed a question of what distinguishes a short-term risk model from detection of a cancer. In discussions with FDA, who approves clinical devices, it was clear that FDA defines cancer detection as the identification of a malignant lesion in the breast. Based on that, I defined short-term risk assessment, using mammographic features, as the identification of the breast with a malignant predisposition, but without identifying any specific lesion or region in the breast.

In this study, a case-cohort was used instead of the full cohort. The advantage is that a smaller group of women can be used to perform the time-consuming image analysis. By up weighting the controls, the full cohort can be recreated using robust standard error estimates. The analysis was also performed using logistic regression instead of Cox regression to estimate the risk factors. A Cox regression could possibly have been preferred but was more time-consuming to perform. The timeframe was fixed to two years in the analysis which would lead to comparably smaller difference in estimates. In model updates that will include five-year projection, time time-varying regression will be used. The analysis was controlled for age.

Short-term risk versus detection also brings up a question whether the model identifies women with already existing cancers, or cancers that are developing quickly within the two-year time interval. To clarify this matter, it is valuable to discuss cancer development. There is a point in time when a cancer is initiated. Cancer initiation happens at the first genetic mutation,

provided that this mutation later leads to a cancer. This subclinical phase is difficult to estimate but is normally assumed to be ten years. There is a second point in time when a cancer theoretically could be detected based on the detection tool that is used to identify cancers. In the mammography screening context this point in time is referred to as the sojourn time. Studies show that this time is on average three years prior to diagnosis. There is a third point in time when the cancer is diagnosed. Interval cancers consists mainly of two groups, i.e. missed cancers and fast-growing cancers. They are approximately in a proportion of 50-50. This means that the short-term risk tool targets both interval cancers that are missed at prior screen and cancers that are fast-growing. The short-term risk model targets interval and large cancers at the next screen that was not detected at prior screen. It should be underlined that from a clinical point of view the question whether the cancer is already existing is of lesser importance. The screened woman is sent at home without additional follow-up.

A study showed that women who were recalled and were found to be false positives had a doubled risk for developing a breast cancer within four years after the false-positive recall [210]. Another study showed that women diagnosed with in-situ carcinoma have a four-fold risk to later in life develop invasive carcinoma [28]. Recent studies showed that the breast cancer incidence may be underestimated in mammography screening based on digital mammography [211, 212]. At the same time, over diagnosis of clinically irrelevant cancers is a major challenge in mammography screening [4].

No study has shown that supplemental screening reduces breast cancer mortality. However, today there is no established clinical praxis for when to perform supplemental screening and how to perform the procedure. This challenges a systematic investigation of the value of supplemental screening.

The main challenge when developing risk models is to judge the model accuracy in independent populations. Statistical regulation is performed to that end on the training dataset to make the model more generalizable. Our study was limited by not having access to more than two Swedish external datasets. The model accuracy is affected by different screening strategies, screening modalities, ages included, screening intervals, personal screening histories, and population ethnicities. A common strategy is to further train the model by including in the training dataset those population settings to enhance the model further.

Additional external validation is needed in screening settings with different screening routines, screening modalities, and ethnicities.

6.3 STUDY III

We showed that lower doses of tamoxifen could reduce mammographic density to the same extent as standard 20 mg dose with less severe vasomotor symptoms. Prior to the study was performed, tamoxifen was used in the adjuvant setting to reduce recurrence of breast cancer and in the neo-adjuvant setting to reduce tumor burden prior to surgery. Tamoxifen was also in limited use in the preventive setting to reduce breast cancer incidence. The main challenge was the severe side effects that caused women to discontinue the treatment. There was a need

to investigate alternative approaches to increase uptake in the population in the preventive setting. A study had shown that a lower dose could be used to reduce intraepithelial neoplasia, but it was unknown which women benefit from the treatment [173]. It was known that mammographic density in the adjuvant and preventive setting was an early marker for which women benefit from the therapy. A lower dose could potentially lead to less severe side effects, but the efficacy of the density reduction needed further investigation.

Our study was limited by a low participation rate and that no difference was seen in adherence between the study arms. Women with a family history of breast cancer were more likely to participate and potentially more loyal to the study protocol. The low participation rate did not likely lead to bias as no factor was identified that affected both participation rate and the study outcome. Furthermore, a low participation rate and adherence do not necessarily affect the results of a dose-determination randomized controlled trial. The potential for increasing uptake in the population when using low-dose tamoxifen has been discussed recently [213].

Further studies are needed to test if a reduction of breast cancer incidence and an increase in uptake in the population follows from the use of low-dose tamoxifen.

6.4 STUDY IV

We presented the first study of the tamoxifen effect on mammography screening sensitivity. We modelled the influence of 2.5 mg tamoxifen on mammographic sensitivity in premenopausal women. At the time, paper III had shown that low-dose tamoxifen reduced mammographic density as efficient as full dose of tamoxifen. We hypothesized that the mammographic density, in addition to be used for reducing breast cancer incidence, could also improve screening sensitivity. Low-dose tamoxifen could decrease mammographic density and reduce the number of interval cancers by identifying the cancers at smaller tumor sizes already at prior screen. The group of women that could potentially benefit from low-dose tamoxifen are women at high-risk of breast cancer that in addition have dense breast and a masking problem. It has been shown that all subtypes of cancer, not only oestrogen-receptor positive cancers, have a higher probability to be detected in women with low mammographic density compared to women that have high mammographic density. Women using low-dose tamoxifen could therefore benefit from early detection regardless of which cancer subtype they develop, an aggressive or a non-aggressive tumor. The modelling study was a feasibility study to investigate the potential use of low-dose tamoxifen to increase mammography sensitivity.

Our study estimated the number of interval cancers that potentially could be reduced, assuming that a certain threshold of mammographic density reduction was sufficient to identify an interval cancer already at the prior regular screening mammogram. The median relative reduction of mammographic density in the study was ~20%. Studies on prevention have shown that women with at least a median density reduction responds to therapy with a reduced breast cancer incidence [112, 114]. However, it has not been shown that mammographic density or this specific cut-off of relative density reduction will improve screening by reducing the number of interval cancers.

The study was limited by the potential outcome analysis. Mammographic density responses were available from the KARISMA randomized controlled trial and were used as the reference of the expected density responses that would be observed in a large screening cohort, if they had used low-dose tamoxifen. Due to the limited number of density responses in KARISMA the density response variance was limited in KARMA.

Further studies are needed to confirm that the density reduction induced by low-dose tamoxifen results in increased sensitivity that in turn translates into earlier detection of breast cancers of any subtype and thereby reduces breast cancer mortality.

7 METHODOLOGICAL CONSIDERATIONS

7.1 BIAS, CONFOUNDING, AND VALIDITY

Selection (sampling) bias

Statistical methods are well developed for assessing the probability that a point estimate is with in the confidence intervals. This estimation is commonly reported with a confidence of 95%, i.e. it is a 95% certainty that the point estimate is within the given interval. The point estimate certainty is based on the assumption that the data is a representative sample, a random sample of sufficiently large size, of the underlying population which the point estimate is trying to assess. In the medical research this assumption is often violated due to difficulties of selecting individuals for the study.

The possibility to perform a random draw from an underlying population is challenged by the availability of the entire pool of individual for the draw. Studies often rely on volunteering study participants. This concept jeopardizes the availability of the entire pool for the draw. Another challenge is that research budgets struggle with the cost for achieving a sufficiently large sample to test the research question. Power analyses helps in this respect but are at the same time assuming the possibility to make random draws, which often is challenged. Factors affecting the selection of individuals could affect how exposure and outcome are selected and could therefore result in a bias. A low participation rate per se does not imply selection bias. However, selection bias could affect the association between exposure and outcome if the factors that is responsible for the selection bias is associated with both exposure and outcome.

Misclassification (information) bias

In order to study the association between exposures and outcomes, the exposures and outcomes needs to be defined. There will always be gap between the data definitions and the true values. This means that the exposure and outcome definitions could include misclassifications. All information that is collected, measured, or categorized is subject to this bias. e.g. information from images, biological samples, survey-based interview data, medical records, or register data.

The misclassification is either differential or non-differential. Differential misclassification means that the exposure or outcome definitions are not equally wrong in the two contrasting groups usually referred to as cases and controls. Example of an outcome misclassification is what occurs in a retrospective study based on survey data, where women with breast cancer are more likely to be aware of what exposures affects breast cancer as compared to women who remained healthy during the study period. The breast cancer cases are therefore more likely to better recall breast cancers among their relatives as compared to the women who did not develop breast cancer. This means that the exposure definitions would be more accurate among the breast cancer cases as compared to the non-breast cancer cases.

Lead-time bias

A special case of misclassification in screening studies is lead time bias. Mammography screening is an example where women are diagnosis with a cancer earlier than otherwise would have been the case if the women did not participate in mammography screening. When estimating survival from breast cancer the lead time from diagnosis to death will be longer for screened women compared to non-screened women because of earlier detection. This systematic difference in lead time is referred to as the lead time bias where survival time is overestimated due to early detection.

Length-time bias

A special case of selection bias is length-time bias. Mammography screening is by design sensitive to detect cancers that are slowly growing. In comparison, women who do not attend mammography screening will detect slow growing cancers symptomatically at a later time. The time from detecting a slowly growing cancer to death is systematically longer in women attending mammography screening compared with the corresponding time among women not attending screening. This means that mammography screening leads to a potential overdiagnosis of cancers that are progressing slowly.

Confounding

The association between exposures and outcomes could be affected by a common cause. This leads to that the association between exposures and outcomes could be wrongly estimated. This bias is referred to as confounding. The true effect could therefore be outside the confidence intervals that was estimated from the biased model. If the confounder of an association is known, it can be controlled. A study could handle confounders by design through matching cases and controls on the confounders. In an analysis situation, associations could also be estimated using regression analysis or stratified analysis after controlling for the confounders. Confounding could also occur due du unknown factors skewing the association results. This could be addressed by study design in a randomized controlled trial. A randomized controlled trial assures that any possible known or unknown common cause of the studied association between exposures and outcome is controlled for at baseline of the trial. This means that all pre-exposures are equally distributed in all study arms and, are non-differential to the trial exposure and outcome.

Validity

Bias could lead to overestimation or underestimation of the true effects. Bias could be reduced by increasing validity, i.e. by decreasing systematic error. This is done by reducing selection bias, differential misclassification, and confounding. An increase in sample size will not affect the validity if the bias is present in the sample.

7.2 STUDY I

We addressed the concern that large screening cohorts were available for potential analysis of mammographic density and breast cancer, but there did not exist a tool for measuring

mammographic density on processed images. In order to develop the mammographic density tool, three cohorts (KARMA, LIBRO1, and SASBAC) were used with a good representation of mammograms of different types from different vendors.

The mammographic density tool could not be developed based on the contrast between breast cancer cases and healthy controls or the contrast between screen detected cancers and interval detected cancers, as this would lead to a training effect that could overestimate the risk association between mammographic density and breast cancer or screening sensitivity. The mammographic density tool was therefore developed using healthy controls in the KARMA and SASBAC cohorts. The KARMA and SASBAC cohorts included digital and analogue mammograms from different vendors.

The FDA approved iCAD iReveal mammographic density measurement tool was used as the reference measure. iReveal measured mammographic density on original raw mammograms. Through the data collection in KARMA both raw and processed images were collected from the same women. This made it possible to use iReveal as the reference measure on the raw mammogram and train the STRATUS mammographic density tool to measure mammographic density on the corresponding processed mammogram.

STRATUS learned to measure mammographic density based on image features and image tags information in the processed images using supervised machine learning. Mammographic features were analysed, and a lasso penalized linear regression models were used to predict the reference mammographic density measure. The mammographic features were weighted differently for different mammogram types and vendors. The weights could be estimated using penalized lasso regression, including a weight of zero. This meant that the model by itself could find those features that were relevant. The regularization technique with cross-validation also made the model more generalizable. Other model techniques were also tested, such as support vector machines, but the penalized lasso regression model performed as well as the other modelling techniques.

The STRATUS tool was tested on an independent selection of women in the KARMA, LIBRO1, and SASBAC cohorts. Spearman correlation statistics and Bland-Altman fit plots were used to assess the measurement accuracy on square root transformed density measurements to achieve an approximation of normal distribution.

Based on the development and validation of the density tool, the association between mammographic density and breast cancer was estimated. Through the literature it was known that the association between mammographic density and breast cancer was confounded by age, BMI, menopausal status, family history of breast cancer, and use of hormonal replacement therapy. An augmented case—control study design was used to include the three cohort KARMA, LIBRO1, and SASBAC where available breast cancer cases were matched in one-year age bands. Logistic regression models were fit with breast cancer case status as the outcome and mammographic density as exposure, adjusted for the known confounding factors.

An alignment protocol was developed to better measure mammographic density changes over time. By observing numerous mammograms, it became clear that the radiographers who take mammograms work with a degree of imprecision when it comes to compressing and placing the breast into the mammography machine. Through estimation of breast positioning on the digital mammogram we saw that approximately 10% of the mammograms substantially deviated from recommended breast placement recommendation. The KARMA cohort was used for the study, including one dataset for training the alignment protocol and one dataset for evaluating the alignment protocol. Rigid global registration was the simplest method and as good or better than the alternative registration methods.

The training was performed on women who had repeated screening mammograms taken within minutes from each other due to that the radiographer considered the first mammogram to be suboptimal. This is data that is not stored at the hospitals, but due to our data collection protocol this phenomenon could be studied.

7.3 STUDY II

We addressed the need for developing a short-term risk model to improve mammography screening. A case-cohort was constructed based on the available breast cancers in the KARMA cohort at the register linkage in 2017 (N=974) and a random sample of 9,376 controls. The women in the random draw that were breast cancer cases at the register linkage in 2017 were included as controls in the study.

Mammographic features, (mammographic density, microcalcifications, masses) and differences in the occurrences of mammographic features between left and right breasts were measured for the women. In addition, lifestyle risk factors and family history of breast cancer was collected, and a polygenic risk score were estimated.

The mammographic features included bilateral asymmetry between left and right breasts. These features have a statistical interesting property, because both breasts have been exposed to the same personal and familial history including germline genetics, lifestyle factors, and family history of breast cancer. This means that the remaining factor that stands out that differs between the two breasts is the disease.

Risk models were developed to predict breast cancer diagnosis in the 2-year time interval after a negative mammogram on the basis of the mammographic, lifestyle, familial, and polygenic risk factors. The first risk model included mammographic features and age only. The second model included model one plus lifestyle and familial breast cancer risk factors. The third included model 2 and in addition a polygenic risk score. Logistic regression was used to estimate the associations between the risk factors and the breast cancer outcome. Absolute risks were calculated based on the achieved estimates in addition to Swedish national incidence rates and competing mortality risks, and risk factor exposure prevalence in the KARMA cohort. The absolute risks were further categorized using the NICE guidelines into high, moderate, and general risk. A fourth low-risk category was defined in the low end of the general risk category.

External validation was performed in the CSAW and MBTST screening cohorts including approximately ten thousand women each and in total approximately 700 incident breast cancers. A third independent validation cohort was extracted from KARMA including breast cancer cases diagnosed after the breast cancer that were included in the KARMA case-hort. The independent validation cohort also included a sample of approximately ten thousand controls from the KARMA cohort.

The short-term risk model was compared with established long-term risk models Tyrer-Cuzick, Gail, and PRS designed for breast cancer prevention.

The generalizability of a model is the ability predict the risk accurately in new data. In the model training, statistical generalization techniques were used based on women that attend the Swedish mammography screening setting. New data to predict can be new women in the same screening setting or new women in other screening settings. Population ethnicity, screening modality, mammographic feature risk associations, screening intervals, personal screening history, and incidence rates are examples of factors that determines a screening setting. A model construct that is developed in one such screening setting has the potential to generalize to new women in that setting. Further model development should be needed to generalize in other screening settings.

7.4 STUDY III

We addressed the need to efficiently reduce breast cancer incidence by using a medication. For this purpose, a study would be needed that included approximately 100,000 women due to the low incidence of 0.5% breast cancer cases per 1,000 screened women. Previous research showed that mammographic density is an early marker of women that responds to tamoxifen and experience a lower risk of developing breast cancer. To this end, 159,027 women in the Swedish mammography screening program, aged 40-74, were invited to the six-months double-blind six-arm randomized placebo-controlled dose-determination KARISMA trial. Approximately 1.4% or 2,314 of the women volunteered and were investigated for eligibility. A randomization procedure was used that assured that the medication allocation was masked for all parties involved in the study execution. Placebo, 1, 2.5, 5, 10, and 20 mg of tamoxifen was tested. The main outcome was defined as a non-inferior reduction of mammographic density causing fewer severe side-effects in the lower doses compared with the standard dose of 20 mg. Mammographic density was measured after aligning time series of mammogram, to account for radiographer variability in performing mammograms. Both intention to treat and per protocol populations were analysed. A sensitivity analysis was performed on the full population after multiple imputation. Post hoc analyses identified that menopausal status was a key to understand differential mammographic density effects among the study participants.

7.5 STUDY IV

We considered the potential use of low-dose tamoxifen to improve mammogram sensitivity in a pilot study. It is known that women who use the oestrogen hormonal replacement therapy increase in mammographic density and have lower screening sensitivity. We assumed that

tamoxifen, an anti-oestrogen proven to reduce mammographic density, could improve screening sensitivity. The density response to tamoxifen was known from the KARISMA trial and the density dependent screening sensitivity were known in the KARMA cohort. Potential outcome analysis was performed including the available 28,282 premenopausal women in the KARMA prospective screening cohort. Mammographic density is shown to be associated with screening sensitivity and the tumor size that is registered at the time of diagnosis. For this reason, two models were fitted to estimate the screening sensitivity and tumor size dependence on mammographic density in the KARMA cohort. The density response in the KARISMA trial were thereafter investigated for any association with lifestyle and familial risk factors. No association was found and the density response in the 2.5 arm in the KARISMA trial was therefore applied to the KARMA screening cohort using a random distribution. Analysis was performed where the KARMA women were compared to themselves with and without the estimated density effect of tamoxifen. Interval cancer rates and tumor sizes were compared before and after a potential exposure to tamoxifen.

8 ETHICAL CONSIDERATIONS

Key concepts for performing research is to assure the safety and integrity of the study participant. The study participants shall have the right to a well-balanced study information prior to actively choosing to accept or not to accept to participate in the study. The European Council established a convention on human rights and biomedicine to regulate the ethical process including the operation of ethical review boards that approve research. The General Data Protection Regulation (GDPR) legislates the use of sensitive personal data such as informed consent and data based on surveys, mammograms, blood samples, and registers. Personal identifying information must be removed from the data prior to the research use of the data.

In the KARMA study women who attend mammography screening were recruited at study centres located near the mammography screening units. Women were given detailed written and oral information before giving informed consent to participate in the study. The women contributed with survey-based information, mammograms, donated blood, and gave the permission to access register-based data and medical records. The data was collected and handled according to the officially available and GDPR compliant routines at the department of Medical Epidemiology and Biostatistics, Karolinska Institutet. The women in the LIBRO1 study and the CAHRES study similarly contributed with survey-based information, mammograms, donated blood, and gave the permission to access register-based data including cancer and vital status, and medical records. The data was collected and handled according to the routines held by the department at these time periods. The KARISMA trial included women who were recruited when performing mammography screening. The same procedure was followed as for the KARMA women. The CSAW study was based on register data including mammograms and cancer and vital status register data. The MBTS study similarly was based on register data including mammograms and cancer status register data. Personal identifying IDs were exchanged with study specific IDs to protect the study participants from revealing personal identifying information.

Ethical approval was granted by the regional ethics board at Karolinska Institutet for all studies but for MBTS that was granted ethical approval by the regional ethics board at Lund University:

- KARMA 2010/958-31/1 and 2013/2090-32
- KARISMA 2016/65-31/2
- LIBRO1 2009/254-31/4 and 2011/2010-32
- CAHRES/SASBAC 155/93, 2006/1350-32
- CSAW 2016/2600-31
- MBTS 2009/770

Register-based data is governed under the same regulation as all sensitive personal data. However, due to routines of de-identifying the data some of the data may not be classified as

sensitive personal data. For this reason, informed consent may not be required. Under no circumstances could personal identifying information be revealed in the research data.

9 CONCLUDING REMARKS

In this thesis I investigate ideas for improving mammography screening and prevention of breast cancer. Considering that one woman in eight develops breast cancer during the lifetime in the Western world, measures are warranted for reducing mortality and to prevent breast cancer. In this thesis I developed tools for assessing mammographic density and breast cancer risk. We also developed one low-dose tamoxifen concept that reduces mammographic density for the potential use in breast cancer prevention, and I developed one low-dose tamoxifen concept that reduces mammographic density for the potential use to improve screening sensitivity.

The study populations that were used to develop the tools and concepts were mainly KARMA and KARISMA. The prospective KARMA cohort are women attending mammography screening in the Stockholm area and southern of Sweden in 2011 till today. Approximately 35% of the women who attend mammography screening participated in the KARMA study. An increased proportion of women with a family history of breast cancer was observed, which may have affected our results in addition to other non-measured confounding. Due to the prospective study design, the women were not aware of a later cancer outcome, which makes any classification bias of risk factors and breast cancer outcome less likely differential.

Women who attended mammography in the Stockholm area were invited to participate in the KARISMA study. Approximately 1.4% of the invited women participated in the KARISMA study between 2016-2019. This phase II study was a double-blind randomized and placebo-controlled non-inferiority dose determination study. This means that the baseline characteristics of the women were equally distributed in the study arms. In consequence, this means that the observed differences in density reductions between the study arms after tamoxifen exposure have low bias.

In study I, I developed and evaluated a mammographic density tool for automated assessment of radio dense fibro-glandular tissue. The measurement could be used on processed and raw images and on digital and analogue mammograms origin from different vendors. This made it possible to access vast resources on the hospitals that was previously difficult to include in research studies. The measurements were also suited for longitudinal studies to follow density change over time. The tool was valuable for assessing the three key concepts risk, masking, and therapy response to therapy.

In study II, I developed and evaluated a risk assessment tool for assessing short-term risk of breast cancer. The work introduced the concept of constructing a risk model using mammograms as the main component. The rational was to make use of the infrastructure that is available at mammography screening units. The work constructed the first risk model with the aim to improve mammography screening. This concept is now recognized as the second clinical use case for a risk model.

In study III, the lowest tamoxifen dose was identified that was non-inferior to standard dose of 20 mg tamoxifen to reduce mammographic density and to show less side-effects. The

overarching aim for performing the project was to improve adherence of a medication that could be used for breast cancer prevention.

In study IV, I developed a concept based on low-dose tamoxifen to improve screening sensitivity and to reduce interval cancers and large cancers. The rational for performing the study was that it is known that low-dose tamoxifen reduces mammographic density to the same extent, but has less side-effects, compared with standard 20 mg dose. It is also known that screening sensitivity is higher for any type of breast cancer in women who have low mammographic density. We modelled the effect of how interval cancers and large cancers could be reduced by identifying those cancers at prior mammography screen. The study was done as a pilot to investigate the feasibility of performing a large-scale study on the same subject.

The concepts developed in this thesis have a huge potential for clinical use. Any follow-up of the use potential of this research requires clinical prospective trials to validate the risk assessment tool and the low-dose tamoxifen therapy.

10 FUTURE PERSPECTIVES

Today, one woman in eight develop breast cancer in her lifetime. Mammography screening reduces mortality by approximately 20%. Prevention strategies are scarce due to severe side-effects. Women who develop breast cancer in a short time after a negative mammogram are between two chairs of clinical strategies; detection of breast cancer and risk assessment for prevention of breast cancer.

In the thesis I describe development of tools that are needed to move from age based to risk based screening by improving mammography screening and to improve breast cancer prevention by reducing therapy side-effects. A risk model was developed to improve mammography screening of interval and large breast cancers. The model could potentially be used to reduce breast cancer mortality. A low-dose tamoxifen therapy was developed to reduce mammographic density. Low-dose tamoxifen was shown to reduce mammographic density to the same extent as the clinically accepted full dose. A mammographic density reduction could contribute to future screening efforts by increasing the sensitivity of a mammogram. A density decrease has also proven to be a good proxy for tamoxifen therapy response thereby indicating that low-dose tamoxifen has a potential preventive effect.

The thesis is the starting point for potential validation studies. The work for the 2020s is outlined as follows.

As a first follow-up study, I suggest a clinical prospective trial for evaluating the risk assessment tool. The study should be done in multi-centre European and US settings. Several countries in Europe provide organized national mammography screening programs, while US provides regional screening programs and opportunistic screening. In Asian countries opportunistic or no mammography screening is most common. Given the international difference in screening routines, at least two main trials are needed, ideally more. A European trial should include mammography screening units from several countries to account for differences in screening routines and differences in screening modalities and ethnicities of the women attending screening. In Europe, the breast cancer detection rate is approximately 0.5% per 1,000 women screened and the recall rate differs between approximately 3-7%. The screening modality that is used is mainly digital mammography. The screening interval is between one and three years and the screening age is most commonly 50-69, but some countries start screening at 40 and ends at 74 years. The screening attendance is on average approximately seventy percent.

The trial should assess the risk of breast cancer in women attending mammography screening in two screening rounds. Two arms will be included, one for risk assessment and one for the regular screening routine group. The outcome should be interval and aggressive cancers diagnosed before or at the last screening round. The comparison is done between the two arms.

A potential result from the trial is that high-risk women are identified who a) benefit from a follow-up using a more sensitive modality if they have dense breasts and b) benefit from a

follow-up using more frequent mammography screening if they have non-dense breasts. Fewer interval cancers and fewer aggressive cancers should be seen in the risk arm compared to the regular arm. More cancers are expected to be seen at the screening visits.

A low-risk group could also be identified who may not benefit from screening and potentially could have less frequent screening. The best definition of low risk could be reached by including genetic determinants into the image-based risk model. This opens a possibility to assess the breast cancer risk already at age 40 using the full risk model. Women with low risk could potentially be recommended to start their regular screening at age 50. The clinical experience for the radiologists is important to follow-up through interviews to understand any potential for an implementation phase.

In the US setting there are key differences in the diversity of ethnicities at screening units. It is more common to use one-year screening intervals and tomosynthesis modalities are a more commonly used. The screening attendance is low below 50%, and there is much higher recall rates of 10-30%. Each of these differences potentially affects the risk model and the model needs to be adapted to these screening settings.

As a second follow-up study, I suggest a clinical prospective trial for evaluating the low-dose therapy use in prevention and screening. The low-dose tamoxifen effect on breast cancer prevention and screening sensitivity could be assessed in the same study. Given that the low-dose phase II trial for reducing mammographic density already is done, a phase III trial should be performed where a reduction of breast cancer incidence is tested. The main aim is to reduce breast cancer incidence in the healthy population. Approximately 25,000 premenopausal women are required to perform the study in a 4-year and two screening round randomized trial. Women receive medication for the first two years and are observed for the next two years. With an incidence rate of ~0.3% approximately 300 breast cancer cases will be developed in the cohort.

With regards to the large study size, it should be investigated if women should be selected based on risk using the risk model from study II. The model is designed to identify approximately ten percent of all screened women where approximately eight times more cancers will be identified. This means that the group of 25,000 women could be reduced to approximately 3,000 women with a similar statistical power.

The main outcomes are a) cancer sub type by oestrogen-receptor status because tamoxifen is shown to reduce oestrogen-receptor positive cancers, and b) cancer detection rates and recall rates because the potential effect that tamoxifen has on reducing mammographic density. Additional outcomes are mammographic density change, adherence, and side-effects. The main estimates for point a) are oestrogen-receptor positive cancers compared between the arms. The main estimates for point b) are incidence of cancer of any subtype compared between the arms. In addition, cancer detection rates, recall rates, and screening sensitivity at year two and four are estimated. Mammographic density change, adherence, and side-effects are a key factor in both analyses.

A potential result from the trial is that fewer oestrogen-receptor women will be observed in the low-dose tamoxifen arms. It remains to be shown whether there will be a difference in number of oestrogen-positive cancers comparing all low-dose tamoxifen arms. It also remains to be shown if the potential reduction is seen at year 2 and at year 4. A second potential result from the trial is that the detection rate will increase for cancers of any subtype, even after accounting for the potential reduction of oestrogen-receptor positive cancers. Screening sensitivity may increase, although it is less clear how recall rates would be affected. How adherence and side-effects will develop is less easy to judge but they are key factors for interpreting the results. The clinical experience for the radiologists is important to follow-up through interviews as a first step towards understanding any potential for an implementation phase. Ethical discussions will be performed with study participants in focus groups on the concept of medicating non-cancerous women.

In summary, a clinical prospective trial could evaluate the risk assessment tool in a European and a US setting to assess the efficacy of reducing interval cancers and large breast cancers. The study will also assess any change in sensitivity, specificity, and recall using the model. In Europe current recall rates are at low levels compared with the high recall rates in US. A second clinical trial could evaluate the low-dose tamoxifen therapy efficacy of reducing oestrogen-receptor positive cancers and increasing screening sensitivity of any type of breast cancer by administering low-dose tamoxifen to healthy women.

The industry is carefully following late developments in the research fields of improved screening and prevention. New screening modalities such as contrast-enhanced mammography should be investigated as an affordable alternative to magnetic resonance imaging for the use in screening follow-up of women at high risk of breast cancer. Additional studies are needed to evaluate the risk model use in the risk-based screening setting in developing countries. Developing countries are likely to benefit the most from risk-based screening with life-changing consequences of reducing breast cancer mortality. New breast cancer risk reducing medication should be investigated such as endoxifen for its potential to reduce breast cancer incidence similar to tamoxifen but with less side-effects. Through the combined efforts of research and the industry, it is feasible to say that in 2030 we will see a break in the trend of increased breast cancer incidence by the prevention initiative and a further reduction of breast cancer mortality by the improved screening.

ABSTRACT IN SWEDISH

En av åtta kvinnor utvecklar bröstcancer under sin livstid i västvärlden. Åtgärder behövs därför för att minska dödligheten och för att förhindra bröstcancer. Mammografiscreening minskar dödligheten genom tidig upptäckt. Cirka en fjärdedel av kvinnorna som utvecklar bröstcancer diagnostiseras dock inom två år efter en normal mammografiundersöknng. Det finns därför ett behov av att identifiera den kortsiktiga risken för bröstcancer för att bättre kunna vägleda ett klinisk beslut för vilka kvinnor som behöver bättre uppföljning. En annan nackdel med mammografiskreening är att den enbart fokuserar på tidig upptäckt och inte på förebyggande åtgärder för bröstcancer. Idag är det känt att kvinnor som deltar i screening kan delas in i kvinnor med hög och låg risk för bröstcancer. Kvinnor med hög risk skulle kunna erbjudas förebyggande åtgärder såsom en låg dos av tamoxifen för att minska risken för att utveckla bröstcancer. Kvinnor med låg risk har inte samma behov av skreening och skulle därför kunna erbjudas mindre frekvent skreening.

I studie I utvecklade jag mammografitäthets-mätverktyget STRATUS som är till för att möjliggöra att mammogramresurser på sjukhus kan användas för storskaliga epidemiologiska studier för att studera risk och prognos för bröstcancer. STRATUS är även till för att studera risken för att missa en cancer samt att använda mammografisk täthet som en markör för att påvisa om man svarar på en behandling som kan förebygga bröstcancer.

I studie II utvecklade jag en riskmodell för att bedöma korttidsrisken för att bli diagnostiserad med bröstcancer. Modellen var baserad på mammografiska markörer såsom mammografisk täthet, mikroförkalkningar, knölar i bröstet och skillnader i mammografiska fynd mellan vänster och höger bröst. Modellen kan därtill utvidgas med riskfaktorer relaterat till livsstil och genetiska riskmarkörer. Baserat på resultaten visade vi att bland kvinnor som ej hade en cancer vid den nuvarande mammografiscreen, så identifierade korttidsrisk-verktyget kvinnor som senare blev diagnosiserade med bröstcancer före eller vid nästa screening. Vi visade också att traditionella riskmodeller inte var anpassade för att identifiera de kvinnor som inom en kort tid efter en riskbedömning blev diagnostiserade med bröstcancer.

I studie III genomförde vi en fas II-studie där den lägsta dosen tamoxifen identifierades som kunde användas för att minska den mammografiska tätheten. En sänkning av den mammografisk tätheten är en tidig markör för att kvinnan har en minskad risk för att utveckla bröstcancer. Kvinnorna påvisade lägre grad av biverkningar än standarddosen 20 mg som används normalt. En tamoxifendos på 2,5 mg tamoxifen var tillräcklig för att minska den mammografiska tätheten lika mycket som standarddosen. Till följd av den lägre dosen rapporterade kvinnorna cirka 50% mindre allvarliga vasomotoriska biverkningar, dvs. mindre vallningar och svettningar.

I studie IV undersökte jag om lågdos-tamoxifen även kan användas för att underlätta läsningen av mammografibilderna till följd av att tamoxifen sänker den mammografisk tätheten som kan dölja en cancer i mammografibilden. Vi visade att 24% av intervallcancer potentiellt skulle kunna upptäckas vid det tidigare ordinarie skreeningtillfället.

Sammanfattningsvis utvecklade vi verktyg för att mäta mammografisk täthet och korttidsrisk för bröstcancer. Dessutom utvecklade vi två koncept med lågdostamoxifen, där det första konceptet kan förhindra östrogenpositiva bröstcancerar från att utvecklas och det andra konceptet kan förbättre möjligheten för att upptäcka en cancer vid mammografiskreeningen. Kliniska prospektiva valideringar behöver genomföras för att utvärdera verktyget för riskbedömning samt en låg dos av tamoxifen mer ingående för att de ska kunna användas vid förebyggande av bröstcancer och för att minska dödlighet i bröstcancer.

11 ACKNOWLEDGEMENTS

I want to direct my sincere gratitude to all who contributed to my thesis.

Per Hall, my main supervisor. I am honored by the extraordinary introduction I received to the world of breast cancer epidemiology and top researchers around the world. By allowing me to work on "agnostic" research ideas I gained invaluable learnings. An editorial elegance and a focus on simple research questions accompanied with definite clinical end points is an experience to remember for life.

Kamila Czene, my co-supervisor. It is a once in a lifetime event to cross roads with such a brilliant researcher with the unique ability to read the scientific language between the lines. A character for refining scientific dares into research findings in black and white made my journey more successful than otherwise ever possible.

Doug Easton, world leading geneticist and head of BCAC with both feet on the ground who piloted me deep into the genetic risks and strongly supported me in my risk model approach.

Antonis Antoniou, world leading risk modeller who provided invaluable insights to risk constructs and the means to heal my sun burns.

Jack Cuzick, father of the most famous risk model who gave me profound understanding of risk factor assessment and automated tool constructs.

Marjanka Schmidt, leading cancer sub type researcher and head of B-CAST who didn't cast any shadow of the values for using mammographic density and image-based risk assessment to improve risk assessment in screening, but rather promoted this to the front line.

Montse Santori, leading risk epidemiologists who travelled all the way to my workplace to learn about the research platform and promoted it as the raw model at NCI, and in addition honorably refined it further.

Emily Conant, a lovely leading US radiologist at U-Penn with a natural mentoring image and a formidable interest in exploring the risk model in the research lab and clinical praxis.

Paul Pharoah, world leading prognosis modeller who honarably showed how risk transform suggestions make a difference for model fit.

Nico Karssemeijer, world leading cancer image analysist for very encouraging discussions of the need of a unique path for image-based risk models.

Celine Vachon, leading breast cancer researcher who encouraged me in the development of STRATUS for the use in research and in the clinic.

Carla van Gils, leading epidemiologist in studying MRI in a screening setting and promoter of the STRATUS density tool evaluation in BRCA patients. Thanks for the soothing discussions at conferences.

Soo Theo, leading epidemiologist bridging Asian and European populations differences into insights of the similarities who performed the first STRATUS based GWAS with very promising results.

Jennifer Stone, leading Australian epidemiologist with no challenge untouched and a great supporter of my risk model efforts.

John Hopper, leading Australian geneticist and a fantastic supporter of my research projects. Thanks for inspiring discussions at conferences.

Jenny Chang-Claude, key BCAC professor with whom I had many discussions about the best large-scale data approach for meta analyses.

Martin Widschwendter, head of ENVISION and invaluably promoter of the idea for the short-term risk model approach to improve screening.

JoAnn Pushkin, patient advocate and founder of DenseBreast-info.org for warm support of the risk model and for the energizing discussions.

Adam Brentnall, key developer of the Tyrer-Cuzick model for interesting meetings at Wolfson and insightful information exchanges.

Anthony Howell, charming professor and risk prevention specialist who provided highly interesting feedback in research collaboration on risk prevention.

Susan Astley, expert mammographic density researcher with whom I had many cheerful conference chats and dinners.

Gretchen Gierach, expert breast density researcher who gave insightful discussion points at conferences and inspiring events in the Hawaiian Islands.

Karin Leifland, head of the South General hospital radiology unit who invited me to stimulating discussions on how the screening unit work flow is setup and for showing interest in the potential for improvements.

Paolo Giorgi Rossi, brilliant Italian epidemiologist and mammography screening program developer for discussions of remarkable research project prospects.

Axel Gräwingholt, German radiologist with unmistaken devotion to improve mammography screening and an invaluable fascination to evaluate "the latest" in the field.

Nasim Mavaddat, expert geneticist risk modeller who opened my eyes for the great interest in combining late genetic risk determinants with mammographic density as discussed at nice dinners.

Manjeet Bolla, main data manager in BCAC with a very nice appearance and nice lunch companionship for discussions of the data complexities in vast projects.

Jean Wang, key data manager in BCAC with great patience in pursuing loose ends in the data yarn to facilitate research prospects even further.

Michael Lush, IT responsible in Cambridge with a devoted support for guaranteeing the uptime of the STRATUS density tool and for cheerful chats at the coffee machine.

Joe Dennis, geneticist in Cambridge with a formidable programming skill and interest in the further development of density. Nice party conversations.

John Shepherd, leading image analysist and conference organizer for our breakthrough discussions of future development paths in image analysis on our many fold Thursday pancake lunches.

Signe Borgquist, oncologist and epidemiologist for insightful dinner discussions and research discussions, always with a sparkling touch.

Fiona Gilbert, leading UK radiologist for vivid conference discussions of the landscape of on-going studies and the possibilities for the better positioning.

Sophia Zackrisson, leading Swedish radiologist with very productive flight discussion and following fruitful research collaborations.

Fredrik Strand, radiologist and prior PhD colleague who offered very valuable support for performing a key paper in my thesis.

Steven Quay, MIT lecturer and head of the Persian queen therapeutics company who shared enormous skills in patent work with unprecedented dedication to prevention therapies.

Ralph Highnam, head of Volpara and legendary density tool developer with whom I spent many white-board sessions to push the density development one more step forward.

Douglas Katz, scientific editor with skills unprecedented to my experience who gave me memorable learnings towards perfecting a manuscript.

Linda Moy, US radiologist and screening regulatory committee member who gave me crucial screening regulatory insights and great panel discussions for further risk model direction.

Andrea DeCensi, legendary prevention researcher and clinician who inspired me in critical analysis of tamoxifen doses in balancing effect, side-effects, and prior knowledge in the field in a great panel discussion.

Jingmei Li, associate professor in Singapore who quickly connects the dots to a full picture with flowering ideas and great follow-up research collaborations.

Wei He, prior post doc in our research group with a very pleasant approach and sharing of nice research ideas, along with Chinese habits.

Hatef Darabi, prior post doc in our group with prediction skills to infinity and beyond who introduced me to the martial arts.

Natalie Holowko, prior post doc in our research group with especial interest in genetic determinants of mammographic features for the productive research collaborations.

Martin Eklund, associate professor at the department with an exquisite ability to stay on course in delicate analyses under challenging conditions.

Andrea Discacciati, a post doc at the department with excellent statistical skills who elegantly and cheerfully showed how all imputations and bootstraps should be realized.

Felix Grossmann, post doc in our research group with extraordinary large-scale analysis abilities and insightful discussions of analysis problems over oatly spiced coffee.

Shadi Azam, PhD fellow for all discussions with easy laughs and several collaborative research projects.

Maya Alsheh Ali, post doc in our research group with high image analysis skills with special interest in automated microcalcification analysis and fruitful collaborations.

Haomin Yang, prior PhD fellow for fun discussions on complementary use cases of mammographic features with s novel clinical potential.

Roxanna Hellgren, radiologist and head of radiology at South General hospital for nice walks and talks about potential research questions and their challenges.

Amber Wilcox, US PhD student who kept me on edge with multitudes of risk model assessments.

Nadja Rajaram, Malaysian PhD student who clarified the similarities and differences of mammographic densities in a fruitful collaboration using STRATUS in international cohorts with a focus on traditional risk factors and physical activity. Soya is the latest prospect.

Linda Rainey, PhD fellow in the Netherlands for nice discussions on analysis challenges and quality of life interpretations.

Marie O'Reilly, professor at the department and an excellent teacher who extended my view of what is a good study design.

Matteo Bottai, professor at IMM and a distinguished teacher who made teaching of elements a work of art.

Abbas Cheddad, prior post doc in the research group and skilled image analysist for interesting meetings on analysis challenges during summers.

Senthil Periaswamy, industry leading deep learning expert with ROC curves and patent skills covering the office walls like a true rock star. Thanks for fruitful collaborations and coffee coupons.

Kourosh Jafari, deep learning expert with unusual empathic image and modelling expertise with unravelled performances and future prospects. Thanks for nice chats.

Mike Klein, head of the world leading cancer detection tool for mammography screening that warmly embraced the expansion of the portfolio into my image based short-term risk model.

Rodney Hawkins, vice president of industry leading detection tool and the first to adopt the short-term risk tool for evaluation and later implementation into a clinically valid product.

Andjela Azabagic, project manager for the short-term risk tool and is devoted 24 hours a day pushing towards to establish the first short-term risk tool for improving womens' breast health in the screening setting.

Jeffrey Hoffmeister, scientist and research study responsible for clinical implementations who has shown extraordinary engagement in promoting the risk tool.

Alicia Garfield, project manager in the risk tool project who luckily keep us all on the right track with time-lines and deliveries.

Jim Pike, research responsible for the clinical risk tool and devoted into adopting the tool to the different and very challenging screening settings.

Rives Bird, computer aided detection account responsible who provided the first machine that later led to history.

Simon Lee, IT specialist who with unprecedented record in delivering solutions to problems that occur in the clinical application phase.

Daniel Hayden, programming expert with a probable Viking gene hidden and language skills that makes any coding translation a breeze.

Thomas Harrison, programming expert with skills stretching also into the deep learning area and strong tool prototypes.

Linda Thorén, PhD fellow with sparkling bottle promises soon to be realized after extensive and nice research discussions on improving breast cancer patient health using refined early markers for therapy response.

Jacob Järås, statistician and breast cancer register expert for an always happy and reliable delivery and quality strategy.

Daniel Nyqvist, Pfizer representative with great interest in the research prospects and valuable discussions on future endeavors.

Malin Gavelin, editor of Hemmets Journal for a fruitful discussion and introduction to media exposure.

And not the least, I want to direct my sincere gratitude to each woman that contributed with trust, time, information, and samples to the research work. **And in addition**, to the Entire study personnel group in KARMA, KARISMA, LIBRO1, SASBAC, MBTST, and CSAW.

12 REFERENCES

1. *Statistics on Cancer Incidence 2016. The National Board of Health and Welfare, 2017, publikationer/2017-12-31*. 2017.
2. Plevritis, S.K., et al., *Association of Screening and Treatment With Breast Cancer Mortality by Molecular Subtype in US Women, 2000-2012*. JAMA, 2018. **319**(2): p. 154-164.
3. Olsson, S., et al., *Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme*. J Med Screen, 2000. **7**(1): p. 14-8.
4. Welch, H.G., et al., *Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness*. N Engl J Med, 2016. **375**(15): p. 1438-1447.
5. *Årsrapport 2015 från Nationella Bröstcancerregistret; Regionala Cancercentrum i samverkan, 2016*.
6. Mocellin, S., A. Goodwin, and S. Pasquali, *Risk-reducing medications for primary breast cancer: a network meta-analysis*. Cochrane Database Syst Rev, 2019. **4**: p. CD012191.
7. Land, S.R., et al., *Symptoms and QOL as predictors of chemoprevention adherence in NRG Oncology/NSABP Trial P-1*. Journal of the National Cancer Institute, 2016. **1;108(4):djv365**.
8. Martini, F.H., M.J. Timmons, and R.B. Tallitsch, *Human anatomy 6th ed*. San Francisco: Pearson Benjamin Cummings, 2009.
9. Javed, A. and A. Lteif, *Development of the human breast*. Semin Plast Surg, 2013. **27**(1): p. 5-12.
10. Russo, I.H. and J. Russo, *Pregnancy-induced changes in breast cancer risk*. J Mammary Gland Biol Neoplasia, 2011. **16**(3): p. 221-33.
11. Radisky, D.C. and L.C. Hartmann, *Mammary involution and breast cancer risk: transgenic models and clinical studies*. J Mammary Gland Biol Neoplasia, 2009. **14**(2): p. 181-91.
12. Zhang, J. and S.N. Powell, *The role of the BRCA1 tumor suppressor in DNA double-strand break repair*. Mol Cancer Res, 2005. **3**(10): p. 531-9.
13. Clamp, A., S. Danson, and M. Clemons, *Hormonal and genetic risk factors for breast cancer*. Surgeon, 2003. **1**(1): p. 23-31.
14. Michailidou, K., et al., *Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer*. Nat Genet, 2015. **47**(4): p. 373-80.
15. Michailidou, K., et al., *Association analysis identifies 65 new breast cancer risk loci*. Nature, 2017. **551**(7678): p. 92-94.
16. Holm, J., et al., *Associations of Breast Cancer Risk Prediction Tools With Tumor Characteristics and Metastasis*. J Clin Oncol, 2016. **34**(3): p. 251-8.
17. Mavaddat, N., et al., *Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes*. Am J Hum Genet, 2019. **104**(1): p. 21-34.
18. Stephens, P.J., et al., *The landscape of cancer genes and mutational processes in breast cancer*. Nature, 2012. **486**(7403): p. 400-4.
19. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.
20. Wellings, S.R., H.M. Jensen, and R.G. Marcum, *An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions*. J Natl Cancer Inst, 1975. **55**(2): p. 231-73.

21. Robinson, D., L. Holmberg, and H. Moller, *The occurrence of invasive cancers following a diagnosis of breast carcinoma in situ.* Br J Cancer, 2008. **99**(4): p. 611-5.

22. Innos, K. and P.L. Horn-Ross, *Risk of second primary breast cancers among women with ductal carcinoma in situ of the breast.* Breast cancer research and treatment, 2008. **1;111(3):531-40**.

23. De Souza, I., et al., *Growth-hormone dependence among human breast cancers.* The Lancet, 1974. **27;304(7874):182-4**.

24. *http://pathology.jhu.edu/breast/types.php, Accessed Nov 2020.*

25. Lichtenstein, P., et al., *Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland.* N Engl J Med, 2000. **343**(2): p. 78-85.

26. Buell, P., *Changing incidence of breast cancer in Japanese-American women.* J Natl Cancer Inst, 1973. **51**(5): p. 1479-83.

27. Clemons, M. and P. Goss, *Estrogen and the risk of breast cancer.* N Engl J Med, 2001. **344**(4): p. 276-85.

28. Dupont, W.D. and D.L. Page, *Risk factors for breast cancer in women with proliferative breast disease.* N Engl J Med, 1985. **312**(3): p. 146-51.

29. Li, J., et al., *Family History, Reproductive, and Lifestyle Risk Factors for Fibroadenoma and Breast Cancer.* JNCI Cancer Spectr, 2018. **2**(3): p. pky051.

30. Schneider, R., *Comparison of age, sex, and incidence rates in human and canine breast cancer.* Cancer, 1970. **26**(2): p. 419-26.

31. Howlader, N., et al., *SEER Cancer Statistics Review, 1975-2017, National Cancer Institute, https://seer.cancer.gov/csr/1975_2017/, Accessed Nov 2020.*

32. Lilienfeld, A.M. and E.A. Johnson, *The age distribution in female breast and genital cancers.* Cancer, 1955. **8**(5): p. 875-82.

33. van den Brandt, P.A., et al., *Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk.* Am J Epidemiol, 2000. **152**(6): p. 514-27.

34. Cleary, M.P. and N.J. Maihle, *The role of body mass index in the relative risk of developing premenopausal versus postmenopausal breast cancer.* Proc Soc Exp Biol Med, 1997. **216**(1): p. 28-43.

35. Ahlgren, M., et al., *Growth patterns and the risk of breast cancer in women.* N Engl J Med, 2004. **351**(16): p. 1619-26.

36. Kelsey, J.L., M.D. Gammon, and E.M. John, *Reproductive factors and breast cancer.* Epidemiol Rev, 1993. **15**(1): p. 36-47.

37. Collaborative Group on Hormonal Factors in Breast, C., *Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies.* Lancet Oncol, 2012. **13**(11): p. 1141-51.

38. Bodicoat, D.H., et al., *Timing of pubertal stages and breast cancer risk: the Breakthrough Generations Study.* Breast Cancer Res, 2014. **16**(1): p. R18.

39. Friebel, T.M., S.M. Domchek, and T.R. Rebbeck, *Modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: systematic review and meta-analysis.* J Natl Cancer Inst, 2014. **106**(6): p. dju091.

40. Kahlenborn, C., et al., *Oral contraceptive use as a risk factor for premenopausal breast cancer: a meta-analysis.* InMayo Clinic Proceedings, Elsevier, 2006. **Vol. 81, No. 10, pp. 1290-1302**.

41. Lambertini, M., et al., *Reproductive behaviors and risk of developing breast cancer according to tumor subtype: A systematic review and meta-analysis of epidemiological studies.* Cancer Treat Rev, 2016. **49**: p. 65-76.

42. Pan, H., et al., *Reproductive factors and breast cancer risk among BRCA1 or BRCA2 mutation carriers: results from ten studies.* Cancer epidemiology. **1;38(1):1-8**.

43. Chowdhury, R., et al., *Breastfeeding and maternal health outcomes: a systematic review and meta-analysis.* Acta Paediatr, 2015. **104**(467): p. 96-113.

44. Narod, S.A., *Hormone replacement therapy and the risk of breast cancer.* Nat Rev Clin Oncol, 2011. **8**(11): p. 669-76.

45. Beral, V. and C. Million Women Study, *Breast cancer and hormone-replacement therapy in the Million Women Study.* Lancet, 2003. **362**(9382): p. 419-27.

46. Tempfer, C.B., et al., *Side effects of phytoestrogens: a meta-analysis of randomized trials.* Am J Med, 2009. **122**(10): p. 939-46 e9.

47. Takahashi, T.A. and K.M. Johnson, *Menopause.* Med Clin North Am, 2015. **99**(3): p. 521-34.

48. Sievert, L.L., *Menopause: a biocultural perspective*. 2006: Rutgers University Press.

49. Press, D.J., et al., *Breast cancer risk and ovariectomy, hysterectomy, and tubal sterilization in the women's contraceptive and reproductive experiences study.* Am J Epidemiol, 2011. **173**(1): p. 38-47.

50. Longnecker, M.P., *Alcoholic beverage consumption in relation to risk of breast cancer: meta-analysis and review*. Cancer Causes Control, 1994. **5**(1): p. 73-82.

51. Smith-Warner, S.A., et al., *Alcohol and breast cancer in women: a pooled analysis of cohort studies.* JAMA, 1998. **279**(7): p. 535-40.

52. Jung, S., et al., *Alcohol consumption and breast cancer risk by estrogen receptor status: in a pooled analysis of 20 studies.* Int J Epidemiol, 2016. **45**(3): p. 916-28.

53. Trinh, T., et al., *Background risk of breast cancer influences the association between alcohol consumption and mammographic density.* British journal of cancer, 2015. **113(1):159-65.**

54. Johnson, K.C., et al., *Active smoking and secondhand smoke increase breast cancer risk: the report of the Canadian Expert Panel on Tobacco Smoke and Breast Cancer Risk (2009).* Tob Control, 2011. **20**(1): p. e2.

55. Morabia, A., et al., *Relation of smoking to breast cancer by estrogen receptor status.* Int J Cancer, 1998. **75**(3): p. 339-42.

56. Van Voorhis, B.J., et al., *The effects of smoking on ovarian function and fertility during assisted reproduction cycles.* Obstet Gynecol, 1996. **88**(5): p. 785-91.

57. Michnovicz, J.J., et al., *Increased 2-hydroxylation of estradiol as a possible mechanism for the anti-estrogenic effect of cigarette smoking.* New England Journal of Medicine, 1986. **20;315(21):1305-9**.

58. Audrain-McGovern, J. and N.L. Benowitz, *Cigarette smoking, nicotine, and body weight.* Clin Pharmacol Ther, 2011. **90**(1): p. 164-8.

59. Vachon, C.M., et al., *Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States).* Cancer Causes Control, 2000. **11**(7): p. 653-62.

60. Caspersen, C.J., K.E. Powell, and G.M. Christenson, *Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research.* Public Health Rep, 1985. **100**(2): p. 126-31.

61. Ainsworth, B.E., et al., *Compendium of physical activities: an update of activity codes and MET intensities.* Med Sci Sports Exerc, 2000. **32**(9 Suppl): p. S498-504.

62. Wu, Y., D. Zhang, and S. Kang, *Physical activity and risk of breast cancer: a meta-analysis of prospective studies.* Breast Cancer Res Treat, 2013. **137**(3): p. 869-82.

63. Irwin, M.L., et al., *Physical activity, body mass index, and mammographic density in postmenopausal breast cancer survivors.* J Clin Oncol, 2007. **25**(9): p. 1061-6.

64. Qureshi, S.A., et al., *Physical activity and mammographic density in a cohort of postmenopausal Norwegian women; a cross-sectional study.* Springerplus, 2012. **1**(1): p. 75.

65. Boyle, P., *Triple-negative breast cancer: epidemiological considerations and recommendations.* Ann Oncol, 2012. **23 Suppl 6**: p. vi7-12.

66. Mavaddat, N., et al., *Prediction of breast cancer risk based on profiling with common genetic variants.* J Natl Cancer Inst, 2015. **107**(5).

67. Antoniou, A.C., et al., *The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions.* Br J Cancer, 2008. **98**(8): p. 1457-66.

68. Tomasetti, C., L. Li, and B. Vogelstein, *Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention.* Science, 2017. **355**(6331): p. 1330-1334.

69. Li, J., et al., *Prevalence of BRCA1 and BRCA2 pathogenic variants in a large, unselected breast cancer cohort.* Int J Cancer, 2019. **144**(5): p. 1195-1204.

70. Antoniou, A., et al., *Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies.* Am J Hum Genet, 2003. **72**(5): p. 1117-30.

71. Sobin, L.H. and M.K.W. Gospodarowicz, C., *TNM classification of malignant tumours.* 2011: John Wiley & Sons.

72. Elston, C.W. and I.O. Ellis, *Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.* Histopathology, 2002. **41**(3A): p. 154-61.

73. Ahn, S., et al., *HER2 status in breast cancer: changes in guidelines and complicating factors for interpretation.* J Pathol Transl Med, 2020. **54**(1): p. 34-44.

74. Grabau, D., *KVAST dokument bröstumörer.* Vol. volume 3.1.3. 2014: Svensk förening för patologi - Svensk förening för klinisk cytologi.

75. Vetto, J., et al., *Use of the "triple test" for palpable breast lesions yields high diagnostic accuracy and cost savings.* The American journal of surgery, 1995. **1;169(5):519-22**.

76. Radiology, A.C.o., *Breast Imaging Reporting and Data System® (BI-RADS®) 5*. 2013.

77. Obenauer, S., K.P. Hermann, and E. Grabbe, *Applications and literature review of the BI-RADS classification.* Eur Radiol, 2005. **15**(5): p. 1027-36.

78. Ballard-Barbash, R., et al., *Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database.* AJR Am J Roentgenol, 1997. **169**(4): p. 1001-8.

79. Zheng, Y., et al., *Parenchymal texture analysis in digital mammography: A fully automated pipeline for breast cancer risk assessment.* Med Phys, 2015. **42**(7): p. 4149-60.

80. Winkel, R.R., 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. **1;16(1):414**.

81. Malkov, S., et al., *Mammographic texture and risk of breast cancer by tumor type and estrogen receptor status*. Breast Cancer Res, 2016. **18**(1): p. 122.

82. Martinez, V. and J.G. Azzopardi, *Invasive lobular carcinoma of the breast: incidence and variants*. Histopathology, 1979. **3(6):467-88**.

83. Saadatmand, S., et al., *Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients*. Bmj, 2015. **6;351:h4901**.

84. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.

85. Cheang, M.C., et al., *Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer*. J Natl Cancer Inst, 2009. **101**(10): p. 736-50.

86. *Cancer Facts & Figures 2020, National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program, Accessed Nov 2020*.

87. Kamangar, F., G.M. Dores, and W.F. Anderson, *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): p. 2137-50.

88. Nason, L.K., *Breast Imaging: The Requisites*. Academic Radiology, 2011. **2011;9(18):1201-2.**

89. Li, T., et al., *The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer*. Cancer Epidemiol Biomarkers Prev, 2005. **14**(2): p. 343-9.

90. Gabrielson, M., et al., *Amount of stroma is associated with mammographic density and stromal expression of oestrogen receptor in normal breast tissues*. Breast cancer research and treatment, 2016. **1;158(2):253-61**.

91. Conklin, M.W. and P.J. Keely, *Why the stroma matters in breast cancer: insights into breast cancer patient outcomes through the examination of stromal biomarkers*. Cell adhesion & migration, 2012. **1;6(3):249-60**.

92. Wolfe, J.N., *Risk for breast cancer development determined by mammographic parenchymal pattern*. Cancer, 1976. **37**(5): p. 2486-92.

93. Tabar, L. and P.B. Dean, *Mammographic parenchymal patterns. Risk indicator for breast cancer?* JAMA, 1982. **247**(2): p. 185-9.

94. Boyd, N.F., et al., *Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study*. J Natl Cancer Inst, 1995. **87**(9): p. 670-5.

95. Burnside, E.S., et al., *The ACR BI-RADS experience: learning from history*. J Am Coll Radiol, 2009. **6**(12): p. 851-60.

96. Li, J., et al., *High-throughput mammographic-density measurement: a tool for risk prediction of breast cancer*. Breast Cancer Res, 2012. **14**(4): p. R114.

97. Highnam, R.P., J.M. Brady, and B.J. Shepstone, *Estimation of compressed breast thickness during mammography*. Br J Radiol, 1998. **71**(846): p. 646-53.

98. Shepherd, J.A., et al., *Volume of mammographic density and risk of breast cancer*. Cancer Epidemiol Biomarkers Prev, 2011. **20**(7): p. 1473-82.

99. Astley, S.M., et al., *A comparison of five methods of measuring mammographic density: a case-control study.* Breast Cancer Res, 2018. **20**(1): p. 10.

100. Gastounioti, A., et al., *Evaluation of LIBRA Software for Fully Automated Mammographic Density Assessment in Breast Cancer Risk Prediction.* Radiology, 2020. **296**(1): p. 24-31.

101. Keller, B.M., et al., *Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation.* Medical physics, 2012. **39(8):4903-17**.

102. Boyd, N., et al., *A longitudinal study of the effects of menopause on mammographic features.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(10 Pt 1): p. 1048-53.

103. Huo, C.W., et al., *Mammographic density-a review on the current understanding of its association with breast cancer.* Breast Cancer Res Treat, 2014. **144**(3): p. 479-502.

104. Boyd, N.F., et al., *Mammographic density and the risk and detection of breast cancer.* N Engl J Med, 2007. **356**(3): p. 227-36.

105. Strand, F., et al., *Localized mammographic density is associated with interval cancer and large breast cancer: a nested case-control study.* Breast cancer research and treatment, 2019. **1;21(1):8**.

106. Fowler, P.A., et al., *Cyclic changes in composition and volume of the breast during the menstrual cycle, measured by magnetic resonance imaging.* Br J Obstet Gynaecol, 1990. **97**(7): p. 595-602.

107. Boyd, N.F., et al., *Mammographic densities and the prevalence and incidence of histological types of benign breast disease. Reference Pathologists of the Canadian National Breast Screening Study.* Eur J Cancer Prev, 2000. **9**(1): p. 15-24.

108. Azam, S., et al., *Determinants of Mammographic Density Change.* JNCI Cancer Spectr, 2019. **3**(1): p. pkz004.

109. Maskarinec, G., et al., *A longitudinal investigation of mammographic density: the multiethnic cohort.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(4): p. 732-9.

110. Vachon, C.M., et al., *Longitudinal trends in mammographic percent density and breast cancer risk.* Cancer Epidemiol Biomarkers Prev, 2007. **16**(5): p. 921-8.

111. van Gils, C.H., et al., *Changes in mammographic breast density and concomitant changes in breast cancer risk.* Eur J Cancer Prev, 1999. **8**(6): p. 509-15.

112. Cuzick, J., et al., *Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: a nested case-control study.* J Natl Cancer Inst, 2011. **103**(9): p. 744-52.

113. Cuzick, J., et al., *Tamoxifen for prevention of breast cancer: extended long-term follow-up of the IBIS-I breast cancer prevention trial.* Lancet Oncol, 2015. **16**(1): p. 67-75.

114. Li, J., et al., *Mammographic density reduction is a prognostic marker of response to adjuvant tamoxifen therapy in postmenopausal patients with breast cancer.* Journal of Clinical Oncology, 2013. **20;31(18):2249**.

115. Marquardt, D.W., *An algorithm for least-squares estimation of nonlinear parameters.* Journal of the society for Industrial and Applied Mathematics, 1963. **11(2):431-41**.

116. Nyante, S.J., et al., *The association between mammographic calcifications and breast cancer prognostic factors in a population-based registry cohort.* Cancer, 2017. **15;123(2):219-27**.

117. Gulsun, M., F.B. Demirkazik, and M. Ariyurek, *Evaluation of breast microcalcifications according to Breast Imaging Reporting and Data System criteria and Le Gal's classification.* Eur J Radiol, 2003. **47**(3): p. 227-31.

118. Aebi, S., et al., *Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2011. **22 Suppl 6**: p. vi12-24.

119. Li, C.I., et al., *Risk of invasive breast carcinoma among women diagnosed with ductal carcinoma in situ and lobular carcinoma in situ.* Cancer, 2006. **15;106(10):2104-12**.

120. Tabar L, T.C.H., Amy Yen MF, Tot T, Tung TH, Chen LS, Chiu YH, Duffy SW, Smith RA., *Mammographic tumor features can predict long-term outcomes reliably in women with 1–14-mm invasive breast carcinoma: Suggestions for the reconsideration of current therapeutic practice and the TNM classification system.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2004. **15;101(8):1745-59**.

121. Ferranti, C., et al., *Relationships between age, mammographic features and pathological tumour characteristics in non-palpable breast cancer.* The British journal of radiology, 2000. **73(871):698-705**.

122. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.

123. Creighton, C.J., J.C. Chang, and J.M. Rosen, *Epithelial-mesenchymal transition (EMT) in tumor-initiating cells and its clinical implications in breast cancer.* J Mammary Gland Biol Neoplasia, 2010. **15**(2): p. 253-60.

124. Scimeca, M., et al., *Microcalcifications in breast cancer: an active phenomenon mediated by epithelial cells with mesenchymal characteristics.* BMC Cancer, 2014. **14**: p. 286.

125. Tse, G.M., et al., *Calcification in breast lesions: pathologists' perspective.* J Clin Pathol, 2008. **61**(2): p. 145-51.

126. Cen, D., et al., *BI-RADS 3-5 microcalcifications: prediction of lymph node metastasis of breast cancer.* Oncotarget, 2017. **8**(18): p. 30190-30198.

127. London, S.J., et al., *A prospective study of benign breast disease and the risk of breast cancer.* JAMA, 1992. **267**(7): p. 941-4.

128. Hartmann, L.C., et al., *Benign breast disease and the risk of breast cancer.* N Engl J Med, 2005. **353**(3): p. 229-37.

129. Dupont, W.D., et al., *Long-term risk of breast cancer in women with fibroadenoma.* N Engl J Med, 1994. **331**(1): p. 10-5.

130. Mullooly, M., et al., *Investigation of relationships between breast cancer risk factors and bilateral mammographic breast density asymmetry among women undergoing diagnostic image-guided breast biopsies.* Cancer Research, 2020. **10.1158/1538-7445.AM2020-NG15**.

131. Hankinson, S.E., G.A. Colditz, and W.C. Willett, *Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones.* Breast Cancer Res, 2004. **6**(5): p. 213-8.

132. Rice, M.S., et al., *Does mammographic density mediate risk factor associations with breast cancer? An analysis by tumor characteristics.* Breast cancer research and treatment, 2018. **1;170(1):129-41**.

133. Byrne, C., et al., *Mammographic Density Change With Estrogen and Progestin Therapy and Breast Cancer Risk.* J Natl Cancer Inst, 2017. **109**(9).

134. Nevo, D., X. Liao, and D. Spiegelman, *Estimation and Inference for the Mediation Proportion.* Int J Biostat, 2017. **13**(2).

135. McCulloch, W.S. and W. Pitts, *A logical calculus of the ideas immanent in nervous activity. 1943.* Bull Math Biol, 1990. **52**(1-2): p. 99-115; discussion 73-97.

136. Kelley, H.J., *Gradient theory of optimal flight paths.* Ars Journal, 1960. **30(10):947-54**.

137. Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning*. 2017: Cambridge MA, MIT press.

138. Kooi, T., et al., *Large scale deep learning for computer aided detection of mammographic lesions.* Med Image Anal, 2017. **35**: p. 303-312.

139. Doll, R. and A.B. Hill, *Mortality in Relation to Smoking: Ten Years' Observations of British Doctors.* Br Med J, 1964. **1**(5396): p. 1460-7 CONCL.

140. Shmueli, G., *To explain or to predict?* Statistical science, 2010. **25(3):289-310**.

141. Yerushalmy, J., *Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques.* Public Health Rep, 1947. **62**(40): p. 1432-49.

142. Altman, D.G. and J.M. Bland, *Diagnostic tests. 1: Sensitivity and specificity.* British Medical Journal, 1994. **11;308(6943):1552**.

143. Hosmer, D.W., et al., *A comparison of goodness-of-fit tests for the logistic regression model.* Stat Med, 1997. **16**(9): p. 965-80.

144. *National Institute for Health and Care Excellence. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer CG164. NICE; 2013. https://www.nice.org.uk/guidance/cg164. Accessed November 2020.*

145. Siu, A.L., *Screening for breast cancer: US Preventive Services Task Force recommendation statement.* Annals of internal medicine, 2016. **16;164(4):279-96**.

146. Brentnall, A.R. and J. Cuzick, *Risk models for breast cancer and their validation.* Statistical science: a review journal of the Institute of Mathematical Statistics, 2020. **3;35(1):14**.

147. Gail, M.H., et al., *Projecting individualized probabilities of developing breast cancer for white females who are being examined annually.* Journal of the National Cancer Institute, 1989. **20;81(24):1879-86**.

148. Dignam, J.J., Q. Zhang, and M. Kocherginsky, *The use and interpretation of competing risks regression models.* Clinical Cancer Research, 2012. **15;18(8):2301-8**.

149. Fine, J.P. and R.J. Gray, *A proportional hazards model for the subdistribution of a competing risk.* Journal of the American statistical association, 1999. **1;94(446):496-509**.

150. Cintolo-Gonzalez, J.A., et al., *Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications.* Breast Cancer Res Treat, 2017. **164**(2): p. 263-284.

151. Tyrer, J., S.W. Duffy, and J. Cuzick, *A breast cancer prediction model incorporating familial and personal risk factors.* Stat Med, 2004. **23**(7): p. 1111-30.

152. van Veen, E.M., et al., *Use of single-nucleotide polymorphisms and mammographic density plus classic risk factors for breast cancer risk prediction.* JAMA oncology, 2018. **1;4(4):476-82**.

153. MacInnis, R.J., et al., *Comparing Five-Year and Lifetime Risks of Breast Cancer in the Prospective Family Study Cohort.* J Natl Cancer Inst, 2020.

154. Pashayan, N., et al., *Personalized early detection and prevention of breast cancer: ENVISION consensus statement.* Nat Rev Clin Oncol, 2020. **17**(11): p. 687-705.

155. Louro, J., et al., *A systematic review and quality assessment of individualised breast cancer risk prediction models.* Br J Cancer, 2019. **121**(1): p. 76-85.

156. Tabar, L., et al., *Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare.* Lancet, 1985. **1**(8433): p. 829-32.

157. Tabar, L., et al., *Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial.* Cancer, 1995. **75**(10): p. 2507-17.

158. Gotzsche, P.C. and K.J. Jorgensen, *Screening for breast cancer with mammography.* Cochrane Database Syst Rev, 2013(6): p. CD001877.

159. Independent, U.K.P.o.B.C.S., *The benefits and harms of breast cancer screening: an independent review.* Lancet, 2012. **380**(9855): p. 1778-86.

160. Lind, H., et al., *Breast Cancer Screening Program in Stockholm County, Sweden - Aspects of Organization and Quality Assurance.* Breast Care (Basel), 2010. **5**(5): p. 353-357.

161. Berg, W.A., 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): p. 1394-404.

162. Cozzi, A., S. Schiaffino, and F. Sardanelli, *The emerging role of contrast-enhanced mammography.* Quant Imaging Med Surg, 2019. **9**(12): p. 2012-2018.

163. *European breast cancer guidelines, https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines. Accessed Nov 2020.*

164. Klemi, P.J., et al., *Mammography screening interval and the frequency of interval cancers in a population-based screening.* Br J Cancer, 1997. **75**(5): p. 762-6.

165. Tupper, C., et al., *Breast screening interval and the characteristics of screen-detected cancers.* Breast cancer research and treatment, 2015. **17(1):1-3**.

166. Britt, K.L., J. Cuzick, and K.A. Phillips, *Key steps for effective breast cancer prevention.* Nat Rev Cancer, 2020. **20**(8): p. 417-436.

167. Liede, A., et al., *Risk-reducing mastectomy rates in the US: a closer examination of the Angelina Jolie effect.* Breast Cancer Res Treat, 2018. **171**(2): p. 435-442.

168. Cuzick, J., et al., *Overview of the main outcomes in breast-cancer prevention trials.* Lancet, 2003. **361**(9354): p. 296-300.

169. Cuzick, J., *Progress in preventive therapy for cancer: a reminiscence and personal viewpoint.* Br J Cancer, 2018. **118**(9): p. 1155-1161.

170. Houssami, N., J. Cuzick, and J.M. Dixon, *The prevention, detection, and management of breast cancer.* Med J Aust, 2006. **184**(5): p. 230-4.

171. He, W., et al., *Predictors of Discontinuation of Adjuvant Hormone Therapy in Patients With Breast Cancer.* J Clin Oncol, 2015. **33**(20): p. 2262-9.

172. Thurston, R.C., et al., *Menopausal symptoms and cardiovascular disease mortality in the Women's Ischemia Syndrome Evaluation (WISE).* Menopause, 2017. **24**(2): p. 126-132.

173. DeCensi, A., et al., *Randomized Placebo Controlled Trial of Low-Dose Tamoxifen to Prevent Local and Contralateral Recurrence in Breast Intraepithelial Neoplasia.* J Clin Oncol, 2019. **37**(19): p. 1629-1637.

174. Day, R., et al., *Health-related quality of life and tamoxifen in breast cancer prevention: a report from the National Surgical Adjuvant Breast and Bowel Project P-1 Study.* J Clin Oncol, 1999. **17**(9): p. 2659-69.

175. He, W., et al., *CYP2D6 Genotype Predicts Tamoxifen Discontinuation and Prognosis in Patients With Breast Cancer.* J Clin Oncol, 2020. **38**(6): p. 548-557.

176. Powles, T.J., et al., *Twenty-year follow-up of the Royal Marsden randomized, double-blinded tamoxifen breast cancer prevention trial.* J Natl Cancer Inst, 2007. **99**(4): p. 283-90.

177. Kim, W.H., et al., *Mammographic density changes following discontinuation of tamoxifen in premenopausal women with oestrogen receptor-positive breast cancer.* Eur Radiol, 2018. **28**(8): p. 3176-3184.

178. Sauter, E.R., *Breast Cancer Prevention: Current Approaches and Future Directions.* Eur J Breast Health, 2018. **14**(2): p. 64-71.

179. Gabrielson, M., et al., *Cohort Profile: The Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA).* Int J Epidemiol, 2017. **46**(6): p. 1740-1741g.

180. Zackrisson, S., et al., *One-view breast tomosynthesis versus two-view mammography in the Malmo Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study.* Lancet Oncol, 2018. **19**(11): p. 1493-1503.

181. KARMA. *https://karmastudy.org, Accessed Nov 2020*.

182. *https://karmastudy.org/contact/data-access/data-sources/, Accessed Nov 2020*.

183. *https://dceg.cancer.gov/research/cancer-types/breast-cancer/confluence-project, Accessed Nov 2020*.

184. Barlow, L., et al., *The completeness of the Swedish Cancer Register: a sample survey for year 1998*. Acta Oncol, 2009. **48**(1): p. 27-33.

185. Wennman-Larsen, A., et al., *Can breast cancer register data on recommended adjuvant treatment be used as a proxy for actually given treatment?* European Journal of Oncology Nursing, 2016. **1;22:1-7**.

186. Johansson, L.A., C. Björkenstam, and R. Westerling, *Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995.* Journal of clinical epidemiology, 2009. **1;62(11):1202-9**.

187. Tornberg, S., et al., *A pooled analysis of interval cancer rates in six European countries.* Eur J Cancer Prev, 2010. **19**(2): p. 87-93.

188. Conant, E.F., et al., *Improving accuracy and efficiency with concurrent use of artificial intelligence for digital breast tomosynthesis.* Radiology: Artificial Intelligence, 2019. **31;1(4):e180096**.

189. *U.S. Patent No. 8,855,388, entitled MICROCALCIFICATION DETECTION CLASSIFICATION IN RADIOGRAPHIC IMAGES, issued 10/2/2014.*

190. Rothman, K.J., S. Greenland, and T.L. Lash, *Modern epidemiology*. 2008: Lippincott Williams & Wilkins.

191. Pearce, N., *Classification of epidemiological study designs.* Int J Epidemiol, 2012. **41**(2): p. 393-7.

192. Concato, J., N. Shah, and R.I. Horwitz, *Randomized, controlled trials, observational studies, and the hierarchy of research designs.* N Engl J Med, 2000. **342**(25): p. 1887-92.

193. Greenland, S. and H. Morgenstern, *Matching and efficiency in cohort studies.* Am J Epidemiol, 1990. **131**(1): p. 151-9.

194. Prentice, R.L., *A case-cohort design for epidemiologic cohort studies and disease prevention trials.* Biometrika, 1986. **1;73(1):1-1**.

195. Barlow, W.E., et al., *Analysis of case-cohort designs.* J Clin Epidemiol, 1999. **52**(12): p. 1165-72.

196. Schulz, K.F. and D.A. Grimes, *Case-control studies: research in reverse.* Lancet, 2002. **359**(9304): p. 431-4.

197. Hennessy, S., et al., *Factors influencing the optimal control-to-case ratio in matched case-control studies.* Am J Epidemiol, 1999. **149**(2): p. 195-7.

198. Goldberger, A.S., *Best linear unbiased prediction in the generalized linear regression model.* Journal of the American Statistical Association, 1962. **1;57(298):369-75**.

199. Tolles, J. and W.J. Meurer, *Logistic Regression: Relating Patient Characteristics to Outcomes.* JAMA, 2016. **316**(5): p. 533-4.

200. Tibshirani, R., *Regression shrinkage and selection via the lasso.* Journal of the Royal Statistical Society: Series B (Methodological), 1996. **58(1):267-88**.

201. Greenland, S., *Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies.* Am J Epidemiol, 2004. **160**(4): p. 301-5.

202. Cox, D.R., *Regression models and life-tables.* Journal of the Royal Statistical Society: Series B (Methodological), 1972. **34(2):187-202**.

203. Kalbfleisch, J.D. and R.L. Prentice, *The statistical analysis of failure time data.* John Wiley & Sons, 2011. **2011 Jan 25**.

204. Hastie, T., J. Friedman, and R. Tibshirani, *Model Assessment and Selection, The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Eds. New York, NY: Springer New York, 2001. **pp. 193–224**.

205. Cawley, G.C. and N.L.C. Talbot, *On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation.* Journal of Machine Learning Research, 2010. **vol. 11, no. Jul, pp. 2079–2107**.

206. D'Agostino, R.B., Sr., J.M. Massaro, and L.M. Sullivan, *Non-inferiority trials: design concepts and issues - the encounters of academic consultants in statistics.* Stat Med, 2003. **22**(2): p. 169-86.

207. Rubin, D.B., *Causal inference using potential outcomes: Design, modeling, decisions.* Journal of the American Statistical Association, 2005. **1;100(469):322-31**.

208. Adami, H.O., et al., *Time to abandon early detection cancer screening.* Eur J Clin Invest, 2019. **49**(3): p. e13062.

209. Rosner, B.A., et al., *Simplified breast risk tool integrating questionnaire risk factors, mammographic density, and polygenic risk score: development and validation.* Cancer Epidemiol Biomarkers Prev, 2020.

210. Castells, X., et al., *Risk of Breast Cancer in Women with False-Positive Results according to Mammographic Features.* Radiology, 2016. **280**(2): p. 379-86.

211. Zuckerman, S.P., et al., *Multicenter Evaluation of Breast Cancer Screening with Digital Breast Tomosynthesis in Combination with Synthetic versus Digital Mammography.* Radiology, 2020: p. 200240.

212. Weinstein, S.P., et al., *Abbreviated Breast Magnetic Resonance Imaging for Supplemental Screening of Women With Dense Breasts and Average Risk.* J Clin Oncol, 2020: p. JCO1902198.

213. Fabian, C.J., *Will a Low-Dose Option Improve Uptake of Tamoxifen for Breast Cancer Risk Reduction?* J Clin Oncol, 2019. **37**(19): p. 1595-1597.