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

EPIDEMIOLOGICAL STUDIES ON BREAST CANCER RISK FACTORS AND SCREENING

Xinhe Mao

毛新荷

Karolinska Institutet

Stockholm 2023

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2023

ISBN 978-91-8017-110-6
Cover illustration: 'Embracing a Bright Future' by Xinhe Mao with essential assistance of DALL·E 2 by OpenAI, 2023

Epidemiological studies on breast cancer risk factors and screening

Thesis for Doctoral Degree (Ph.D.)

By

Xinhe Mao

The thesis will be defended in public at Lecture Hall Samuelsson, Karolinska Institutet, Tomtebodavägen 6, Stockholm, October 23rd 2023, 13:00

Principal Supervisor:
Professor Kamila Czene
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Co-supervisor(s):
Professor Wei He
Zhejiang University
School of Public Health
and
Affiliated to research
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Professor Keith Humphreys
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Associate Professor Haomin Yang
Fujian Medical University
School of Public Health
and
Affiliated to research
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Opponent:
Professor Giske Ursin
University of Olso
Institute of Basic Medical Sciences

Examination Board:
Senior Lecturer Jette Möller
Karolinska Institutet
Department of Global Public Health

Professor John Öhrvik
Karolinska Institutet
Centre for Clinical Research

Docent Renske Altena
Karolinska Institutet
Department of Oncology-Pathology

Sapere aude.

Dare to know.

- Horace

I wish I could show you when you are lonely or in darkness the astonishing light of your own being.

- Hafiz

Abstract

This thesis aims to enhance cancer prevention by investigating the factors and outcomes associated with false-positive (FP) mammography recalls, as well as understanding the association between breast cancer risk factors of women and cancer risk among their relatives. Specifically, four studies were conducted using data from Swedish national registers, the Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA) cohort, and the Linné-Bröst1 (Libro-1) cohort.

In **Study I**, we characterized factors associated with FP mammography recalls, comparing women with a FP recall to those who were not recalled and to those who had a true-positive recall (screen-detected cancer). We found that several mammographic and non-mammographic factors, as well as high breast cancer risk scores, were associated with having a FP recall. However, these factors were either equally or more strongly associated with having a true-positive recall.

In **Study II**, using a matched-cohort design, we examined the risk of subsequent breast cancer among women with a FP mammography recall. We observed a long-term increased breast cancer risk after a FP recall, compared with women who were not recalled. The elevated breast cancer risk differed by age and mammographic density at the matching mammography. In addition, the increased risk for breast cancer diagnosed on the ipsilateral side to the FP recall decreased over time and was highest within the first four years of follow-up.

In **Study III**, we investigated whether specific breast cancer risk factors in women were associated with their sisters' breast cancer incidence. We found that for women with high breast cancer risk prediction scores, benign breast disease (BBD), and high mammographic density, there was an increased risk of breast cancer for their sisters.

In **Study IV**, we investigated the associations of both carriership of protein-truncating variants (PTV) in eight genes and breast cancer polygenic risk scores (PRS) in women, with the risk of cancers in their first-degree relatives. We observed an elevated breast cancer risk among female relatives of women with PTVs, and among those with high breast cancer PRS. Additionally, we found a slightly elevated risk of cancers related to hereditary breast and ovarian cancer syndrome (HBOC)—other than breast cancer—among relatives of women with either high PRS or PTVs in the studied genes.

In summary, this thesis provides valuable information for both screening processes and genetic counseling. Although none of the studied factors are viable for interventions aimed at reducing FP recalls—due to the risk of simultaneously missing tumors—our results may aid in tailoring individualized surveillance plans for women with a FP recall. Additionally, our results suggest that women's breast density and breast cancer risk scores—information that will be available at screening—may be useful for estimating the breast cancer risk in their sisters. Furthermore, PTVs in non-*BRCA* genes might offer insights into cancer aggregation in families. Overall, this thesis advances evidence-based cancer prevention in the era of precision medicine.

List of scientific papers

I. **Mao X**., He W., Humphreys K., Eriksson M., Holowko N., Strand F., Hall P., Czene K. *Factors associated with false-positive recalls in mammography screening.* J Natl Compr Canc Netw. 2023 Feb;21(2):143-152.e4.

II. **Mao X**., He W., Humphreys K., Eriksson M., Holowko N., Yang H., Hall P., Czene K. *Breast cancer incidence after a false-positive mammography result: a population-based study with over 20 years of follow-up.* Submitted.

III. **Mao X**., He W., Eriksson M., Lindström L., Holowko N., Lagercrantz S., Hammarström M., Grassmann F., Humphreys K., Easton D., Hall P., Czene K. *Prediction of breast cancer risk for sisters of women attending screening.* J Natl Cancer Inst. 2023 May 27:djad101. Epub ahead of print.

IV. Xiao Q., **Mao X**., Ploner A., Grassmann F., Rodriguez J., Eriksson M., Hall P., Czene K. *Cancer risks among first-degree relatives of women with genetic predisposition to breast cancer: a population-based cohort study*. Submitted.

Scientific paper not included in the thesis

I. Zhang Y., **Mao X.**, Yu X., Huang X., He W., Yang H. Bone mineral density and risk of breast cancer: A cohort study and Mendelian randomization analysis. Cancer. 2022 Jul 15;128(14):2768-2776.

Contents

List of abbreviations

BBD	Benign breast disease
BI-RADS	Breast imaging reporting and data system
BMI	Body mass index
BOADICEA	Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm
CAD	Computer-aided detection
CEDM	Contrast-enhanced digital mammography
CI	Confidence interval
DBT	Digital breast tomosynthesis
ER	Estrogen receptor
FP	False-positive
GWAS	Genome-wide association studies
HBOC	Hereditary Breast and Ovarian Cancer syndrome
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
HRT	Hormone replacement therapy
ICD	International Classification of Diseases
KARMA	Karolinska Mammography Project for Risk Prediction of Breast Cancer
Libro-1	Linné-Bröst1
OR	Odds ratio
PR	Progesterone receptor
PRS	Polygenic risk score
PTV	Protein-truncating variant
SNOMED	Systematized Nomenclature of Medicine
SNP	Single nucleotide polymorphism
TNBC	Triple-negative breast cancer
TNM	Tumor, Node, Metastasis
TP	True-positive

1 Introduction

Breast cancer has long been the most common cancer and one of the leading causes of cancer death among females [1-3]. Consequently, enormous efforts have been made to prevent breast cancer and reduce mortality from the disease. Primary prevention of breast cancer focuses on preventing the disease before it occurs by modifying lifestyles or employing chemoprevention, while secondary prevention aims to detect and treat the disease in its early stages to achieve better prognosis [4, 5].

Mammography screening is widely used as a secondary prevention method to detect breast cancer early, and it has been implemented in many developed countries for several decades [6]. Studies consistently show that mammography screening has reduced breast cancer death rates by over 20% [6-9]. However, false-positive (FP) results, in which women are recalled for further examinations but are found to be free of breast cancer, are common [6, 10-12]. Since a FP result can lead to a psychological burden and may influence subsequent participation rates in screening, it is considered one of the major harms of mammography screening [13-15]. Currently, there are no interventions to reduce FP mammography recalls in screening, and understanding of the subsequent long-term breast cancer risk is limited.

Predicting cancer risk has become a critical focus, as identifying individuals at high risk can potentially facilitate more effective primary or secondary prevention strategies. In recent decades, many genetic and non-genetic risk factors for breast cancer have been identified [16, 17]. Based on these findings, risk prediction tools have been developed to provide more personalized assessments of breast cancer risk [18-21]. While a breast cancer diagnosis in a woman can be useful to estimate the breast cancer risk for her relatives, it is still unclear whether other factors of a woman can be used to estimate her relatives' risk of developing the disease. Considering that both common and rare variants of breast cancer are associated with cancers at other sites [22, 23], we also examine the association between women's genetic risk factors and the risk of cancer at other sites in their relatives.

This thesis is dedicated to contributing to the prevention of breast cancer by exploring several key areas of concern. Specifically, we investigate whether breast cancer risk factors can be potentially harnessed to minimize FP mammography results, whether and what surveillance programs might be beneficial for women who received a FP result, and whether information on breast cancer risk factors of women could prove useful to their relatives in terms of assessing both breast cancer risk and cancer risk at other sites.

2 Background

2.1 Breast cancer epidemiology

Breast cancer has consistently been recorded as the most common cancer diagnosed in females [1-3, 24]; in 2020, with an estimated more than 2.3 million cases diagnosed, female breast cancer surpassed the number of new lung cancer cases and became the most commonly diagnosed cancer overall [1]. It is estimated that one in eight to ten females will be diagnosed with breast cancer during their lifetime [25, 26]. The incidence of breast cancer has been increasing for decades, a trend that is attributed to changes in lifestyle factors, reproductive patterns, and increases in screening programs and heightened disease awareness [27].

Breast cancer is the leading cause of death from cancer in women [1]. It was estimated that globally, around 0.68 million women died from breast cancer in 2020 [1]. Developed countries have seen mortality rates decrease for decades, primarily due to advancements in breast cancer treatments and the implementation of mammography screening programs [25, 28].

In Sweden, breast cancer is the most frequent malignancy among females [29, 30]. Similar to the trend observed in other developed countries, the incidence of breast cancer in Sweden has increased in the last half-century, while the mortality rate has decreased, as shown in Figure 2.1.

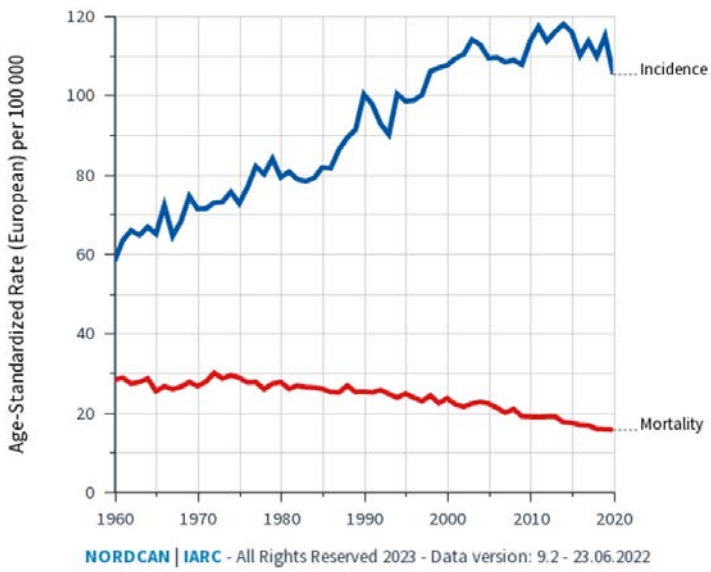

Figure 2.1 *Breast cancer incidence and mortality among females in Sweden [31].*

2.2 Risk factors of breast cancer

Both genetic and environmental factors play important roles in breast cancer carcinogenesis. The following are common risk factors for breast cancer, classified into categories including

family history of breast cancer, genetic factors, hormone-related factors, lifestyle factors, mammographic features, and history of benign breast diseases (BBD).

2.2.1 Family history of breast cancer

Family history of breast cancer has long been recognized as a strong risk factor for breast cancer [32, 33]. Compared to women without any relatives affected by breast cancer, women with one affected first-degree family member (including mothers, sisters and daughters) are associated with around a 1.8-fold increase in risk [34]. Factors such as the degree of kinship (first or second), the number of affected family members, and relatives' age at breast cancer diagnosis influence the magnitude of this risk [32, 34, 35]. Generally, risk is higher when the kinship is closer, the number of affected relatives is greater, and the age of onset is younger (e.g., <50 years old). For example, compared with women without a positive family history, the risk of breast cancer increases 1.5-fold for those with one affected second-degree relative and 2.93-fold for those with two affected first-degree relatives [34]. If multiple cases of breast cancer exist within a family, these cancers can be categorized as familial breast cancer [36]. This familial aggregation of breast cancer, to a large extent, is due to inherited factors [37-39].

2.2.2 Genetic risk factors for breast cancer

The heritability of breast cancer has been estimated to be 25-31 percent in twin or family-based studies [40-42]. Genetic risk predisposition to breast cancer can be classified into three categories: highly penetrant but rare variants, moderately penetrant but rare variants, and low-penetrance but common variants. In Figure 2.2, selected genes are used to illustrate these categories by plotting the magnitude of each variant's association with breast cancer against its minor allele frequency.

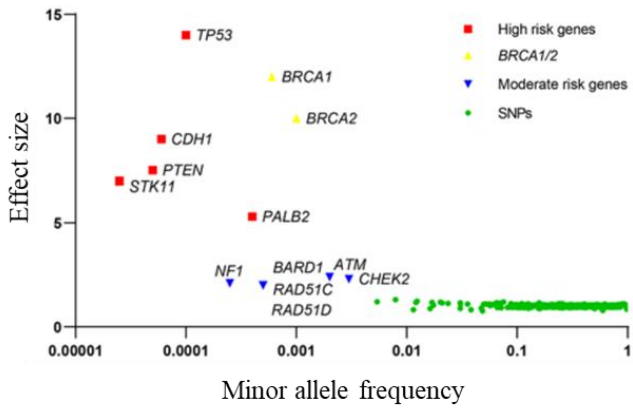

Figure 2.2 *Breast cancer susceptibility by minor allele frequency and estimated relative risk.* Image sourced from [43].

Approximately 5-10% of all breast cancer are caused by inherited high- or moderate-penetrance variants and are known as inherited breast cancer [44]. The most common causes of inherited breast cancer are genetic mutations in either *BRCA1* or *BRCA2* [45, 46]. Since the discovery of *BRCA1* and *BRCA2* mutations in linkage studies [47-49], other rare variants (the

minor allele frequency of <0.01), which confer a high or moderate increase in the risk of breast cancer, have been identified. These include mutations in genes such as *PTEN*, *TP53*, *CDH1*, *STK11*, *CHEK2*, *BRIP1*, *ATM*, and *PALB2* [50-58].

For each gene, the reported relative risk associated with the variants varied depending on several factors including the specific loci, type of variants (such as protein-truncating or missense variants), sample size of the study, and the length of the follow-up period. Dorling et al. conducted a study with the largest sample size to date, potentially providing the most precise association of risk estimates [59]. They found strong evidence of an increased breast cancer risk in protein-truncating variants (PTV) of nine genes, including *ATM*, *BRCA1*, *BRCA2*, *CHEK2*, *PALB2*, *BARD1*, *RAD51C*, *RAD51D*, and *TP53* [59]. Owing to the rarity of missense variants, Dorling and colleagues found limited new information regarding these variants, with the exception of clear evidence related to missense variants in the *CHEK2* gene [59].

Studies have also shown that some of these genetic mutations can result in the clustering of other cancers within the family. For example, *BRCA1* and *BRCA2* mutations can lead to Hereditary Breast and Ovarian Cancer syndrome (HBOC), meaning that carriers are more likely to develop breast, ovarian, prostate, pancreatic cancers and melanoma than non-carriers in their lifetime [22, 60]. Inherited mutations in *TP53* can lead to Li-Fraumeni Syndrome, which increases the likelihood that carriers develop early-onset cancers, including breast cancer, leukemia, brain tumors, and sarcomas [45, 46, 61].

In the recent decade, thanks to advanced laboratory techniques and sequencing capabilities, a rising number of single nucleotide polymorphisms (SNPs) associated with breast cancer risk have been identified in genome-wide association studies (GWAS) carried out by big international consortia [62-65]. A SNP is a substitution of a single nucleotide at either a gene or intergenic region, occurring in 1% or more of the population [65]. Though the risk of breast cancer associated with each SNP is mild, their combined effects (computed into a weighted score) have great potential in predicting breast cancer. This combined weighted score is usually referred to as a polygenic risk score (PRS) [66, 67]. In Mavaddat et al's study, the results showed that women in the top centile of the best-performing breast cancer PRS (based on 313 SNPs) were associated with an over 30% lifetime probability of developing breast cancer. [68]. In a recent GWAS study, 32 new loci were identified, and when coupled with 178 risk loci identified previously, these 210 allele variants could explain around 18% of the two-fold breast cancer familial risk [64]. GWAS have also provided evidence that some SNPs can affect more than one phenotype [69, 70], and associations between many diseases and breast cancer susceptibility genes, or vice versa, have been identified [23, 71, 72].

2.2.3 Hormone-related factors

Hormone-related factors can be categorized into endogenous hormone-related factors and exogenous hormonal medicines. Both early age at menarche and late age at menopause are positively associated with breast cancer risk, as they can prolong the duration of a woman's reproductive years [73, 74]. In a meta-analysis study, researchers found that an additional

reproductive year due to earlier menarche was associated with a greater risk than an additional reproductive year due to later menopause. This finding suggests that the effect on breast cancer risk might not be solely attributed to the lengthening of a woman's reproductive years [75].

Reproductive factors are also well-studied as breast cancer risk factors. Nulliparity has consistently been shown as a risk factor [73, 74], whereas having more children, early age at full-term pregnancy are shown as protective factors against developing breast cancer [74, 76-78]. Women who have breastfed have a lower risk of breast cancer compared to women who have never breastfed. Moreover, studies suggest that breastfeeding serves as a protective factor against hormone-receptor-negative breast cancer [79-81].

Exogenous hormones usually come in two forms: the use of hormone replacement therapy (HRT) and hormonal contraceptives [82]. The use of HRT, also referred to as menopausal hormonal treatment, is shown to be a risk factor for breast cancer [83-85]. The harm-benefit profile of HRT has been controversially argued [86]. HRT was first introduced in clinical settings in the 1940s to treat symptoms of menopause and gained popularity in the late 1960s. Its use declined abruptly due to a study that found HRT to be associated with an increased risk of breast cancer and coronary heart disease [86, 87]. In several developed countries, this decline in HRT usage partially contributed to a fall in breast cancer incidence rates in the early 2000s [27]. However, more recent data suggest a favorable harm-benefit profile for HRT in young women (age<60), including reduced risks of coronary heart disease and all-cause mortality [86]. In terms of breast cancer risk, studies indicate that the use of estrogen alone, rather than in combination with progestogen, may be of less concern [86]. Taking either combined oral contraceptives or progestogen-only hormonal contraceptives is associated with an increased breast cancer risk, though the magnitude of this risk is considered to be slight [88-90].

2.2.4 Lifestyle factors

Lifestyle factors are significant contributors to the onset of breast cancer. The global increase in breast cancer incidence in recent decades is largely attributed to lifestyle changes, including increased alcohol consumption, smoking, higher body mass index (BMI), and sedentary behaviors [91-98]. Even moderate or light alcohol consumption is associated with an increased breast cancer risk [94, 95]. Smoking is moderately associated with breast cancer risk, particularly for those who started smoking at a young age or before giving birth to their first child [92, 93]. Studies have confirmed an increased risk of breast cancer due to high BMI among postmenopausal women [99, 100], whereas an inverse association is generally observed between obesity and breast cancer risk among premenopausal women [101, 102]. However, results indicated the positive association between obesity and premenopausal breast cancer risk in the Asian population [101, 103]. While sedentary behavior is associated with an elevated breast cancer risk [104], high levels of physical activity have consistently been shown to be associated with a significant reduction in breast cancer risk [91, 96-98].

2.2.5 Mammographic features

With the implementation of mammography screening programs in many developed countries, mammographic features have gained increasing attention. Common mammographic features, which are also included in this thesis, include mammographic density, microcalcifications, and masses. Mammographic density is a strong and well-known risk factor for breast cancer; women with density over 75% of the breast have a 4-5 times greater risk of breast cancer than women with no or little density in the breast [105]. Mammographic microcalcifications on mammograms are considered an early sign of breast cancer [106-108]. While the majority of masses, particularly those in young women, result from benign diseases, some suspicious masses are signs of breast cancer [109]. More detailed information on mammographic features is summarized in the mammography screening section.

2.2.6 Benign breast disease

BBD is a term describing a broad category of breast diseases. Generally, the risk of developing BBD begins to rise during one's 20s and peaks in one's 40s to 50s. This is quite different from the risk profile of breast cancer, which is low at a younger age and increases after menopause [110]. The associations between a family history of breast cancer, hormonal factors, reproductive factors, and BMI with BBD are complex and vary by subtypes. For example, at premenopausal ages, women with nulliparity have a reduced risk of epithelial proliferation without atypia but an increased risk of cysts [111]. Since the implementation of mammography screening, the diagnosis of BBD has become common. Usually, based on histologic classification, BBD is categorized into proliferative diseases with atypia, e.g., epithelial proliferation with atypia; proliferative diseases without atypia, e.g., epithelial proliferation without atypia, fibroadenoma; and non-proliferative diseases, e.g., cysts [112, 113]. Proliferative BBDs, especially those with atypia, have long been recognized as a strong risk factor for breast cancer and are associated with a 2-5 fold increased risk of the disease. In contrast, non-proliferative BBDs are associated with only a slightly increased breast cancer risk [112, 114].

2.3 Prediction tools

Many breast cancer risk prediction models have been developed based on identified risk factors. These models aim to identify individuals at high risk of developing breast cancer, thereby guiding medical counseling and cancer prevention. The first breast cancer prediction model was developed in 1989, known as the Breast Cancer Risk Assessment Tool or the Gail model [115]. The Tyrer-Cuzick and Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) models were later developed in 2004 [116, 117]. These risk models have been updated several times to improve their performance by including additional risk factors. In recent years, due to the growing recognition of the predictive potential of mammographic density and PRS, these two factors have been incorporated into the latest versions of these models [19, 118, 119]. The Tyrer-Cuzick and BOADICEA models are among the most validated and well-performing [120, 121]. In a large independent cohort comprising

more than 15,000 women and featuring long-term follow-up (with a median follow-up of around 11 years), the BOADICEA and Tyrer-Cuzick models performed better than the Gail and BRCAPRO models, which are also commonly used, in breast cancer risk prediction [119].

While the majority of risk models aim to predict breast cancer risk over a 5- or 10-year period, our group has developed the KARMA model, which focuses on predicting short-term risk (a 2-year risk) to aid in improving breast cancer screening. This model uniquely incorporates computer-aided detection (CAD) of suspicious masses and microcalcifications [121-123]. Because the KARMA, Tyrer-Cuzick, and BOADICEA models are potential tools for future risk-based breast cancer screening and each incorporates different factors, we included them in this thesis. Table 2.1 provides an overview of the factors included in the most updated versions of these three risk models.

Table 2.1 *Summary of risk factors included in the three risk models.*

Risk factors	BOADICEA	Tyrer-Cuzick	KARMA
Age	Yes	Yes	Yes
Birth year	Yes	NA	NA
Ashkenazi origin	Yes	Yes	NA
Country of birth	Yes	NA	NA
Family history of breast cancer	Yes	Yes	Yes
Family history of ovrian cancer	Yes	Yes	NA
Family history of prostate, male breast and pancreatic cancer	Yes	NA	NA
BRCA1/2 mutations	Yes	Yes	NA
PALB2, *CHEK2*, *ATM* mutatioins	Yes	NA	NA
Polygenic risk score	Yes	Yes	Yes
Age at menarche	Yes	Yes	NA
Menopause status	NA	Yes	Yes
Age at menopause	Yes	NA	NA
Nulliparity (yes or no)	NA	Yes	NA
Parity (number of live births)	Yes	NA	NA
Age at first birth	Yes	Yes	NA
Hormone replacement therapy	Yes	Yes	Yes
Oral contraceptives	Yes	NA	NA
Alcohol	Yes	NA	Yes
Smoking	NA	NA	Yes
Height	Yes	NA	NA
Body mass index	Yes	Yes	Yes
Benign breast disease	NA	Yes	NA
Breast tumor pathology	Yes	NA	NA
Mammographic density	Yes	Yes	Yes
Mass	NA	NA	Yes
Microcalcification	NA	NA	Yes

Note: NA, not available.
BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer.

2.4 Mammography screening

2.4.1 Mammography screening

Mammography is a technology that uses low-dose X-rays to detect breast cancer. The mammography screening program is an effective secondary preventive method that has contributed to a 20-40% reduction in the mortality rate of breast cancer by detecting the disease at early stages [7-9]. The starting age for mammography screening and the intervals between screenings vary from country to country. In most European countries, women aged 50 and older are invited for mammography screening every two years until turning 69 or 74 [124-130]. In Britain, women aged 50 to 70 are invited for breast cancer screening every three years [131]. In the US, there is neither a centrally organized breast cancer screening program nor a single standard recommendation [131]. In May 2023, the United States Preventive Services Task Force drafted an updated recommendation for breast cancer screening, encouraging women aged 40 to 74 to have mammograms at 2-year intervals [132]. Meanwhile, the American Cancer Society recommends that women begin annual mammography at age 45 (with the option to start at 40) until age 54, and then participate in mammography screening every other year thereafter [133]. Figure 2.3 provides information on the screening process.

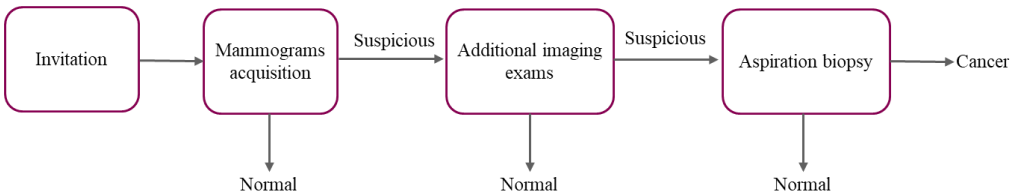

Figure 2.3 *An overview of the screening process.*

In Sweden, the updated National Board of Health and Welfare recommends that women aged 40 to 74 years old have a free mammogram every 18 to 24 months [134]. The starting age for screening and the length of the screening interval vary by region [135]. In Stockholm County, the mammography screening program started in January 1989, and all women aged 50 to 69 years old were invited to have mammograms every 24 months [124, 136]. Since 2005, women aged 40 to 49 in Stockholm have been invited for mammography screening at 18-month intervals (which was changed to a 2-year interval in 2012). Since 2012, women aged 70 to 74 have also been included in the program [124]. In this thesis, when screening information was used in the analysis, only women who were invited to attend mammography screening in the Stockholm region were included.

2.4.2 Screening process and screening outcomes

For every woman at each screening, mammograms are taken from both mediolateral oblique and craniocaudal views. According to the European guideline [137], the mammograms are independently read by two radiologists, who each give an interpretation (normal or suspicious). If the results differ, the radiologists discuss until they reach a consensus. For women whose mammograms are suspicious for breast cancer, further imaging or cytology examinations are

to be conducted [124]. Based on the result at each screening, the screening outcomes can be classified into four categories. The definition and frequency of screening outcomes are shown in Figure 2.4.

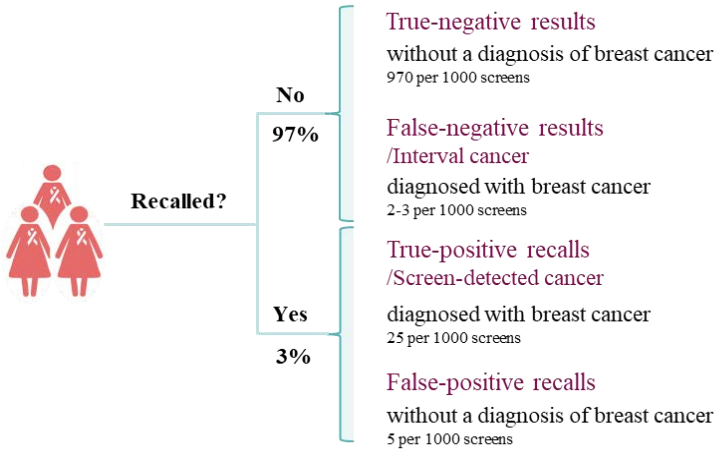

Figure 2.4 *Mammography screening outcomes.* Note: The frequencies of the screening outcomes are based on the data presented in this publication [124] and represent the frequencies in the Stockholm mammography screening program. To define the four screening outcomes at each screening round, we assessed whether a diagnosis of breast cancer had been made or not before the end of the normal screening interval or before the next scheduled screening.

As mentioned in the Introduction, a FP recall is one of the major concerns of mammography screening programs. This issue stems from the high rate of FP recalls [10, 11], the potential psychological burden they can impose on women, and the subsequent influence they may have on re-attendance rates to the program [13-15]. Addressing the occurrence of FP mammography recalls is a critical challenge and an essential focus of ongoing research and consideration. Furthermore, although various published studies have confirmed an increased risk of breast cancer [127, 138, 139], the long-term risk associated with a FP recall remains poorly understood. In this thesis, we present results which help better understand factors associated with it and long-term outcomes of it.

2.4.3 Mammographic features

2.4.3.1 Mammographic density

Due to the distinct X-ray attenuation features of fat, epithelium, and stroma tissues, breasts appear uneven in brightness on mammograms [140]. Fat tissue appears dark and radiolucent on mammograms, while epithelial cells and stroma appear light [141]. Radiologically dense breasts are described as those with more epithelium and less fat; by contrast, breasts with less epithelium and more fat are referred to as non-dense breasts. Mammographic density is a highly heritable trait [142, 143] and can be influenced by many breast cancer risk factors, largely through effects on women's sex hormone levels [105, 142, 144]. Additionally, mammographic density decreases with age [145-148].

Several systems have been created to describe mammographic density either qualitatively or quantitatively. Among them, one of the most commonly used classifications is the Breast Imaging Reporting and Data System (BI-RADS), developed by the American College of Radiology based on radiological assessment [149]. The BI-RADS system categorizes density into four groups. Definitions and illustrations for each group can be found in Figure 2.2.

With the aim of developing a completely automated density measurement method, STRATUS was developed in our group [150, 151]. Based on ImageJ software, STRATUS measures mammographic breast density as both absolute dense area and percent mammographic density [150]. The density data used in this thesis were measured by this method.

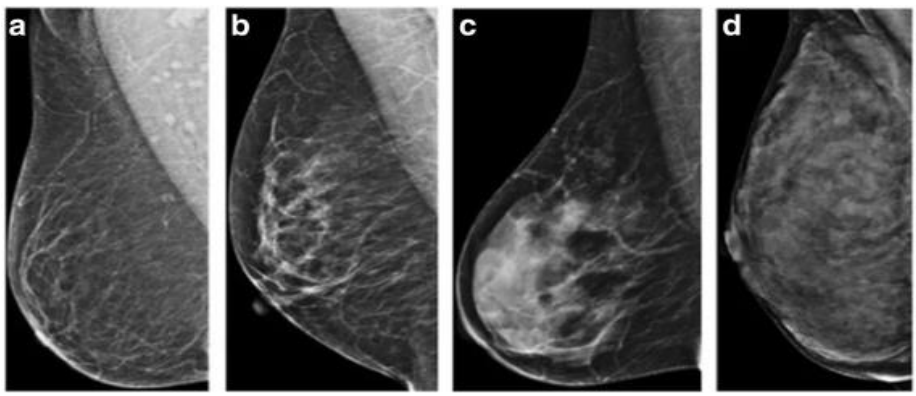

Figure 2.5 *Categories for mammographic density according to BI-RADS, 5th edition.* Image from [152]. BI-RADS, Breast Imaging Reporting and Data System.

Dense breasts not only lead to reduced sensitivity and specificity in mammography screening due to masking [105], but also, as mentioned in the previous section, are associated with an increased risk of breast cancer. In over 30 states in the U.S., women are informed about their breast density when they have a mammogram, and by September 2024, all women will receive mammography reports with density information [153]. In Europe, women do not typically receive breast density information. However, in 2022, the European Society of Breast Imaging recommended that women should be informed about their breast density to enable them to make choices [154].

2.4.3.2 Microcalcifications and masses

Breast calcification refers to the accumulation of calcium in breast tissue, appearing as either macro or micro features. Microcalcifications, which are deposits of calcium smaller than 1 mm, are associated with invasive breast cancer and/or ductal carcinoma in situ [149, 155, 156]. The potentially malignant microcalcifications can be identified by their morphology and distribution. Microcalcifications that appear as coarse heterogeneous, fine pleomorphic, or in segmental distribution or clusters tend to be suspicious and may be malignant [149, 157]. Microcalcifications serve as indicators for around 50% of non-palpable breast cancers in the context of mammography screening [158].

Breast masses are another common feature on mammograms, and they could represent both benign and malignant changes in the breasts. Similar to microcalcifications, benign and malignant masses can often be distinguished by their shape and margin. Oval or round masses are usually benign, while those with irregular, indistinct, or spiculated margins suggest a malignant finding [149].

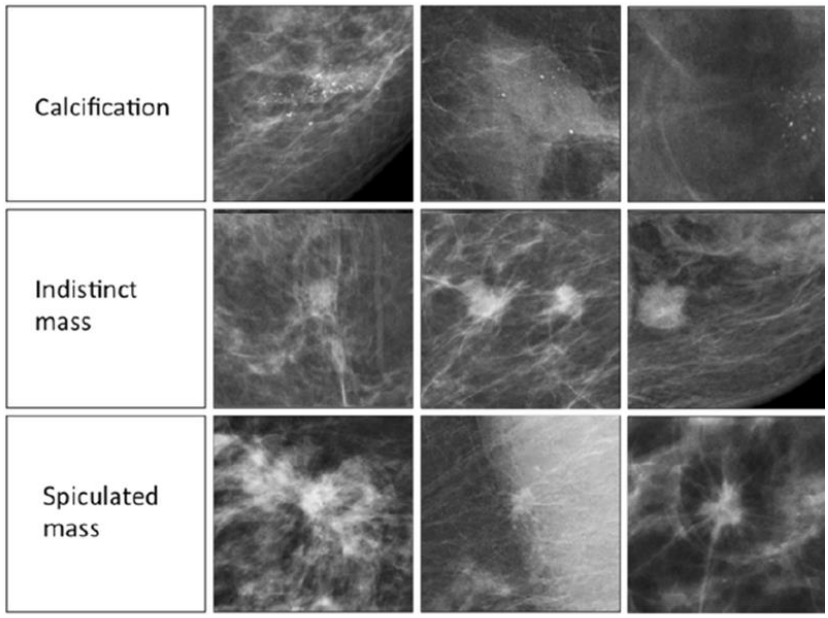

Figure 2.6 *Examples of calcification and masses on mammograms.* Adapted from [159].

Computer-Aided Detection (CAD) software, an FDA-approved Class III device, can identify microcalcifications and masses with a high probability of being malignant, corresponding to suspicious morphologies in BI-RADS scores 3-5 [160, 161]. In this thesis, the data for microcalcifications and masses were obtained using this software.

2.4.4 Other modalities in breast imaging

Besides digital mammography, there are many other modalities used in breast imaging. Some are currently used as supplemental imaging options or for special occasions, while others have the potential to become screening modalities in the future. Table 2.2 provides an overview of breast imaging modalities, along with their pros and cons [162-164].

Table 2.2 *Summary of common breast imaging modalities.*

Technology	Modalities	Pros [a]	Cons [a]	Current use
Ultrasound	Ultrasound	Widely available Higher sensitivity No ionizing radiation	Higher operator dependency Lower specificity	Used as a supplemental imaging modality in screening
Magnetic field	Magnetic Resonance Imaging (MRI)	Higher sensitivity No ionizing radiation	Expensive and not widely available Intravenous gadolinium injection Lower specificity	Used in women with high risk (with a rare genetic mutation)
X-ray	Digital breast tomosynthesis (DBT)	Widely available Higher sensitivity Higher specificity	Higher ionizing radiation	Used as a screening modality; used in the diagnosed setting
X-ray	Contrast-enhanced digital mammography (CEDM)	Widely available Higher sensitivity Higher specificity	Intravenous contrast injection Higher ionizing radiation	Used in the diagnostic setting

Note: [a] All comparisons of sensitivity, specificity and ionizing radiation are made with digital mammography.

2.5 Tumor characteristics and prognosis of breast cancer

Tumor characteristics and the prognosis of breast cancer were only studied in Study II and were not the main focus of this thesis. Since the aim of mammography screening is to reduce breast cancer mortality, a brief summary of these topics is provided here for completeness.

Several tumor characteristics serve as prognosticators for breast cancer. Notable among these are tumor size, tumor grade, lymph node involvement, histological grade, estrogen receptor (ER) status, and human epidermal growth factor receptor 2 (HER2) [165-168]. The progesterone receptor (PR), a steroid hormone receptor, is commonly measured in labs using immunohistochemistry techniques, while the clinical value of PR status remains uncertain [169, 170]. Overexpression of HER2 in breast cancer is linked to cancer relapse and poor survival [167]. However, with advancements in HER-2 targeted therapy, the prognosis has notably improved [171]. Breast cancer that is negative for ER, PR, and HER2 status is termed triple-negative breast cancer (TNBC), known for its aggressive nature and poor prognosis [25, 172].

Four molecular subtypes, based on gene expression patterns, are recognized: Luminal A & B, HER2-enriched, and basal-like breast cancer. These molecular subtypes can also be identified using the immunohistochemical markers mentioned above [173]. Following a consensus from the St Gallen International Expert, these molecular subtypes have become important in guiding therapy for early-stage breast cancer in clinical settings [174, 175].

Women with breast cancer detected during screening (screen-detected breast cancer) generally exhibit less aggressive tumor characteristics compared to those with non-screen-detected cancer [176-182]. In other words, screen-detected breast cancers are more likely to be ER-positive, have no lymph node involvement, be at earlier stages, have smaller tumor sizes, and be at a lower grade at diagnosis. As a result, patients with screen-detected cancers typically have a more favorable prognosis than interval and clinical breast cancer patients [177, 179, 181, 183].

3 Aims and research questions

This thesis aims to better understand and promote cancer prevention by studying the following aspects. First, we focused on factors and outcomes associated with FP recalls — a major issue within the mammography screening program. Second, we studied associations between a woman's breast cancer risk factors and the cancer risk in her relatives. To address these topics, we conducted four studies, each designed to answer the following research questions:

I. What factors are associated with FP recalls in mammography screening; can any of these factors be used to reduce FP recalls? (Study I)

II. Is there a long-term breast cancer risk after a FP recall; does this risk vary by baseline factors, tumor characteristics, and length of time following the FP recall? (Study II)

III. Can breast cancer risk factors and risk scores of women attending breast cancer screening be useful for estimating the breast cancer risk in their sisters? (Study III)

IV. Can both rare and common genetic predispositions to breast cancer in women be useful for estimating the risk of breast cancer and other cancers among first-degree relatives? (Study IV)

4 Data source and linkages

In this thesis, the study population included participants from the mammography screening program in Stockholm-Gotland region as well as from the KARMA and Libro-1 cohorts. For Studies III and IV, first-degree relatives of the participants from the KARMA and Libro-1 cohorts were identified through a linkage to the Multi-generation Register and were also included as part of the study population. All these women and their relatives were linked to various Swedish nationwide registers at an individual level. Figure 4.1 provides an overview of the data sources and the study population used in this thesis.

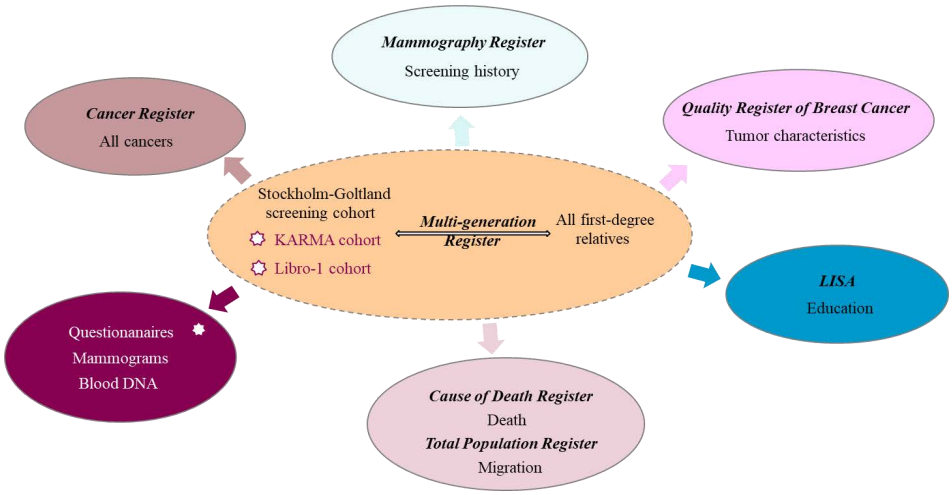

Figure 4.1 *Overview of data sources and study population.* Note: Data from questionnaires, mammograms, and blood DNA were available only for participants from the KARMA and Libro-1 cohorts. KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer; Libro-1, Linné-Bröst1; LISA, Longitudinal Integrated Database for Health Insurance and Labor Market Studies.

4.1 Swedish national registers

Sweden is one of the countries with high-quality, nationwide registers. In 1947, a personal number system was introduced throughout the country [184]. This system became the foundation for national registers and facilitated linkages between them.

In this thesis, the following registers are used.

4.1.1 Multi-generation Register

All individuals born after 1932 and alive in 1961 (if they were born before 1961) are included in the Multi-generation Register. People's personal numbers, as well as those of their parents, were recorded. Based on the parents' information, siblings and other relatives can be identified. The missing data for mothers is around 3%; whilst for fathers, it is about 5% [184].

4.1.2 Mammography Screening Register (Stockholm-Gotland region)

In the Stockholm-Gotland region, the screening program was initiated in 1989. Currently, women in the region aged 40 to 74 are invited to attend the mammography screening every 24 months. Detailed information about screening invitations, attendance, and screening outcomes (for those who participated) has been recorded for each woman at every screening round since 1989. If women were recalled for further examinations, that information was also documented in detail [124, 136].

4.1.3 Swedish Cancer Register

Since 1958, all new cancer diagnoses have been recorded in the Swedish Cancer Register. According to the guidelines set by the Swedish National Board of Health and Welfare, health care providers must report every new cancer case [185]. The Swedish Cancer Register includes the following information: age at diagnosis, date of diagnosis, invasiveness of the cancer, tumor, node, metastasis (TNM) stage, International Classification of Disease (ICD) code, and Systematized Nomenclature of Medicine (SNOMED) code.

4.1.4 Breast Cancer Quality Register

In 1976, the Stockholm region launched its regional breast cancer quality register, recording detailed information on tumor characteristics, treatment, and metastasis for each breast cancer case. The other five regional cancer centers in Sweden initiated similar registers around the same time [186]. In 2008, all six regional quality registers combined to initiate the national breast cancer quality register. The completeness of the national breast cancer quality register was high, with a coverage rate of 99.9% between 2010 and 2014. Missing values were less than 5% for most of the variables [187].

4.1.5 Other registers

The Swedish Cause of Death Register was initiated in 1952; virtually all deaths have been included in this register since then. The underlying cause of death was also recorded for each individual [188]. The Total Population Register was launched in 1967, and migration data, among other information, were recorded for each individual [189]. The Longitudinal Integrated Database for Health Insurance and Labor Market Studies (LISA) register contains information on socioeconomic status, such as educational levels [190].

4.2 The KARMA cohort

KARMA is a prospective mammography screening cohort study that includes 70,877 women. Between 2011 and 2013, women undergoing mammography screening or clinical mammography in Stockholm and Skåne were invited to participate. Upon the study entry, participants completed a detailed web-based questionnaire regarding breast cancer risk factors. The questions covered topics such as personal medical history, medication history, reproductive history, lifestyle factors, and family history of diseases related to breast cancer. Approximately 98% of participants also donated blood samples at the KARMA baseline. In

addition, mammograms of the participants were continuously collected each time the women attended the screening after joining the cohort [191].

4.3 The Libro-1 cohort

Libro-1 is a case-only cohort that includes 5,715 women who were diagnosed with breast cancer in the Stockholm region from 2001 to 2008 and were still alive in 2009. Similar to participants in KARMA, those in Libro1 answered detailed questionnaires about their breast cancer risk factors. Blood samples were also collected from these participants [192, 193].

5 Study designs and methods

Figure 5.1 presents a brief overview of the study design, data sources, main measurements, and statistical analyses for the studies included in this thesis.

Study	Study design	Data source of study population	Exposure & Outcome	Main statistical analysis
I	Matched case-control study	KARMA cohort	Exposure: breast cancer risk factors Outcome: false-positive and true-positive recalls	Conditional logistic regression; Logistic regreesion
II	Matched cohort study	Stockholm Mammography Screening Register and KARMA cohort	Exposure: false-positive recalls Outcome: incident breast cancer; mortality	Stratified cox regression; Flexible parametric model
III	Population-based cohort study	KARMA cohort and Swedish Multi-generation Register	Exposure: breast cancer risk factors Outcome: incident breast cancer	Cox regression
IV	Population-based cohort study	KARMA and Libro-1 cohorts and Swedish Multi-generation Register	Exposure: breast cancer genetic predispositions Outcome: incident cancer	Cox regression

Figure 5.1 *Overview of the study designs of constituent studies in the thesis.* KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer; Libro-1, Linné-Bröst1.

5.1 Study design and study population

5.1.1 Matched case-control study

A matched case-control design is commonly used in public health research to assess whether specific exposures are associated with certain outcomes. In this design, for each case, a fixed number of controls — people who do not have the outcome — are matched to the case based on certain criteria [194]. In Study I we selected controls using incidence density sampling [195], where (matched) controls for each case were randomly selected at the exact time point the outcome (for the case) occurred. Figure 5.2 illustrates density sampling for this design.

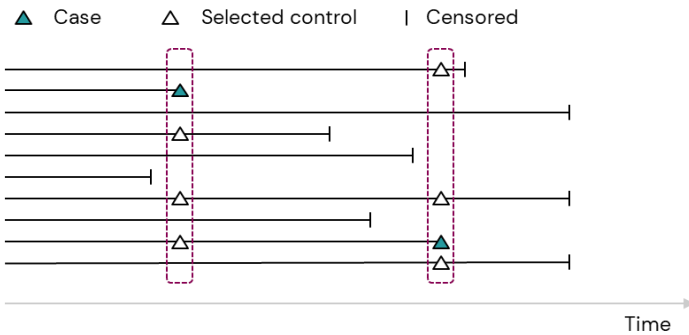

Figure 5.2 *Illustration of density sampling in a matched case-control design.* In this figure, each case is matched to three controls.

5.1.1.1 Study I

In Study I, a total of 29,129 KARMA participants were eligible to be included. Detailed inclusion and exclusion criteria are presented in Figure 5.3. We identified 1,550 women who were recalled at a screening round between 2011 and 2015. For each woman recalled, we randomly selected and matched her with five individuals who were not recalled but were of the same age and screened in the same calendar year. Neither the cases nor the controls had previously been recalled in the mammography screening program. We examined the association between various breast cancer risk factors and a FP recall. We also explored the association between these same risk factors and a TP recall for comparison.

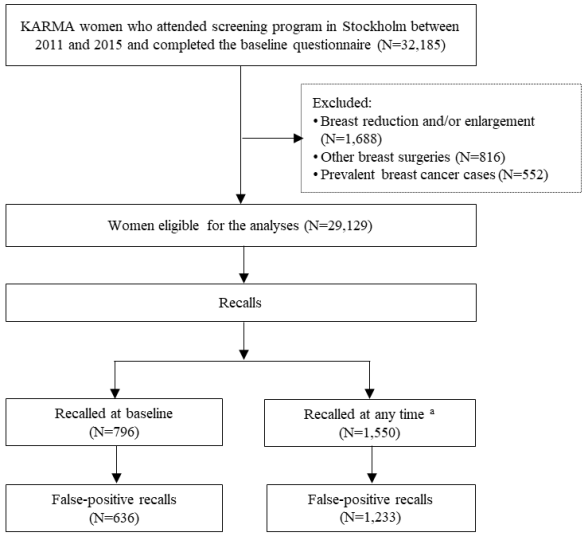

Figure 5.3 *Illustration of the study population in Study I.* Adapted from eFigure1 in Study I.
Note: [a] Women received their first mammography recalls at or after their enrollment in KARMA.

5.1.2 Population-based cohort study and matched cohort design

A cohort study is another common design used in epidemiological research to examine the association between a certain exposure and an outcome. In this design, both exposed and unexposed individuals, who are at risk for the event of interest, are followed until that event occurs [196, 197]. Cohort studies can be conducted either prospectively or retrospectively. Cohort studies can provide time-to-event data to be analyzed, because exposure, occurrence of the outcome, study entry date, and an underlying time scale can all be clearly defined. We used a cohort design in Studies III and IV.

The matched cohort design is an extension of the general cohort design. In a matched cohort design, at the start of the follow-up, each exposed individual is matched with a specific number of unexposed individuals. These matches are based on the same levels of certain confounders, such as age and gender [198]. Both the exposed and unexposed groups are then followed until the event of interest occurs. A matched cohort design was employed in Study II.

5.1.2.1 Study II

In Study II, we used a matched cohort design to examine the long-term risk of breast cancer incidence and mortality following a FP mammography recall. A total of 593,886 women who attended the mammography screening program in Stockholm were eligible to be included in this study. Detailed inclusion and exclusion criteria are presented in Figure 5.4. We identified 45,213 women who received a FP recall between 1991 and 2017. For each woman with a FP recall, we randomly selected and matched her with ten individuals who had true-negative results. The matching was based on age and the calendar year of the mammography. None of the women in either group had been recalled in the mammography screening program prior to the matching. We followed these women until the event of interest (breast cancer or death),

emigration, death (if death was not the event of interest), or March 31, 2020, whichever came first. A total of 12,243 KARMA participants with data on mammographic density were also included for the density analyses.

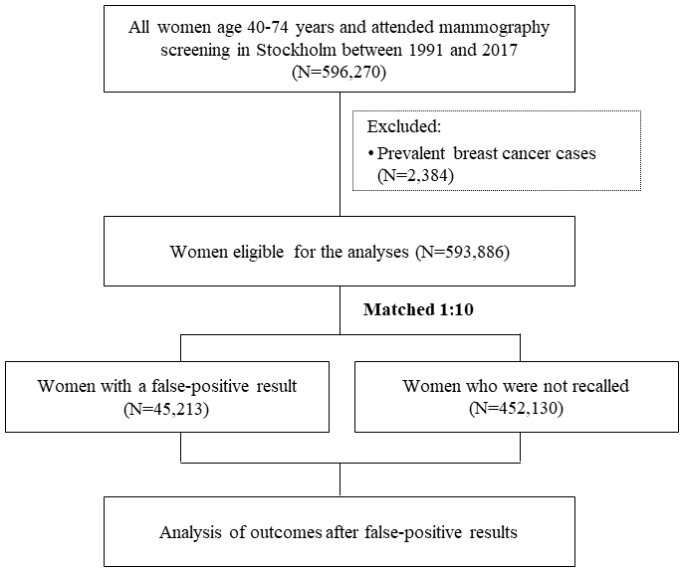

Figure 5.4 *Illustration of the study population in Study II.* Adapted from eFigure1 in Study II.

5.1.2.2 Study III

In Study III, using a cohort design, we investigated whether women's breast cancer risk factors are associated with the risk of breast cancer in their full-sibling sisters. 53,051 KARMA women who were born in Sweden, aged 40-74 years, and had no cancer diagnosis before entering KARMA were eligible for analysis. Through linkage to the Multi-generation Register, we identified 37,998 full-sibling sisters of KARMA women. After exclusions, as described in Figure 5.5, 32,198 sisters were included in the study population. Sisters of the KARMA participants were followed from the enrollment date of the KARMA women until their own breast cancer diagnosis, emigration, death, or October 31, 2019, whichever occurred first. KARMA women were followed in the same way and were analyzed separately for comparison.

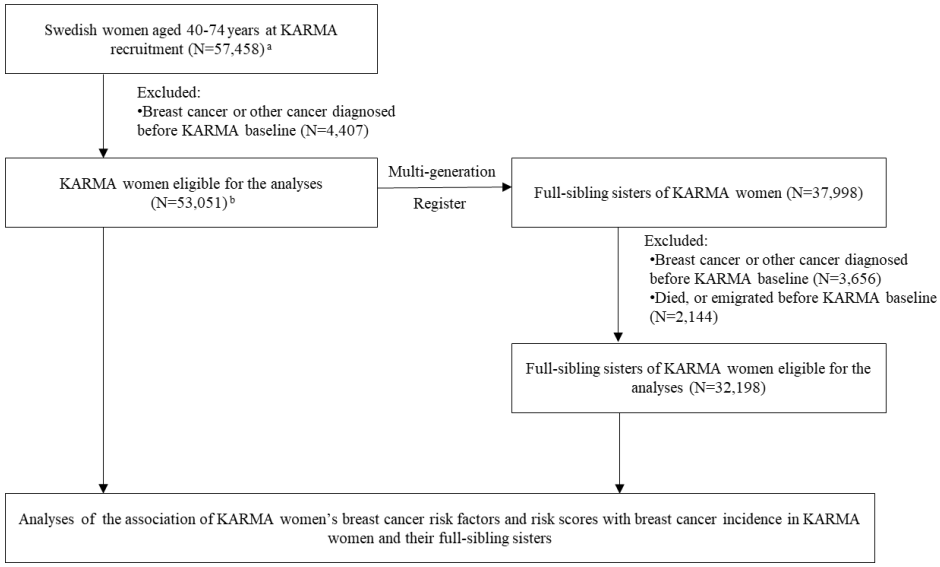

Figure 5.5 *Illustration of the study population in Study III.* Adapted from Supplementary Figure 1 in Study III. Note: [a] Women who were Swedish-born, had completed questionnaires, and were aged 40-74 years at enrollment. [b] For mammographic feature analyses, 3,965 KARMA women who had a breast surgery were excluded. For genetic analyses, only 17,835 women with genotype data were included.

5.1.2.3 Study IV

In Study IV, we used a cohort design to examine the association between women's rare and common genetic risk factors for breast cancer and risk of breast cancer and other cancers among their first-degree relatives. We included 28,362 women with genotype data and 13,226 women with sequencing data from the KARMA and Libro-1 cohorts as index women. Through linkage to the Multi-generation Register, we identified first-degree relatives of index women, which included parents, siblings, and children. Figure 5.6 presents the detailed exclusion criteria. Relatives were followed from January 1, 1958, or from age 20 (whichever came later), until the first diagnosis of cancer, emigration, death, reaching age 80, or December 31, 2017, whichever occurred first.

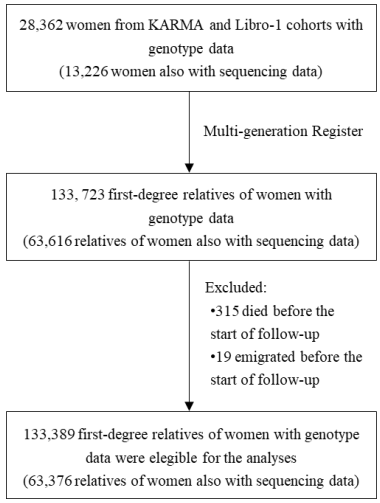

Figure 5.6 *Illustration of the study population in Study IV.* Adapted from Figure 1 in Study IV.

5.2 Main measurements

5.2.1 Screening outcomes

Using data from the Stockholm Mammography Screening Register, we defined screening outcomes for each woman, for every screening round she attended. This information was used in Studies I and II. Among women who were recalled at the screening, those diagnosed with breast cancer by the end of the normal screening interval or by the next screening were referred to as having a TP result or screen-detected breast cancer. Those without a breast cancer diagnosis had a FP result. Among women who were not recalled, those without a breast cancer diagnosis before the next screening or within a normal screening interval had a true-negative result. Those diagnosed with breast cancer before the next screening round had a false-negative result, also known as interval cancer. In Figure 2.4 of this thesis (on Page 10), the definition and frequency of each screening outcome are presented.

5.2.2 Breast cancer and other cancers

Information on the date of breast cancer diagnosis was retrieved from the Swedish Cancer Register and the Breast Cancer Quality Register. This information was used in all the studies in this thesis. Tumor characteristics of breast cancer were defined based on data from the Breast Cancer Quality Register and were used in Study II. Dates of other cancer diagnoses were retrieved from the Swedish Cancer Register and used in Studies III and IV.

5.2.3 Breast cancer risk factors and risk prediction scores

- Common risk factors: Reproductive breast cancer risk factors, use of oral contraceptives and HRT, lifestyle risk factors, BMI, education level, and family history of breast cancer were obtained from the questionnaire data from KARMA participants. This information was used in Studies I and III. In Study II, the education level and

family history of women were sourced from LISA and the Swedish Cancer Register (a breast cancer diagnosis from the women's mothers and/or sisters), respectively.

- Mammographic risk factors: Mammographic density was measured by the Stratus method [199] and was used in Studies I, II, and III. Microcalcifications and masses were measured by the iCAD software [160, 161] and were used in Studies I and III.
- Breast cancer PRS: This score was the summed effect of 313 SNPs reported in this study [68]. The score was weighted by multiplying the per-allele log-odds ratios, which represent the association between each SNP and breast cancer risk, by the number of risk alleles an individual has for each SNP. Breast cancer PRS was used in Studies I, III, and IV.
- PTV in breast cancer risk genes: Information was retrieved from sequencing data using a gene panel of 31 genes associated with breast cancer risk [200]. In Study IV, we focused on PTVs in the following nine genes: *ATM*, *BRCA1*, *BRCA2*, *CHEK2*, *PALB2*, *BARD1*, *RAD51C*, *RAD51D,* and *TP53*. This focus was due to strong evidence of breast cancer risk associated with PTVs in these genes, as found in a large study involving 110,000 women [59]. However, we did not study PTVs in *TP53* because no carriers were found in our study population.
- BBD: Information on the history of BBD (yes or no) was obtained from the questionnaire data of KARMA participants and used in Study I. Through a linkage to the Sympathy pathology record system, we retrieved information on breast biopsies and the date of these records from 1979 to 2015. BBD subtypes were categorized based on European guidelines for breast pathology [201, 202]. We used information on subtypes of BBD and the time of diagnosis from the Sympathy system in Study III.
- Prediction scores: Breast cancer risk scores from the Tyrer-Cuzick, BOADICEA, and KARMA models were calculated and used in Studies I and III [19, 117, 123].

5.3 Main statistical methods

5.3.1 Conditional logistic regression

Logistic regression models are used to examine the association between categorical or continuous variables and a binary outcome/dependent variable. The regression coefficients of a logistic regression model for an (e.g. disease) outcome provide estimates of log odds ratios (OR) assessing the associations between exposures and the outcome [203]. Conditional logistic regression is a special case of logistic regression used for matched case-control studies. In conditional logistic regression, comparisons are performed within each matching stratum – the conditional probability of the observed outcomes of the individuals, given the matching ratio, is evaluated within each stratum, based on assuming a common OR across strata. Matching variables are controlled for in the analyses. In Study I, we estimated the OR with a 95% confidence interval (CI) for a FP or TP based on various breast cancer risk factors using this model.

5.3.2 Cox regression

The Cox proportional hazards regression model is a model that is commonly used for analyzing time-to-event data [204]. It is a semi-parametric model since no assumptions are made regarding the baseline hazard. However, the model does assume proportional hazards over time, i.e., a constant hazard ratio (HR). In Study III, we measured the HR, along with a 95% CI, for breast cancer among women's sisters based on the women's breast cancer risk factors, using this model. In Study IV, we calculated the HR and 95% CI for cancer in women's first-degree relatives based on the women's carriership of PTV in studied genes and breast cancer PRS.

The stratified Cox model is a special type of the Cox model that can be used in matched-cohort studies. Matching identifiers are included in the stratified Cox regression analyses, allowing analyses to be conducted within each matching stratum [198]. A single estimate of HR is obtained after pooling together all HRs from each stratum. In Study II, we used stratified Cox regression analyses to estimate the HR and 95% CI for overall and subtypes of breast cancer, as well as all-cause and breast cancer-specific mortality [205].

5.3.3 Flexible parametric survival method

The flexible parametric survival model is another approach that is used to analyze time-to-event data. Unlike the Cox model, the flexible parametric survival model specifically models the baseline hazard – but by doing so is more flexible in modelling the role of covariates. The flexible parametric survival model uses cubic spline (or other spline) functions and can easily present the HR over time [205, 206]. In Study II, we utilized a flexible parametric model to examine the hazard ratio of breast cancer after a FP recall on both the same and contralateral side as the FP, in comparison to those who were not recalled.

6 Main results

6.1 Factors associated with a false-positive recall

Using the screening records from women at the time of KARMA enrollment, we first aimed to identify factors associated with mammography recalls. Among various factors—such as age, reproductive, hormonal, lifestyle, mammographic, and genetic risk factors, as well as breast cancer risk prediction scores—we found eight risk factors that were positively associated with a mammography recall. These factors were: a family history of breast cancer, history of BBD, mammographic density, the presence of masses and microcalcifications, breast cancer PRS, and risk scores from both the Tyrer-Cuzick and KARMA models.

In contrast, age at mammography was the only risk factor for breast cancer that was negatively associated with mammography recalls. We found that the distribution of FP and TP recalls showed opposite trends by age. Specifically, FP recalls were more frequent among younger women (aged 40-49), whereas TP recalls were more common among older women (aged 60-74) (Figure 6.1). The ratio of FP recalls to TP recalls was close to 10 among younger women, which was significantly higher than the ratio of 2 observed among older women.

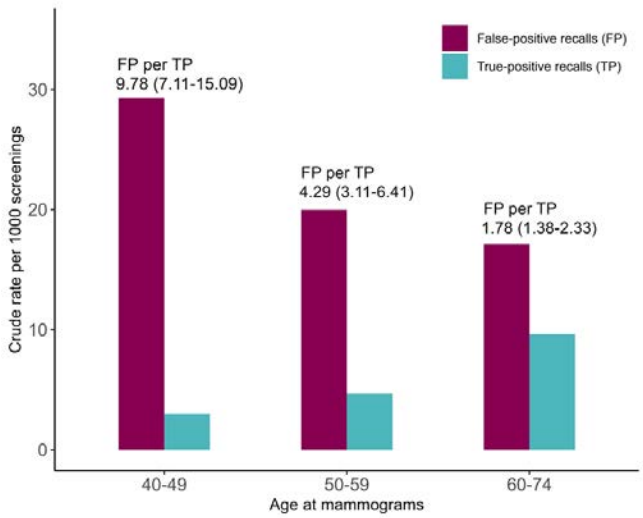

Figure 6.1 *Recall rates of false-positive and true-positive recalls by age at mammography screening.* Adapted from Figure 1 in Study I.

Since age was strongly associated with mammography recalls, we used a matched case-control design (with age as the matching variable) to examine the association between these eight risk factors or risk scores and FP/TP recalls (Figures 6.2 and 6.3). We found that a history of BBD and mammographic dense areas were associated with both FP and TP results to a similar extent when compared to their corresponding matched controls (Figure 6.2).

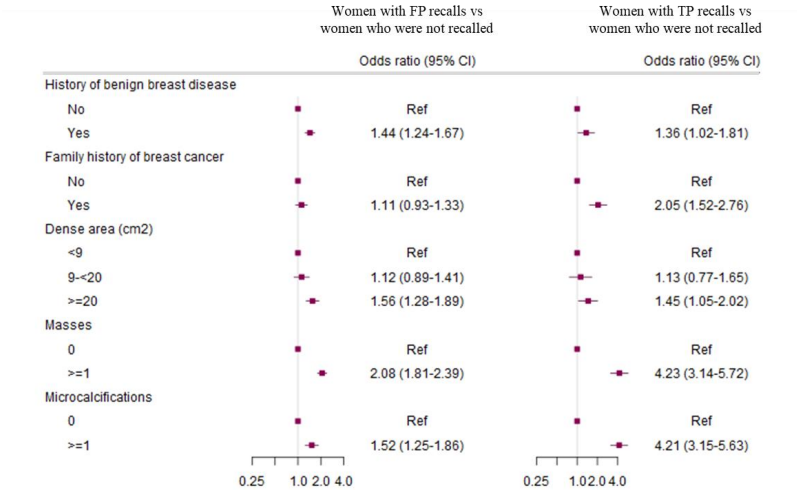

Figure 6.2 *Association between breast cancer risk factors and a FP and TP recall.* Figures were created based on estimates from Table 2 and 3 in Study I. Note: Dense areas, masses and microcalcifications of the recalled breast were used in the analyses. FP, false-positve; TP; true-positive.

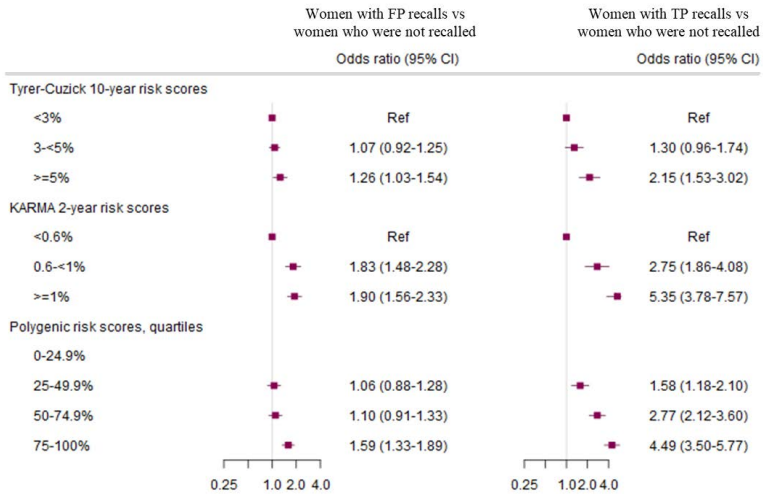

Figure 6.3 *Association between breast cancer risk prediction scores and a FP and TP recall.* Figures were created based on estimates from Table 2 and 3 in Study I. Note: For polygenic risk scores analyses, we defined false/true positive recalls as women who had received false/true positive recalls by 2015. We determined polygenic risk score quartile cutoffs based on the PRS distribution in women who had never been recalled by 2015. FP, false-positve; KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer; TP; true-positive.

Other risk factors and scores, including family history of breast cancer, microcalcifications, masses, PRS, Tyrer-Cuzick, and KARMA risk scores, were more strongly associated with a TP result than with a FP result (Figures 6.2 and 6.3).

None of the eight studied factors showed a stronger association with a FP result than with a TP result. When restricting the analysis to women who were recalled, and studying the association of these factors with an FP result while using a TP result as a control, we observed consistent results (data shown in the manuscript of Study I).

6.2 Risk of breast cancer after a false-positive recall

We observed that women with a FP result were at an increased risk of breast cancer for up to 20 years after the FP, compared to women without a FP (Figure 6.4). The association remained unchanged after adjusting for age and the calendar year of the mammograms, with a HR of 1.61 (95% CI: 1.54-1.68).

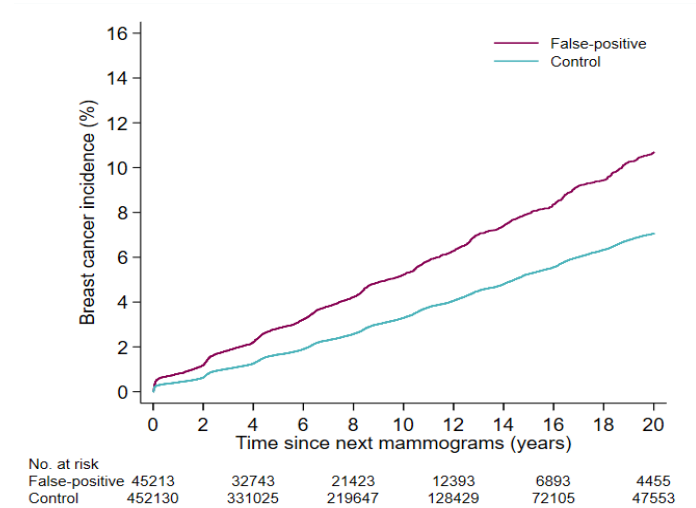

Figure 6.4 *Cumulative breast cancer incidence among women with (plum line) versus without (green line) a FP result at the index mammogram.* Adapted from Figure 1 in Study II.

We further found that the increased risk of breast cancer associated with a FP was more pronounced among women aged 60-74 (HR: 2.02; 95% CI: 1.80-2.26) than among women aged 40-49 (HR: 1.38; 95% CI: 1.23-1.56). Additionally, we observed that the elevated risk was more significant among women with lower mammographic density compared to those with higher mammographic density. The risk did not vary by calendar year, education level, or family history of breast cancer (Table 6.1).

We hypothesized that the risk of breast cancer following a FP result might vary based on the follow-up time and the side on which the cancer was detected. To investigate this, we used a flexible parametric model (Figure 6.5). We found that the risk of breast cancer being detected on the same side as the previous recall was highest at the beginning of the follow-up and decreased sharply in the initial years, gradually leveling off to a HR of around 1.5. In contrast, no increased risk was observed for breast cancer detected on the contralateral side of the FP recall at the start of the follow-up, and the risk increased and remained stable at an HR of around 1.4 for up to 20 years of follow-up (Figure 6.5).

Table 6.1 *Hazard ratio for breast cancer after a FP result, by baseline characteristics.* Adapted from Table 1 in Study II.

	No. of women	No. of cases	HR (95% CI) [a]	P –value for interaction [b]
Age (years) [c]				**<0.01**
40-49	176,231	2,486	1.38 (1.23-1.56)	
50-59	103,840	2,419	1.53 (1.36-1.72)	
60-75	74,250	2,156	2.02 (1.80-2.26)	
Calendar year [d]				0.45
1991-1999	95,755	3,424	1.70 (1.54-1.87)	
2000-2008	98,923	3,732	1.63 (1.48-1.79)	
2009-2017	113,080	2,110	1.80 (1.59-2.03)	
Family history of breast cancer				0.66
Without	353,553	12,357	1.58 (1.50-1.67)	
With	29,650	1,299	1.51 (1.27-1.81)	
Education (years)				0.45
≤9	74,101	3,007	1.68 (1.50-1.88)	
10-12	197,549	6,846	1.54 (1.43-1.66)	
>12	223,340	6,890	1.59 (1.47-1.71)	
Mammographic density [e]				**0.01**
Low (cBI-RADS A & B)	5,108	67	4.65 (2.61-8.29)	
High (cBI-RADS C & D)	7,135	108	1.60 (0.93-2.73)	

Note: [a] Estimated by stratified Cox regression models.
[b] P-values are estimated by adding an interaction term in the models.
[c] Only mammograms taken in 2005 and onwards were included. As women aged 40-49 were invited to screening from 2005 and onwards, by doing this, we made the subgroup analyses comparable.
[d] Only mammograms from women aged 50-69 were included, with a follow-up no longer than ten years. Only women aged 50-69 were invited to screening in all studied calendar years, and women with a mammogram at an earlier than later calendar year had a longer follow-up time. By doing these, we made the subgroup analyses comparable.
[e] Analyses were conducted among KARMA women with available data on mammographic density. cBI-RADS refers to computer-generated BI-RADS (Breast Imaging Reporting and Data System) density score; HR, hazard ratio.

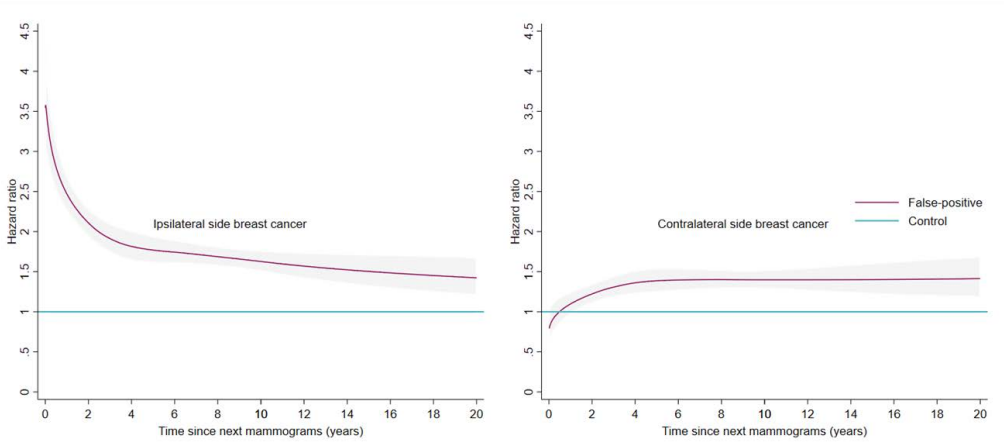

Figure 6.5 *Hazard ratio for breast cancer following a FP result, categorized by side and by follow-up time.* Estimated using flexible parametric models, with adjustments made for age and calendar year of mammogram, positive family history and education level. Adapted from Figure 3 in Study II.

6.3 Association between risk factors of women and breast cancer risk in their sisters

A total of 53,051 KARMA women and 32,198 of their sisters (identified from the Multi-generation Register, including 26,846 non-KARMA participants) were followed for their first incident of breast cancer from the time the KARMA women were enrolled in the cohort. The mean age of KARMA participants and their sisters at the time of KARMA enrollment was 54.4 (SD: 9.4) and 54.5 (SD: 10.6) years, respectively. We found that breast cancer PRS, BBD, and mammographic density were associated with an increased HR of breast cancer for both the index women and their sisters (Table 6.2). By contrast, several reproductive risk factors, as well as masses and microcalcifications of KARMA women, were not associated with an increased risk of breast cancer among their sisters (data shown in the manuscript of Study III).

Table 6.2 *Association between KARMA women's breast cancer risk factors and breast cancer risk for KARMA women and their sisters.* Adapted from Table 2&3 in Study III.

	KARMA women			Full-sibling sisters		
	No. of women	No. of cases	HR (95% CI)	No. of women	No. of cases	HR (95% CI)
Breast cancer PRS [a]						
0-9.9%	1,784	40	**0.39 (0.28-0.53)**	1,078	19	0.71 (0.44-1.14)
10-89.9%	14,268	813	Ref	8,667	210	Ref
90-100%	1,783	205	**2.01 (1.73- 2.35)**	1,057	37	**1.50 (1.06-2.13)**
Benign breast disease [b]						
No	48,894	1 154	Ref	29,610	666	Ref
Yes	4,081	184	**1.89 (1.63-2.20)**	2,538	76	**1.27 (1.01-1.60)**
Mammographic percent density [c]						
cBi-RADS A & B	20,636	524	Ref	12,262	280	Ref
cBi-RADS C	16,829	505	**1.70 (1.49-1.96)**	10,419	260	**1.42 (1.17-1.71)**
cBi-RADS D	4,886	159	**2.46 (1.99-3.05)**	3,002	65	**1.62 (1.19-2.22)**

Note: Age of KARMA women and age of full-sibling sisters were used as the time scale in the Cox regression models.
[a] 17,835 women with available genotype data and their sisters were included. Further adjusted for genotyping method.
[b] The BBD information was sourced from the Sympathy Medical System. BBD was coded as a time-varying exposure. 76 women with breast cancer diagnosed within 6 months of a BBD diagnosis were excluded.
[c] Adjusted for age and BMI of KARMA women at baseline.
HR, hazard ratio; KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer; PRS, polygenic risk score.

Sisters of women with high or moderately increased breast cancer risk, as defined by the KARMA, BOADICEA, and Tyrer-Cuzick risk models, were also observed to be associated with an increased risk of breast cancer (Table 6.3). In clinical settings, a 5-year risk of breast cancer can be used to determine whether women need preventive risk-lowing medications. Therefore, we also examined the 5-year cumulative incidence of breast cancer among sisters of KARMA women, categorized by the breast cancer risk scores of the KARMA women (Figure 6.6). Consistently, we found that the 5-year cumulative incidence rates of breast cancer were statistically significantly higher among sisters of women defined to be at high risk by the KARMA, BOADICEA, and Tyrer-Cuzick risk models. Specifically, these rates were 2.1%, 2.8%, and 2.5%, respectively, compared to 1.5%, 1.4%, and 1.4% among sisters of women classified as being at average risk by those same models.

Table 6.3 *Association between KARMA women's breast cancer risk prediction scores and breast cancer risk for KARMA women and their sisters.* Adapted from Table 4 in Study III.

	KARMA women			Full-sibling sisters		
	No. of women	No.of cases	HR (95% CI)	No. of women	No.of cases	HR (95% CI)
KARMA 2-year risk scores						
Average risk (<0.6%)	29,370	511	Ref	18,062	378	Ref
Moderate risk (0.6-<1.0%)	3,656	121	**1.70 (1.39-2.07)**	2,186	50	0.98 (0.73-1.32)
High risk (≥1.0%)	3,677	172	**2.32 (1.94-2.77)**	2,146	70	**1.35 (1.04-1.74)**
BOADICEA 5-year risk scores						
Average risk (<1.5%)	29,549	548	Ref	18,216	365	Ref
Moderate risk (1.5-<2.5%)	6,133	193	**1.48 (1.25-1.75)**	3,614	106	**1.27 (1.02-1.58)**
High risk (≥2.5%)	1,021	63	**2.98 (2.29-3.89)**	564	27	**2.04 (1.37-3.01)**
Tyrer-Cuzick 5-year risk scores						
Average risk (<1.5%)	22,803	382	Ref	14,574	277	Ref
Moderate risk (1.5-<2.5%)	9,452	217	**1.22 (1.03-1.45)**	5,557	133	1.13 (0.92-1.39)
High risk (≥2.5%)	4,448	205	**2.38 (1.99-2.83)**	2,263	88	**1.78 (1.39-2.26)**

Note: Age of KARMA women and age of full-sibling sisters were used as the time scale in the Cox regression models. BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; HR, hazard ratio; KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer.

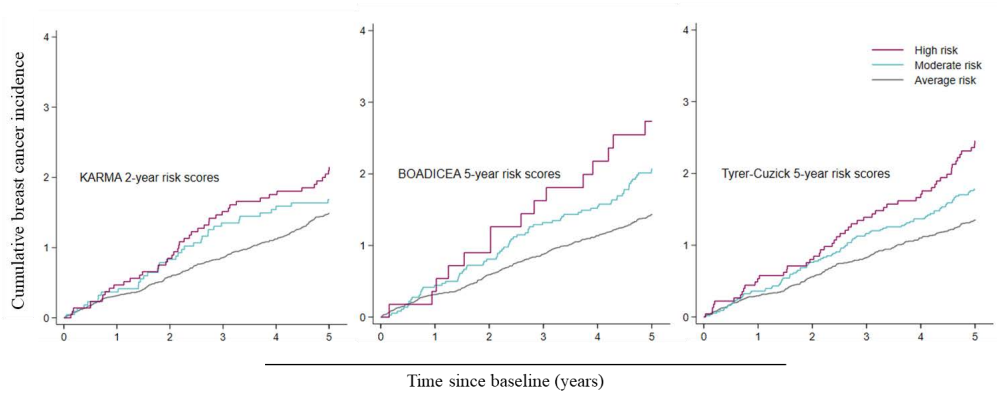

Figure 6.6 *Cumulative incidence rate of breast cancer among full-sibling sisters of KARMA women, by breast cancer risk scores of KARMA women.* Adapted from Figure 1 in Study 3. BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; KARMA, Karolinska Mammography Project for Risk Prediction of Breast Cancer.

6.4 Association between genetic predisposition of women and cancer risk in their first-degree relatives

Among a total of 13,226 index women with sequencing data, 482 women were carriers of PTV in any of the studied risk genes: *ATM, BRCA1, BRCA2, CHEK2, PALB2, BARD1, RAD51C* and *RAD51D*. Among first-degree relatives of these PTV carriers, 11.7% had breast cancer (among female relatives only) and 22.2% had cancer at any site. Among relatives of 28,362 index women with genotyping data, the percentages of both breast cancer and cancer of any type were higher for those whose index women in the top quartile of breast cancer PRS,

compared to those whose index women in the bottom quartile (7.7% vs 4.0% for breast cancer and 18.9% vs 17.0% for any cancer).

Table 6.4 *Hazard ratio (95% CI) of cancer among relatives of index women by genetic predisposition in index women.* Adapted from Table 2 & 3 in Study IV.

	Breast cancer	Non-breast HBOC-related [a]	Any HBOC-related
PTV status by the related index women			
No. cancer cases among relatives	2,094	2,836	4,942
Non-carriers [b]	1.00 (Ref)	Ref	Ref
Any risk gene [c]	**1.85 (1.52-2.27)**	**1.30 (1.06-1.59)**	**1.51 (1.31-1.75)**
Non-*BRCA* risk genes	**1.59 (1.25-2.02)**	1.16 (0.90-1.49)	**1.33 (1.11-1.59)**
BRCA1/2	**2.59 (1.84-3.63)**	**1.69 (1.23-2.32)**	**2.02 (1.60-2.54)**
PRS quartiles by the related index women [d]			
No. cancer cases among relatives	3,897	5,896	9,812
0-24.9%	1.00 (Ref)	Ref	Ref
25-49.9%	**1.26 (1.12-1.41)**	1.00 (0.92-1.09)	**1.08 (1.01-1.16)**
50-74.9%	**1.54 (1.37-1.72)**	1.08 (0.99-1.17)	**1.22 (1.14-1.30)**
75-100%	**1.85 (1.66-2.06)**	1.07 (0.99-1.16)	**1.32 (1.24-1.41)**
Per SD increase	**1.28 (1.23-1.32)**	**1.04 (1.01-1.07)**	**1.13 (1.10-1.15)**

Note: Hazard ratios were estimated via Cox regression, adjusted for birth years of relatives, cohort and breast cancer case-control status of index women at study entry.
[a] Non-breast HBOC-related cancer were defined as prostate, ovarian and pancreatic cancer, as well as melanoma.
[b] Non-carrier status was defined as the absence of PTVs in any studied risk gene, including *CHEK2*, *BRCA2*, *ATM*, *BRCA1*, *PALB2*, *BARD1*, *RAD51C* and *RAD51D*.
[c] Carrier status of any studied risk gene.
[d] PRS quartiles were defined according to breast cancer-free index women at study entry.
HBOC, Hereditary Breast and Ovarian Cancer syndrome; PRS, polygenic risk score; PTV, protein-truncating variant.

PTV carriership of any of the studied genes, of *BRCA1/2*, and of other non-*BRCA* risk genes, as well as higher PRS of index women was associated with breast cancer among female relatives (Table 6.4). Furthermore, we observed a stronger association between PTV carriership in index women and early-onset breast cancer among their relatives (diagnosed before 50 years old) compared to late-onset cases (diagnosed at 50 years old or older). In contrast, we found no significant difference in the associations between PRS quartiles and early- versus late-onset breast cancer among relatives (results detailed in the manuscript of Study IV). We also found increased risk of non-breast HBOC-related cancers (including ovarian, prostate, pancreatic cancer and melanoma) associated with index women's PTV carriership of any of the studied genes (HR: 1.30, 95% CI: 1.06-1.59), as well as a higher PRS ($HR_{per\ SD}$: 1.04, 95% CI: 1.01-1.07) (Table 6.4).

Furthermore, we examined the lifetime risk of breast cancer and non-breast HBOC-related cancer among relatives, categorizing them by the PTV status in any of the studied genes or the PRS quartile of their index women (Figure 6.7). Relatives of women with PTV in any risk gene had statistically higher lifetime risk of breast cancer or non-breast HBOC cancers than relatives

of non-carriers. We observed differences in the lifetime risk of breast cancer among relatives based on the PRS quartiles of their index women, but not in the lifetime risk of non-breast HBOC-related cancers (Figure 6.7).

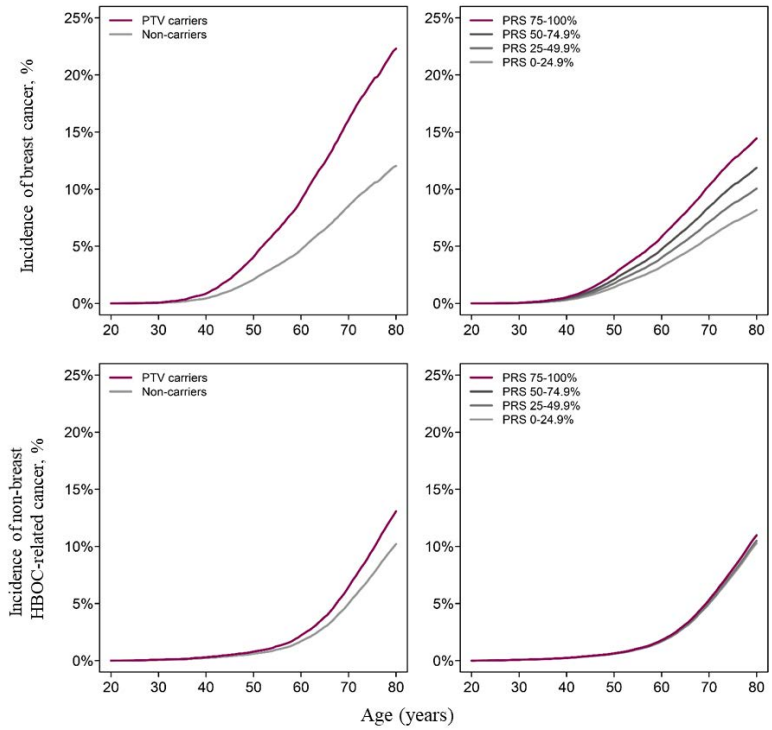

Figure 6.7 *Adjusted cumulative incidence of breast and non-breast HBOC-related cancers among relatives of index women by genetic predisposition of index women.* Adapted from Figure 2 in Study IV. The cumulative incidence was estimated using standardized Cox regression models, with adjustments for the breast cancer case-control status of their index women at the KARMA enrollment. PTV carriers: carrying mutations in any of the eight studied genes. Non-carriers: women without PTVs in any of the eight studied genes. PRS quartiles were defined among index women who were free of breast cancer at the KARMA enrollment. HBOC, Hereditary Breast and Ovarian Cancer syndrome; PRS, polygenic risk score; PTV, protein-truncating variant.

7 Discussion

7.1 Interpretation of main findings

7.1.1 Factors associated with a false-positive recall (Study I)

In Study I, we identified several breast cancer risk factors and risk scores that were associated with a FP recall. However, these factors were equally or more strongly associated with a TP recall. Therefore, they are not suitable for development as targeted interventions to minimize the rate of FP mammography recalls, since such interventions could simultaneously decrease the rate of TPs.

For women aged 40-49 years, the ratio of FP per TP (9.78) was much higher than the ratio for women aged 60-74 years (at 1.78). Our findings suggest that younger women may generally experience a more unfavorable harm-benefit ratio when attending mammography screenings. Tailored breast cancer risk-based screening may help balance the harm-benefit ratio in this age group in the future.

Mammographic features are the key factors for radiologists in making accurate decisions about whether or not to recall a woman for further examination. High mammographic density has been consistently shown to be a risk factor for breast cancer and women with high density are more likely to have a FP recall [105, 207]. Our results confirmed previous studies and showed that density was associated with both FP and TP recalls at a similar magnitude. Suspicious masses and microcalcifications, identified by iCAD, an FDA-approved software, were designed to assist radiologists in detecting breast cancer on mammograms [160]. Consistent with previous literature [208, 209], although these features were associated with a FP result, we found that they were more strongly associated with a TP result. Furthermore, the KARMA and Tyrer-Cuzick breast cancer risk models were associated with FP recalls but were more likely to be associated with TP recalls. The association between high risk scores and a FP recall can probably be explained by the fact that these models directly or indirectly incorporate mammographic features [117, 122, 210].

7.1.2 Risk of breast cancer after a FP (Study II)

It has consistently been shown that women with a FP result are at an increased risk of breast cancer in subsequent years [15, 126, 139, 211]. However, Study II was the first, to our knowledge, to not only confirm a long-term risk of breast cancer but also to examine how it varied based on age, mammographic density, and time since follow-up.

We were the first to report that the risk of breast cancer associated with a FP recall was statistically significantly higher among women with lower mammographic density compared to those with higher density. This may be because women with low and high density are recalled for different reasons. We hypothesized that women with low density were more likely to be recalled due to suspicious findings, while women with high density were more likely to be recalled because of the masking effect [207]. Since mammographic density is negatively

associated with age [105], our results were consistent in showing that the increased risk of breast cancer was more pronounced among older than younger women.

Our study was the first to observe that the increased risk of breast cancer on the ipsilateral side was highest in the initial years of follow-up and decreased over time. Meanwhile, we noted a similar elevated risk of breast cancer on both the ipsilateral and contralateral sides among women with a FP result. The short-term elevated risk on the ipsilateral side could possibly be attributed to tumors being missed during subsequent examinations at the recall [138, 212] or to BBD [127, 138]. Our findings suggest that a short-term surveillance program might benefit these women. Additionally, since these women have an elevated long-term risk of breast cancer, they may benefit more from an extended screening program compared to women without a FP recall.

Furthermore, we observed that women with a FP result were also at an increased risk of breast cancer-specific mortality. However, the prognosis for breast cancer among women with or without a FP recall did not differ (HR: 1.05; 95% CI: 0.89-1.25). Thus, the increased breast cancer mortality is likely driven by the higher incidence of breast cancer among women with a FP recall, rather than by differences in tumor characteristics. Consistently, we found that a FP recall was not associated with any specific tumor characteristics, except for larger tumor size.

7.1.3 Association between risk factors of women and breast cancer risk in their sisters (Study III)

Our findings suggest that mammographic density, a history of BBD, and BC-PRS, as well as breast cancer risk prediction scores, can be useful in estimating breast cancer risk for the sisters of women being assessed. In contrast, other factors of women such as reproductive and hormonal aspects, lifestyle factors, as well as masses and microcalcifications, were not associated with increased risk for their sisters. The association between a woman's specific risk factor and her sister's risk of breast cancer could probably be attributed to two things: the intraclass correlation of the risk factor between sisters and the strength of that factor's association with breast cancer. For instance, the stronger a breast cancer risk factor is and the higher its correlation between sisters, the more pronounced its association with the sister's risk of developing the disease (results detailed in the manuscript of Study III).

The positive association between an index woman's mammographic density and the breast cancer risk in her sisters suggests that informing women with high mammographic density at screening could potentially benefit not only the women themselves but also their sisters. This is important because examination centers in the U.S. already inform all women about their mammographic density [153, 213]. Health care agencies in European countries are considering including information about women's breast density in future mammography reports [154].

Primary care doctors in the U.S. commonly assess women's breast cancer risk using risk models, such as the Tyrer-Cuzick and Gail models, to recommend screening strategies and possible reimbursement options [214]. In this study, we demonstrated effective stratification of breast cancer incidence among the sisters of the index women, based on risk scores from three

different models. Consequently, when a woman is classified as high-risk according to these models during a clinic visit, sharing this information with her sisters could enhance their disease awareness and motivate them to undergo screening as well. However, ethical considerations regarding sharing risk assessments with women who have not requested them must be carefully addressed.

7.1.4 Association between genetic predisposition of women and cancer risk in first-degree relatives (Study IV)

In Study IV, we found that the presence of PTVs in any of the 8 studied risk genes, which are strong risk factors for breast cancer [59], or in *BRCA1/2* alone, was associated with an increased risk of breast cancer among the female relatives of the index women. We observed that this increased risk was significantly higher for early-onset breast cancer compared to late-onset. This observation aligns with previous findings that the risk of early-onset breast cancer associated with PTV in these rare genes is statistically stronger than the risk for late-onset breast cancer in the women themselves [215].

Consistent with the results from Study III of this thesis, we observed that the breast cancer PRS of the women was associated with an elevated breast cancer risk in their first-degree female relatives. This increased risk was largely consistent when we restricted our analysis to only mothers or to sisters and offspring of the index women. Additionally, when we further assessed whether the breast cancer risk differed between early- and late-onset cancer, we did not observe a statistical difference.

We also observed that relatives of women with higher breast cancer PRS had a slightly increased risk of non-breast HBOC-related cancers. Incorporating SNPs associated with other cancers, or considering breast cancer PRS in combination with other factors, may provide more informative results.

We did not find sufficient evidence to suggest that PTV carriership of non-*BRCA* risk genes was associated with non-breast HBOC-related cancer (HR: 1.16; 95%: 0.90-1.49). Upon further investigation of cancer aggregation within families due to PTV carriership of individual genes, we discovered that *PALB2* was associated with multiple types of cancer within families (results shown in the supplements in Study IV). Our findings indicated that these genes may potentially play a role in the cancer development of multiple cancers and might be useful in genetic counseling.

7.2 Methodological consideration

7.2.1 Selection bias

Selection bias, a systematic error, can significantly affect the validity of observational studies by causing the study population to not accurately mirror the source population. Consequently, the conclusions derived from the study population may not be applicable to the source population. In Study II, the main study population was drawn from the screening register, encompassing all screening records for every woman of suitable age in the Stockholm-Gotland

region. Every woman with the exposure (a FP recall) was included. Thus, the study is not prone to selection bias.

A common form of selection bias is the "self-selection bias". Although invitations to participate in a cohort study are typically sent to everyone in a specific area, those who choose to participate are often highly educated and may have a family history of the disease under investigation. This bias is evident in the KARMA cohort. Invitations to attend the KARMA study were sent to 210,233 women who underwent mammography screening or clinical mammography between 2011 and 2013; of these, 70,877 (34%) consented to participate. KARMA participants were more likely to participate in screening and have higher educational levels and a family history of breast cancer compared to the general female population in Sweden. In Studies I and III, the study populations included KARMA participants (in both studies) and their relatives (in Study III). Despite this, the specific research questions in these studies focused on the screening population, which tends to have higher education and a family history of breast cancer compared to non-participants. Therefore, the findings from these two studies remain relevant to those who attend mammography screenings.

In Study IV, one of the main focuses was the association between rare genetic mutations and the risk of cancer among women's relatives. Previous studies on these rare mutations often included breast cancer cases with a family history to improve the power to observe an association, likely introducing selection bias. We avoided such selection in our study. We included breast cancer cases (irrespective of a positive family history) and controls from both Libro-1 and KARMA, along with all first-degree relatives identified in the Multi-generation Register. However, it should be noted the Libro-1 cohort is composed of breast cancer patients diagnosed between 2001 and 2008 who were alive in 2009, and the KARMA cohort includes survivors diagnosed before enrollment. To address this, we presented the cancer risk among relatives based on the genetic predisposition of index women, for both breast cancer patients and individuals who were breast-cancer-free at the time of enrollment in KARMA. We observed similar results based on breast cancer case-control status.

7.2.2 Misclassification bias and recall bias

In this thesis, the four studies primarily used data sourced from various Swedish national registers to define exposures, covariates, and outcomes. Specifically, screening outcomes were sourced from the Stockholm mammography screening register, cancer diagnoses and tumor characteristics from the Swedish Cancer Register and Breast Cancer Quality Register, and data on death and migration from the Cause of Death and Migration registers. As mentioned in the previous section, since Swedish registers are generally considered to be of high quality, the likelihood of having misclassified variables from these sources is minimal. Additionally, variables from these registers were recorded independently of the case-control or exposed-unexposed status. Therefore, even if misclassification does occur, it would likely be non-differential, not significantly altering our conclusions.

Mammographic features, including mammographic density, masses, and microcalcifications, were automatically measured by Stratus or iCAD software for each mammogram. Both common and rare genetic predispositions were derived from genotype and sequencing data. Like variables defined by register data, any misclassification of these variables would likely be random and non-differential.

Studies I and III incorporated data on common breast cancer risk factors, such as hormonal and reproductive factors, which were based on questionnaire data. Consequently, potential misclassification and recall bias should be considered. However, two key points substantially mitigate concerns about these biases. First, women diagnosed with breast cancer before enrolling in KARMA were excluded from both studies. Second, women in the KARMA cohort were not aware of our study objectives—namely, to examine the association between those factors and a FP result, as well as the risk of breast cancer among their sisters. Therefore, the introduction of recall or misclassification biases was unlikely, either due to breast cancer case-control status at baseline or due to participants' awareness of the research aims.

7.2.3 Generalizability

Generalizability refers to the extent to which the results found in studies can be applied to other populations. Regarding findings on FP results from Studies I and II, cautious interpretation is advised. This is because the ages targeted by screening programs, screening intervals, and age distribution of screening populations may vary among countries [124, 131]. Therefore, our findings should be validated within specific populations before implementation. However, several common factors were found to be associated with FPs based on data from different countries [207, 216]. Additionally, an elevated risk of breast cancer following an FP is consistently observed in multiple countries [139, 217]. Hence, it is likely that our findings can be generalized to other developed countries. For findings from Study III, the magnitude of the association between a woman's breast cancer risk factors and the breast cancer risk in her sisters may vary depending on age distribution and prevalence of risk factors. For Study IV, variables such as the incidence rate of cancers and the ancestry of women in other populations might yield variations in the magnitudes of associations.

7.3 Ethical consideration

Ethics are important not only in interventional research but also in observational research. In this thesis, all the constituent studies were observational and used data from national registers or population-based cohorts. Thus, from an ethical perspective, there are two main concerns: obtaining ethical approval and maintaining the privacy and integrity of the data used.

Data related to disease history, genetic mutations, and family histories of specific diseases are all considered sensitive. To manage this sensitive data, we obtained ethical approval for the included studies before collecting data and performing analyses. The Regional Ethical Review Board in Stockholm approved all projects which are included in this thesis. The ethical approvals related to this thesis are as follows: Dnr 2019-04369; Dnr 2010/958-31/1 with extension Dnr 2013/2090-32, and Dnr 2009/254-31/4, with extensions 2011/2010-32 and

2012/465-32. For the register-based study, the requirement for informed consent from the individuals included in the register was waived after obtaining approval from the ethical review board. All participants from the KARMA and Libro-1 cohorts have provided informed consent, and they can withdraw from the study at any time if they want.

All analyses included in the thesis were performed under a data protection framework to avoid disclosing personal data. No individuals were identified during data management or analysis, and no researchers not involved in the projects had access to the data.

Besides the two ethical concerns commonly associated with large epidemiological studies, findings from Studies III and IV highlight a potential ethical issue that should be addressed in the era of precision medicine. Our results suggest that a woman's risk factors may be useful for predicting cancer risk in her relatives. This information might encourage relatives to attend breast cancer screenings or prompt relatives to seek genetic counseling. However, sharing this information with sisters or other relatives may violate their personal integrity. As the implementation of personalized risk-based screening programs is foreseeable, this poses an ethical challenge that healthcare agencies and governments need to discuss in detail and address carefully.

8 Conclusions

There are three main conclusions emanating from the studies included in the thesis:

- We identified several breast cancer risk factors and risk scores that were associated with FP mammography recalls—a major concern in mammography screening. However, none of the studied breast cancer risk factors or risk scores could be used as potential interventions to minimize FP recalls, as such interventions could also risk missing true tumors. Although we did not identify factors that could reduce FP recall rates, we observed that the increased breast cancer risk following an FP recall could persist for over 20 years, was higher in older women, and highest in the first few years of follow-up. These findings can be used to develop personalized surveillance programs for women with a FP recall (Studies I & II).

- Based on a large cohort from a screening population, we found that mammographic density, breast cancer PRS, and a history of BBD, as well as breast cancer risk prediction scores from the KARMA, BOADICEA, and Tyrer-Cuzick models, were also associated with breast cancer risk in the sisters of these women. Our findings suggest that risk assessments performed for women during screening can be useful for estimating their sisters' risk of breast cancer. While this may increase disease awareness among sisters at higher risk, thorough discussions are required, particularly concerning potential ethical issues (Study III).

- We found that relatives of women with higher breast cancer PRS or PTVs in any of the eight examined risk genes were at increased risk for breast and other HBOC-related cancers at a population level. Additionally, our results highlight the potential of non-*BRCA1/2* genetic predispositions, such as *PALB2*, to provide valuable insights into the aggregation of cancer in families (Study IV).

9 Points of perspective

With four studies in the thesis conducted with the aim of contributing to improved cancer prevention, here are some perspectives on the related topics and what should be addressed to improve cancer prevention in the near future.

9.1 What should we do about FP mammography recalls in screening?

We did not find breast cancer risk factors as potential targets to be used to reduce FP mammography recalls. However, this does not mean the adverse issue of FP mammography recalls in screening cannot be addressed. First, FP recall rates might be reduced through using AI-based screening programs [218, 219]. Results from a clinical trial that included AI assistance for the radiologist—instead of having two radiologists read mammograms independently—showed improved breast cancer detection rates and fewer FP recalls [219]. Second, since FP recalls are often more frequent in women with dense breasts, other modalities with high specificity, such as digital breast tomosynthesis and contrast-enhanced spectral mammography, may benefit these women [162].

While current mammography screening program may continue to be used across countries for many years to come, our findings highlight potential surveillance strategies that could benefit women with a FP mammography recall. Future studies could test whether short-term, intensive imaging examinations can help detect cancers at an earlier stage, and whether educational programs could encourage women to maintain long-term awareness of the disease. Given the long-term breast cancer risk for women with a FP recall and the generally longer life expectancy for women in countries with screening programs, extending the screening program beyond the current age limit (74 years in Sweden) might be beneficial for these women.

9.2 An era of risk-based screening?

Regarding breast cancer screening, one of the trending research topics is risk-based screening. Current breast cancer screening methods are age-based and adopt a one-size-fits-all approach, meaning all women of specific ages are invited for screening at same intervals. For example, in Stockholm, all women aged 40-74 are invited to screening every 24 months. A major concern with this age-based approach is that the risk of breast cancer varies among women of the same age, given that genetic and environmental factors differ from person to person. Thus, many researchers believe that risk-based screening, guided by breast cancer prediction scores, could lead to better screening outcomes. Women with moderate or high risk could be screened more frequently than those at lower risk. Findings from our Study III can be applied in this scenario: when women receive information about their own risk, the sisters of those women at high risk may be motivated to also participate in screening and become more aware of the disease.

9.3 Other future perspectives

While screening programs and breast cancer prognosis are generally good, future studies are needed to further improve cancer prevention and so to improve the cancer prognosis.

- For women with moderately or highly increased breast cancer risk, risk-lowering medication could be offered, in addition to more frequent screening [220]. However, although risk-lowering medication is recommended, it is not widely used in the U.S. due to the side effects associated with tamoxifen [221]. Our group is conducting clinical trials to test whether a small dose of tamoxifen is effective in reducing breast cancer risk while also substantially limiting side effects. So far, using reduced mammographic density as a proxy, the results have shown that a lower dose of tamoxifen (2.5 mg) reduces mammographic density as effectively as the standard dose (20 mg), while severe side effects were reduced by 50% [222].

- Women with screen-detected breast cancer are considered the beneficiaries of the screening program. Compared to women with non-screen-detected cancer, these women tend to have smaller tumor sizes, lower grades, and better tumor characteristics. However, not all women participate each time they are invited to screening. We hypothesize that among women with screen-detected cancer, those who previously participated in the screening are likely to have better tumor characteristics. We plan to conduct analyses to study whether tumor characteristics or prognosis differ based on screening history.

- The COVID-19 pandemic has significantly impacted breast cancer screening programs. All cities in Sweden have reported substantially reduced screening attendance. This situation can be considered a 'natural experiment.' It would be interesting to examine the tumor characteristics of screen-detected cancers diagnosed during or shortly after the pandemic, as well as the prognosis for these women. Comparing these data with the tumor characteristics and prognosis of patients diagnosed just before the pandemic may provide the most direct and current evidence regarding the extent to which mammography screening contributes to early breast cancer detection and reduced mortality.

10 Acknowledgements

This section is dedicated to those who have guided, supported, and accompanied me throughout this rewarding PhD journey.

To **Kamila Czene**, my main supervisor. Your intelligence, passion for breast cancer research, and exceptional presentation skills continually impress me. What inspires me the most is your extraordinary problem-solving ability. No matter what problems I have faced, a conversation with you has invariably pointed me toward a solution that far exceeded my initial expectations. I am, and will always be, grateful that you gave me the chance to pursue this PhD. Thank you for all the constructive suggestions which have been pivotal to my progress, for motivating me to excel, and for guiding me toward becoming an independent researcher.

To **Wei He**, my co-supervisor. Thank you for your consistent kindness and consideration, for always finding time to help me, and for your patience in guiding me along the way. I have lost count how many times we have had the following conversation over the years: *Me: 'Do you have time now, or would later be more convenient for me to ask a question?' You: 'I can help you now.'* I am continually amazed by your ideas in research, your insights into how to write clearly and engagingly, and your expertise in delivering innovative presentations.

To **Keith Humphreys**, my co-supervisor. You are incredibly intelligent, yet so humble. Thank you for your patient guidance in helping me solve statistical problems and for your invaluable suggestions on many random and general questions I have had during my PhD journey. I always feel that I can rely on your help when I need it the most. I have always enjoyed our talks about life, culture, books, and movies. Thank you for sharing your thoughtful perspectives on various matters.

To **Haomin Yang**, my co-supervisor. You are a role model to me for your impressive productivity. Thank you for your invaluable suggestions and for your consistent willingness to help during these years. You are also like a 'big academic brother,' and I appreciate that you genuinely shared your own experiences from your PhD journey.

To the co-authors of the four studies included in this thesis, listed in alphabetical order: **Alexander Ploner, Douglas Easton, Felix Grassmann, Fredrik Strand, José Tapia, Juan Rodriguez, Linda S. Lindström, Mattias Hammarström, Mikael Eriksson, Natalie Holowko, Per Hall, Qingyang Xiao,** and **Svetlana Bajalica-Lagercrantz**—thank you all for your timely feedback and suggestions that improved the manuscripts and facilitated the completion of revisions. Special thanks to **Per,** for consistently sharing valuable insights, clinical perspectives, and constructive suggestions that have made the manuscripts more interesting to read and enhanced their clinical relevance. To **Natalie**, in addition to your incredible input on the manuscripts, I appreciate your patient and pedagogical guidance on academic writing. Your warm encouragement and congratulations on every milestone I reached were invaluable to me. I also cherish our conversations over lunch and fika. To **Alexander**, you have an incredible knack for resolving statistical issues and explaining them clearly.

Discussions with you always leave me in a good mood. To **Felix**, thank you for providing sharp and quick feedback, as well as your pedagogical teachings on how to conduct genetic analyses. To **José**, I appreciate your rigorous work ethic and your friendly demeanor. I have also greatly enjoyed our conversations about life.

To **Yi Lu**, my mentor. I enjoyed all the conversations we had. Thank you for being warm and encouraging! Your passion for research inspires me.

To past and present members from Kamila's group who have not been mentioned: **Erwei Zeng, Pui San Tan, Yuqi Zhang, Ziyan Ma, Damien Kaukonen, Ami Rönnberg, Emilio Ugalde Morales,** and **Qiaoli Wang.** Working with you all has been a true delight, and I am grateful for the support from you. Special thanks to **Erwei**, I am so grateful for your companionship. I have enjoyed our countless conversations, playing games together, and our trip to Berlin. To **Pui San**, thank you for being so kind and supportive all these years. I cherish the time having lunch and hanging out together. I feel lucky that our friendship continues. To **Yuqi** and **Ziyan**, thank you for always being positive and supportive, and for all the fun times we have had together. To **Ami**, for being friendly, encouraging, and for always giving timely support.

To **the great teaching group at MEB**, I enjoyed every course I took from our department. Thank you all for teaching epidemiology, biostatistics, and coding languages in such a clear and interesting way. Special thanks to **Arvid Sjölander** and **Mark Clements** for giving me the opportunity to assist in teaching for the courses Epidemiology II and Biostats III. I would also like to thank **Mark** for always being encouraging, kind, and willing to help.

The spirit of MEB is a result of the collective efforts of all MEBers. I sincerely thank **all the past and current MEBers** for creating a work environment that is not only great, open but also a place where support is always readily available. Special thanks to my past and current **officemates**. Thanks to my peers in the **PhD Board group** and **PubMEB group**. It was absolute fun time to work together and organize all the events! To **Alessandra Nanni**, thank you for all the patience, time, and support you have given me over the years. It is always lovely talking to you.

Many thanks go to all my friends in Sweden, China, and around the world. I feel incredibly fortunate to have you all in my life, cherishing our happy moments, our mutual sharing and listening, as well as years of companionship. Special thanks to a selected few who have been particularly supportive during my PhD journey. To **Shuang Hao**, my colleague, friend, sister, and 'family': you are smart and self-disciplined, yet also open, genuine, and fun. Thank you for your tremendous support, care, and invaluable suggestions regarding both my work and life. I appreciate our gatherings and fun times. To **Philippe Weitz**, without our friendship, my PhD experience would have lacked so much fun time. You are also consistently reliable, and you serve as my role model for work-life balance. To my '**Lappis family**', Vanhery, William, Rintaro, and Abhimanyu, thank you for the movie nights, dinner gatherings, and other activities we enjoyed, especially at the beginning of the pandemic. Our continued friendship means a lot to me. Special thanks to Vanhery and William for always being there for me.

To my beloved **dad** and **mom**, thank you for all the love and support you continually give me, for the care and understanding you show, and for granting me the freedom to be who I want to be. My deepest thanks go to my mom. Your unconditional love and faith in me are the greatest treasures that I possess in this world. These form the core of my faith in myself. It is this foundation that continually gives me the courage to explore unknown worlds and sustains me mentally during challenging times in my life.

To **Song**, my boyfriend. Thank you for your unwavering support and caring, for continually encouraging me to pursue my own path. I feel so lucky to 'grow up' together with you in the adult world.

"I shall be telling this with a sigh

Somewhere ages and ages hence:

Two roads diverged in a wood, and I—

I took the one less traveled by,

And that has made all the difference."

--Robert Frost

Many years from now, I imagine my feelings about the decision to step into this major and pursue this Ph.D. will mirror the sentiments expressed in Robert Frost's poem 'The Road Not Taken.' Yet, I will have no regrets, thanks to the enriching experiences and the incredible people I have met along this PhD journey. To all the friends and colleagues I have mentioned in this thesis and to the many more I have not named: Thank you! ♡

11 References

1. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. **71**(3): p. 209-249.

2. Parkin, D.M., J. Stjernswärd, and C.S. Muir, *Estimates of the worldwide frequency of twelve major cancers.* Bull World Health Organ, 1984. **62**(2): p. 163-82.

3. Parkin, D.M., et al., *Global cancer statistics, 2002.* CA Cancer J Clin, 2005. **55**(2): p. 74-108.

4. Kolak, A., et al., *Primary and secondary prevention of breast cancer.* Ann Agric Environ Med, 2017. **24**(4): p. 549-553.

5. Colditz, G.A. and K. Bohlke, *Priorities for the primary prevention of breast cancer.* CA Cancer J Clin, 2014. **64**(3): p. 186-94.

6. Løberg, M., et al., *Benefits and harms of mammography screening.* Breast Cancer Research, 2015. **17**(1): p. 63.

7. Coldman, A., et al., *Pan-Canadian study of mammography screening and mortality from breast cancer.* J Natl Cancer Inst, 2014. **106**(11).

8. Tabar, L., et al., *Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening.* Lancet, 2003. **361**(9367): p. 1405-10.

9. Mandelblatt, J.S., et al., *Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies.* Ann Intern Med, 2016. **164**(4): p. 215-25.

10. Tsuruda, K.M., et al., *Cumulative risk of a false-positive screening result: A retrospective cohort study using empirical data from 10 biennial screening rounds in BreastScreen Norway.* Cancer, 2022. **128**(7): p. 1373-1380.

11. Singh, D., et al., *Cumulative risk of false positive test in relation to breast symptoms in mammography screening: a historical prospective cohort study.* Breast Cancer Res Treat, 2016. **159**(2): p. 305-13.

12. Hubbard, R.A., et al., *Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study.* Ann Intern Med, 2011. **155**(8): p. 481-92.

13. Brewer, N.T., T. Salz, and S.E. Lillie, *Systematic review: the long-term effects of false-positive mammograms.* Annals of internal medicine, 2007. **146**(7): p. 502-510.

14. Brett, J., et al., *The psychological impact of mammographic screening. A systematic review.* Psychooncology, 2005. **14**(11): p. 917-38.

15. McCann, J., D. Stockton, and S. Godward, *Impact of false-positive mammography on subsequent screening attendance and risk of cancer.* Breast Cancer Res, 2002. **4**(5): p. R11.

16. Sun, Y.-S., et al., *Risk factors and preventions of breast cancer.* International journal of biological sciences, 2017. **13**(11): p. 1387.

17. McPherson, K., C. Steel, and J. Dixon, *Breast cancer—epidemiology, risk factors, and genetics.* Bmj, 2000. **321**(7261): p. 624-628.

18. Yala, A., et al., *A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction.* Radiology, 2019. **292**(1): p. 60-66.

19. Lee, A., et al., *BOADICEA: a comprehensive breast cancer risk prediction modelincorporating genetic and nongenetic risk factors.* Genetics in Medicine, 2019. **21**(8): p. 1708-1718.

20. Gail, M.H. and R.M. Pfeiffer, *Breast Cancer Risk Model Requirements for Counseling, Prevention, and Screening.* J Natl Cancer Inst, 2018. **110**(9): p. 994-1002.

21. Eriksson, M., et al., *A risk model for digital breast tomosynthesis to predict breast cancer and guide clinical care.* Sci Transl Med, 2022. **14**(644): p. eabn3971.

22. Yoshida, R., *Hereditary breast and ovarian cancer (HBOC): review of its molecular characteristics, screening, treatment, and prognosis.* Breast Cancer, 2021. **28**(6): p. 1167-1180.

23. Choi, J., et al., *Associations of genetic susceptibility to 16 cancers with risk of breast cancer overall and by intrinsic subtypes.* Human Genetics and Genomics Advances, 2022. **3**(1).

24. Torre, L.A., et al., *Global cancer statistics, 2012.* CA Cancer J Clin, 2015. **65**(2): p. 87-108.

25. Harbeck, N. and M. Gnant, *Breast cancer.* Lancet, 2017. **389**(10074): p. 1134-1150.

26. Howlader, N., et al., *SEER cancer statistics review, 1975–2016.* National Cancer Institute, 2019. **1**.

27. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.

28. DeSantis, C.E., et al., *Breast cancer statistics, 2019.* CA Cancer J Clin, 2019. **69**(6): p. 438-451.

29. Socialstyrelsen. *Statistics on Cancer Incidence 2021*. 2022 [cited 2023 August 3rd]; Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/statistik/2022-12-8309.pdf.

30. Cancerfonden. *Breast cancer statistics*. 2022 [cited 2023 August 3rd]; Available from: https://www.cancerfonden.se/om-cancer/statistik/brostcancer.

31. Nordcan. *Age-Standardized Rate (European) per 100 000 , Incidence and Mortality, Females*. 2022 [cited 2023 August 3rd]; Available from: https://nordcan.iarc.fr/en/dataviz/trends?cancers=180&sexes=2&populations=752&types=0_1&years=1960_2020&num_path=4&years_available=1943_2020&key=asr_e.

32. Pharoah, P.D., et al., *Family history and the risk of breast cancer: a systematic review and meta-analysis.* Int J Cancer, 1997. **71**(5): p. 800-9.

33. Sattin, R.W., et al., *Family history and the risk of breast cancer.* Jama, 1985. **253**(13): p. 1908-1913.

34. Cancer, C.G.o.H.F.i.B., *Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease.* Lancet, 2001. **358**(9291): p. 1389-99.

35. Colditz, G.A., et al., *Family history and risk of breast cancer: nurses' health study.* Breast Cancer Res Treat, 2012. **133**(3): p. 1097-104.

36. Hemminki, K., et al., *Familial Risks and Proportions Describing Population Landscape of Familial Cancer.* Cancers (Basel), 2021. **13**(17).

37. Stratton, M.R. and N. Rahman, *The emerging landscape of breast cancer susceptibility.* Nat Genet, 2008. **40**(1): p. 17-22.

38. Peto, J. and T.M. Mack, *High constant incidence in twins and other relatives of women with breast cancer.* Nat Genet, 2000. **26**(4): p. 411-4.

39. Couto, E. and K. Hemminki, *Estimates of heritable and environmental components of familial breast cancer using family history information.* Br J Cancer, 2007. **96**(11): p. 1740-2.

40. Mucci, L.A., et al., *Familial Risk and Heritability of Cancer Among Twins in Nordic Countries.* Jama, 2016. **315**(1): p. 68-76.

41. Czene, K., P. Lichtenstein, and K. Hemminki, *Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database.* Int J Cancer, 2002. **99**(2): p. 260-6.

42. Möller, S., et al., *The Heritability of Breast Cancer among Women in the Nordic Twin Study of Cancer.* Cancer Epidemiol Biomarkers Prev, 2016. **25**(1): p. 145-50.

43. Woodward, E.R., E.M. van Veen, and D.G. Evans, *From BRCA1 to Polygenic Risk Scores: Mutation-Associated Risks in Breast Cancer-Related Genes.* Breast Care (Basel), 2021. **16**(3): p. 202-213.

44. Apostolou, P. and F. Fostira, *Hereditary Breast Cancer: The Era of New Susceptibility Genes.* BioMed Research International, 2013. **2013**: p. 747318.

45. American Cancer society. *Breast Cancer Risk Factors You Cannot Change*. 2021 [cited 2023 September 13]; Available from: https://www.cancer.org/cancer/types/breast-cancer/risk-and-prevention/breast-cancer-risk-factors-you-cannot-change.html.

46. Garber, J.E. and K. Offit, *Hereditary cancer predisposition syndromes.* Journal of clinical oncology, 2005. **23**(2): p. 276-292.

47. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.

48. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21.* Science, 1990. **250**(4988): p. 1684-1689.

49. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. **265**(5181): p. 2088-2090.

50. Shiovitz, S. and L.A. Korde, *Genetics of breast cancer: a topic in evolution.* Ann Oncol, 2015. **26**(7): p. 1291-9.

51. King, M.C., J.H. Marks, and J.B. Mandell, *Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.* Science, 2003. **302**(5645): p. 643-6.

52. Tan, M.H., et al., *Lifetime cancer risks in individuals with germline PTEN mutations.* Clin Cancer Res, 2012. **18**(2): p. 400-7.

53. Pharoah, P.D., P. Guilford, and C. Caldas, *Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families.* Gastroenterology, 2001. **121**(6): p. 1348-53.

54. Walsh, T., et al., *Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer.* Jama, 2006. **295**(12): p. 1379-88.

55. Seal, S., et al., *Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles.* Nat Genet, 2006. **38**(11): p. 1239-41.

56. Thompson, D., et al., *Cancer risks and mortality in heterozygous ATM mutation carriers.* J Natl Cancer Inst, 2005. **97**(11): p. 813-22.

57. Rahman, N., et al., *PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene.* Nat Genet, 2007. **39**(2): p. 165-7.

58. Gupta, S., et al., *Can selenium be a modifier of cancer risk in CHEK2 mutation carriers?* Mutagenesis, 2013. **28**(6): p. 625-9.

59. Dorling, L., et al., *Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women.* N Engl J Med, 2021. **384**(5): p. 428-439.

60. Petrucelli, N., M.B. Daly, and T. Pal, *BRCA1-and BRCA2-associated hereditary breast and ovarian cancer.* 2022.

61. Pinto, E.M., et al., *TP53-Associated Pediatric Malignancies.* Genes Cancer, 2011. **2**(4): p. 485-90.

62. Michailidou, K., et al., *Large-scale genotyping identifies 41 new loci associated with breast cancer risk.* Nat Genet, 2013. **45**(4): p. 353-61, 361e1-2.

63. Michailidou, K., et al., *Association analysis identifies 65 new breast cancer risk loci.* Nature, 2017. **551**(7678): p. 92-94.

64. Zhang, H., et al., *Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses.* Nat Genet, 2020. **52**(6): p. 572-581.

65. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**(7148): p. 1087-93.

66. Choi, S.W., T.S.H. Mak, and P. O'Reilly, *A guide to performing Polygenic Risk Score analyses.* 2018.

67. Socrates, A., et al., *Polygenic risk scores applied to a single cohort reveal pleiotropy among hundreds of human phenotypes.* bioRxiv, 2017: p. 203257.

68. 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.

69. Hung, R.J., et al., *Cross Cancer Genomic Investigation of Inflammation Pathway for Five Common Cancers: Lung, Ovary, Prostate, Breast, and Colorectal Cancer.* J Natl Cancer Inst, 2015. **107**(11).

70. Rashkin, S.R., et al., *Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts.* Nat Commun, 2020. **11**(1): p. 4423.

71. Yang, H., et al., *Inherited factors contribute to an inverse association between preeclampsia and breast cancer.* Breast Cancer Res, 2018. **20**(1): p. 6.

72. Jiang, X., et al., *Shared heritability and functional enrichment across six solid cancers.* Nature Communications, 2019. **10**(1): p. 431.

73. Hulka, B.S. and A.T. Stark, *Breast cancer: cause and prevention.* Lancet, 1995. **346**(8979): p. 883-7.

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

75. Cancer, C.G.o.H.F.i.B., *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.

76. Rosner, B., G.A. Colditz, and W.C. Willett, *Reproductive risk factors in a prospective study of breast cancer: the Nurses' Health Study.* Am J Epidemiol, 1994. **139**(8): p. 819-35.

77. Li, H., et al., *BMI, reproductive factors, and breast cancer molecular subtypes: A case-control study and meta-analysis.* J Epidemiol, 2017. **27**(4): p. 143-151.

78. Moorman, P.G., et al., *Oral contraceptives and risk of ovarian cancer and breast cancer among high-risk women: a systematic review and meta-analysis.* J Clin Oncol, 2013. **31**(33): p. 4188-98.

79. Fortner, R.T., et al., *Parity, breastfeeding, and breast cancer risk by hormone receptor status and molecular phenotype: results from the Nurses' Health Studies.* Breast Cancer Res, 2019. **21**(1): p. 40.

80. Islami, F., et al., *Breastfeeding and breast cancer risk by receptor status—a systematic review and meta-analysis.* Annals of Oncology, 2015. **26**(12): p. 2398-2407.

81. Ma, H., et al., *Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies.* Breast Cancer Research, 2006. **8**(4): p. R43.

82. Pearlman, M.D., *Exogenous Hormones and Breast Cancer Risk: Contraception, Menopausal Hormone Therapy, and Breast Cancer Survivors.* Clin Obstet Gynecol, 2022. **65**(3): p. 510-523.

83. Tjønneland, A., et al., *Hormone replacement therapy in relation to breast carcinoma incidence rate ratios: a prospective Danish cohort study.* Cancer, 2004. **100**(11): p. 2328-37.

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

85. Rossouw, J.E., et al., *Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial.* Jama, 2002. **288**(3): p. 321-33.

86. Lobo, R.A., *Hormone-replacement therapy: current thinking.* Nature Reviews Endocrinology, 2017. **13**(4): p. 220-231.

87. Investigators, W.G.f.t.W.s.H.I. and W.G.f.t.W.s.H.I. Investigators, *Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial.* Jama, 2002. **288**(3): p. 321-333.

88. Fitzpatrick, D., et al., *Combined and progestagen-only hormonal contraceptives and breast cancer risk: A UK nested case-control study and meta-analysis.* PLoS Med, 2023. **20**(3): p. e1004188.

89. Cancer, C.G.o.H.F.i.B., *Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies.* The Lancet, 1996. **347**(9017): p. 1713-1727.

90. Morch, L.S., et al., *Contemporary Hormonal Contraception and the Risk of Breast Cancer.* N Engl J Med, 2017. **377**(23): p. 2228-2239.

91. Kerr, J., C. Anderson, and S.M. Lippman, *Physical activity, sedentary behaviour, diet, and cancer: an update and emerging new evidence.* Lancet Oncol, 2017. **18**(8): p. e457-e471.

92. Jones, M.E., et al., *Smoking and risk of breast cancer in the Generations Study cohort.* Breast Cancer Res, 2017. **19**(1): p. 118.

93. Gaudet, M.M., et al., *Active smoking and breast cancer risk: original cohort data and meta-analysis.* J Natl Cancer Inst, 2013. **105**(8): p. 515-25.

94. Chen, W.Y., et al., *Moderate alcohol consumption during adult life, drinking patterns, and breast cancer risk.* Jama, 2011. **306**(17): p. 1884-90.

95. Boffetta, P. and M. Hashibe, *Alcohol and cancer.* Lancet Oncol, 2006. **7**(2): p. 149-56.

96. Chen, X., et al., *Physical Activity and Risk of Breast Cancer: A Meta-Analysis of 38 Cohort Studies in 45 Study Reports.* Value Health, 2019. **22**(1): p. 104-128.

97. Pizot, C., et al., *Physical activity, hormone replacement therapy and breast cancer risk: A meta-analysis of prospective studies.* Eur J Cancer, 2016. **52**: p. 138-54.

98. Kyu, H.H., et al., *Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013.* bmj, 2016. **354**.

99. Pfeiffer, R.M., et al., *Proportion of U.S. Trends in Breast Cancer Incidence Attributable to Long-term Changes in Risk Factor Distributions.* Cancer Epidemiol Biomarkers Prev, 2018. **27**(10): p. 1214-1222.

100. Picon-Ruiz, M., et al., *Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention.* CA Cancer J Clin, 2017. **67**(5): p. 378-397.

101. Renehan, A.G., et al., *Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies.* Lancet, 2008. **371**(9612): p. 569-78.

102. Cecchini, R.S., et al., *Body mass index and the risk for developing invasive breast cancer among high-risk women in NSABP P-1 and STAR breast cancer prevention trials.* Cancer Prev Res (Phila), 2012. **5**(4): p. 583-92.

103. Amadou, A., et al., *Overweight, obesity and risk of premenopausal breast cancer according to ethnicity: a systematic review and dose-response meta-analysis.* Obes Rev, 2013. **14**(8): p. 665-78.

104. Zhou, Y., H. Zhao, and C. Peng, *Association of sedentary behavior with the risk of breast cancer in women: update meta-analysis of observational studies.* Annals of epidemiology, 2015. **25**(9): p. 687-697.

105. Boyd, N.F., et al., *Mammographic breast density as an intermediate phenotype for breast cancer.* The Lancet. Oncology, 2005. **6**(10): p. 798-808.

106. Alsheh Ali, M., et al., *Association of Microcalcification Clusters with Short-term Invasive Breast Cancer Risk and Breast Cancer Risk Factors.* Sci Rep, 2019. **9**(1): p. 14604.

107. Stomper, P.C., et al., *Mammographic predictors of the presence and size of invasive carcinomas associated with malignant microcalcification lesions without a mass.* AJR Am J Roentgenol, 2003. **181**(6): p. 1679-84.

108. Azam, S., et al., *Mammographic microcalcifications and risk of breast cancer.* British journal of cancer, 2021. **125**(5): p. 759-765.

109. Apantaku, L.M., *Breast cancer diagnosis and screening.* American Family Physician, 2000. **62**(3): p. 596-602.

110. Guray, M. and A.A. Sahin, *Benign breast diseases: classification, diagnosis, and management.* The oncologist, 2006. **11**(5): p. 435-449.

111. Johansson, A., et al., *Characterization of Benign Breast Diseases and Association With Age, Hormonal Factors, and Family History of Breast Cancer Among Women in Sweden.* JAMA Netw Open, 2021. **4**(6): p. e2114716.

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

113. Dyrstad, S.W., et al., *Breast cancer risk associated with benign breast disease: systematic review and meta-analysis.* Breast Cancer Res Treat, 2015. **149**(3): p. 569-75.

114. Tice, J.A., et al., *Breast Density and Benign Breast Disease: Risk Assessment to Identify Women at High Risk of Breast Cancer.* J Clin Oncol, 2015. **33**(28): p. 3137-43.

115. Gail, M.H., et al., *Projecting individualized probabilities of developing breast cancer for white females who are being examined annually.* J Natl Cancer Inst, 1989. **81**(24): p. 1879-86.

116. Antoniou, A.C., et al., *The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions.* British Journal of Cancer, 2008. **98**(8): p. 1457-1466.

117. 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.

118. Brentnall, A.R., et al., *Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort.* Breast Cancer Research, 2015. **17**(1): p. 147.

119. Terry, M.B., et al., *10-year performance of four models of breast cancer risk: a validation study.* Lancet Oncol, 2019. **20**(4): p. 504-517.

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

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

122. Eriksson, M., et al., *A clinical model for identifying the short-term risk of breast cancer.* Breast Cancer Res, 2017. **19**(1): p. 29.

123. Eriksson, M., et al., *Identification of Women at High Risk of Breast Cancer Who Need Supplemental Screening.* Radiology, 2020. **297**(2): p. 327-333.

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

125. Kemp Jacobsen, K., et al., *Comparing sensitivity and specificity of screening mammography in the United States and Denmark.* Int J Cancer, 2015. **137**(9): p. 2198-207.

126. Hofvind, S., et al., *Interval Breast Cancer Rates and Histopathologic Tumor Characteristics after False-Positive Findings at Mammography in a Population-based Screening Program.* Radiology, 2018. **287**(1): p. 58-67.

127. 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.

128. Ripping, T.M., et al., *Towards personalized screening: Cumulative risk of breast cancer screening outcomes in women with and without a first-degree relative with a history of breast cancer.* Int J Cancer, 2016. **138**(7): p. 1619-25.

129. Parvinen, I., et al., *Mammography screening in three Finnish residential areas: comprehensive population-based study of breast cancer incidence and incidence-based mortality 1976-2009.* Br J Cancer, 2015. **112**(5): p. 918-24.

130. Giordano, L., et al., *Mammographic screening programmes in Europe: organization, coverage and participation.* Journal of medical screening, 2012. **19**(1_suppl): p. 72-82.

131. Williams, J., et al., *Breast cancer screening in England and the United States: a comparison of provision and utilisation.* Int J Public Health, 2015. **60**(8): p. 881-90.

132. Breast cancer research foundation. *What to Know About New Breast Cancer Screening Recommendations*. 2023 [cited 2023 August 9]; Available from: https://www.bcrf.org/blog/uspstf-new-breast-cancer-screening-guidelines-2023/.

133. American Cancer Society. *American Cancer Society Recommendations for the Early Detection of Breast Cancer*. 2022 [cited 2023 August 9]; Available from: https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast-cancer.html.

134. National Board of Health and Welfare, S. *Screening för bröstcancer*. 2023 [cited 2023 August 9]; Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-screeningprogram/2023-2-8360.pdf.

135. 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.

136. Lidbrink, E.K., et al., *The general mammography screening program in Stockholm. Organisation and first-round results.* Acta Oncol, 1994. **33**(4): p. 353-8.

137. Perry, N., et al., *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document.* Ann Oncol, 2008. **19**(4): p. 614-22.

138. von Euler-Chelpin, M., M. Kuchiki, and I. Vejborg, *Increased risk of breast cancer in women with false-positive test: the role of misclassification.* Cancer Epidemiol, 2014. **38**(5): p. 619-22.

139. Román, M., et al., *Risk of breast cancer after false-positive results in mammographic screening.* Cancer Med, 2016. **5**(6): p. 1298-306.

140. 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.

141. Boyd, N.F., et al., *Mammographic densities and breast cancer risk.* Breast Dis, 1998. **10**(3-4): p. 113-26.

142. Holowko, N., et al., *Heritability of Mammographic Breast Density, Density Change, Microcalcifications, and Masses.* Cancer research, 2020. **80**(7): p. 1590-1600.

143. Ursin, G., et al., *The relative importance of genetics and environment on mammographic density.* Cancer Epidemiol Biomarkers Prev, 2009. **18**(1): p. 102-12.

144. Trinh, T., et al., *Background risk of breast cancer influences the association between alcohol consumption and mammographic density.* Br J Cancer, 2015. **113**(1): p. 159-65.

145. Hutson, S.W., P.N. Cowen, and C.C. Bird, *Morphometric studies of age related changes in normal human breast and their significance for evolution of mammary cancer.* J Clin Pathol, 1985. **38**(3): p. 281-7.

146. 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.

147. 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.

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

149. American College of Radiology. *ACR BI-RADS Atlas Fifth Edition*. 2013 [cited 2020 15 Dec]; Available from: https://www.acr.org/-/media/ACR/Files/RADS/BI-RADS/BIRADS-Reference-Card.pdf.

150. 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.

151. Sovio, U., et al., *Comparison of fully and semi-automated area-based methods for measuring mammographic density and predicting breast cancer risk.* British Journal of Cancer, 2014. **110**(7): p. 1908-1916.

152. Partain, N., et al., *Mammographic density changes in surgical weight loss-an indication for personalized screening.* BMC Medical Imaging, 2018. **18**(1): p. 10.

153. American Cancer Society. *Breast Density and Your Mammogram Report*. 2023 March 28, 2023 [cited 2021; Available from: https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/breast-density-and-your-mammogram-report.html.

154. Mann, R.M., et al., *Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI).* European radiology, 2022. **32**(6): p. 4036-4045.

155. Venkatesan, A., et al., *Positive predictive value of specific mammographic findings according to reader and patient variables.* Radiology, 2009. **250**(3): p. 648-57.

156. Ferranti, C., et al., *Relationships between age, mammographic features and pathological tumour characteristics in non-palpable breast cancer.* Br J Radiol, 2000. **73**(871): p. 698-705.

157. Tabar, L., et al., *Mammographic tumor features can predict long-term outcomes reliably in women with 1-14-mm invasive breast carcinoma.* Cancer, 2004. **101**(8): p. 1745-59.

158. Cox, R.F., et al., *Microcalcifications in breast cancer: novel insights into the molecular mechanism and functional consequence of mammary mineralisation.* British Journal of Cancer, 2012. **106**(3): p. 525-537.

159. Strand, F., *Determinants of interval cancer and tumor size among breast cancer screening participants*. 2018, Karolinska Institutet.

160. *ICAD Profound AI solution*. [cited 2022 March 11]; Available from: https://www.icadmed.com/profoundai.html.

161. Azam, S., et al., *Predictors of mammographic microcalcifications.* Int J Cancer, 2020.

162. Coffey, K. and M.S. Jochelson, *Contrast-enhanced mammography in breast cancer screening.* European Journal of Radiology, 2022. **156**: p. 110513.

163. Marinovich, M.L., et al., *Breast cancer screening using tomosynthesis or mammography: a meta-analysis of cancer detection and recall.* JNCI: Journal of the National Cancer Institute, 2018. **110**(9): p. 942-949.

164. Iranmakani, S., et al., *A review of various modalities in breast imaging: technical aspects and clinical outcomes.* Egyptian Journal of Radiology and Nuclear Medicine, 2020. **51**(1): p. 57.

165. Colzani, E., et al., *Prognosis of patients with breast cancer: causes of death and effects of time since diagnosis, age, and tumor characteristics.* J Clin Oncol, 2011. **29**(30): p. 4014-21.

166. Fisher, E.R., et al., *Pathologic findings from the National Surgical Adjuvant Breast Project (Protocol 4). Discriminants for 15-year survival. National Surgical Adjuvant Breast and Bowel Project Investigators.* Cancer, 1993. **71**(6 Suppl): p. 2141-50.

167. Paik, S., et al., *Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: prognostic significance of erbB-2 protein overexpression in primary breast cancer.* J Clin Oncol, 1990. **8**(1): p. 103-12.

168. Rakha, E.A., et al., *Breast cancer prognostic classification in the molecular era: the role of histological grade.* Breast Cancer Res, 2010. **12**(4): p. 207.

169. Purdie, C.A., et al., *Progesterone receptor expression is an independent prognostic variable in early breast cancer: a population-based study.* Br J Cancer, 2014. **110**(3): p. 565-72.

170. Li, Z., et al., *The Role of Progesterone Receptors in Breast Cancer.* Drug Des Devel Ther, 2022. **16**: p. 305-314.

171. Kunte, S., J. Abraham, and A.J. Montero, *Novel HER2-targeted therapies for HER2-positive metastatic breast cancer.* Cancer, 2020. **126**(19): p. 4278-4288.

172. Kumar, P. and R. Aggarwal, *An overview of triple-negative breast cancer.* Archives of gynecology and obstetrics, 2016. **293**: p. 247-269.

173. Lundgren, C., et al., *Agreement between molecular subtyping and surrogate subtype classification: a contemporary population-based study of ER-positive/HER2-negative primary breast cancer.* Breast Cancer Res Treat, 2019. **178**(2): p. 459-467.

174. Goldhirsch, A., et al., *Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013.* Ann Oncol, 2013. **24**(9): p. 2206-23.

175. Coates, A.S., et al., *Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015.* Ann Oncol, 2015. **26**(8): p. 1533-46.

176. Pálka, I., et al., *Tumor characteristics in screen-detected and symptomatic breast cancers.* Pathol Oncol Res, 2008. **14**(2): p. 161-7.

177. Duffy, S.W., et al., *Breast screening, prognostic factors and survival--results from the Swedish two county study.* Br J Cancer, 1991. **64**(6): p. 1133-8.

178. O'Brien, K.M., et al., *Screening status, tumour subtype, and breast cancer survival: a national population-based analysis.* Breast Cancer Res Treat, 2018. **172**(1): p. 133-142.

179. Koh, V.C., et al., *Characteristics and behaviour of screen-detected ductal carcinoma in situ of the breast: comparison with symptomatic patients.* Breast Cancer Res Treat, 2015. **152**(2): p. 293-304.

180. Sihto, H., et al., *Molecular subtypes of breast cancers detected in mammography screening and outside of screening.* Clin Cancer Res, 2008. **14**(13): p. 4103-10.

181. Dawson, S.J., et al., *Molecular characteristics of screen-detected vs symptomatic breast cancers and their impact on survival.* Br J Cancer, 2009. **101**(8): p. 1338-44.

182. Chuang, S.L., et al., *Using tumor phenotype, histological tumor distribution, and mammographic appearance to explain the survival differences between screen-detected and clinically detected breast cancers.* Apmis, 2014. **122**(8): p. 699-707.

183. Joensuu, H., et al., *Risk for distant recurrence of breast cancer detected by mammography screening or other methods.* Jama, 2004. **292**(9): p. 1064-73.

184. Ekbom, A., *The Swedish Multi-generation Register.* Methods Mol Biol, 2011. **675**: p. 215-20.

185. The National Board of Health and Welfare. *The Cancer Registry at the National Board of Health and Welfare*. 2023 [cited 2023 August 11]; Available from: https://cancercentrum.se/samverkan/regional-cancer-centres/knowledge-based-healthcare/cancer-registry/.

186. MATTSSON, B., L.E. RUTQVIST, and A. WALLGREN, *Undernotification of diagnosed cancer cases to the Stockholm Cancer Registry.* International journal of epidemiology, 1985. **14**(1): p. 64-69.

187. Löfgren, L., et al., *Validation of data quality in the Swedish National Register for Breast Cancer.* BMC Public Health, 2019. **19**(1): p. 495.

188. Brooke, H.L., et al., *The Swedish cause of death register.* Eur J Epidemiol, 2017. **32**(9): p. 765-773.

189. Ludvigsson, J.F., et al., *Registers of the Swedish total population and their use in medical research.* Eur J Epidemiol, 2016. **31**(2): p. 125-36.

190. Ludvigsson, J.F., et al., *The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research.* Eur J Epidemiol, 2019. **34**(4): p. 423-437.

191. 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.

192. Li, J., et al., *Breast cancer genetic risk profile is differentially associated with interval and screen-detected breast cancers.* Ann Oncol, 2015. **26**(3): p. 517-22.

193. Li, J., et al., *Differential Burden of Rare and Common Variants on Tumor Characteristics, Survival, and Mode of Detection in Breast Cancer.* Cancer Res, 2018. **78**(21): p. 6329-6338.

194. Ernster, V.L., *Nested case-control studies.* Preventive medicine, 1994. **23**(5): p. 587-590.

195. Richardson, D., *An incidence density sampling program for nested case-control analyses.* Occupational and environmental medicine, 2004. **61**(12): p. e59-e59.

196. Mann, C., *Observational research methods. Research design II: cohort, cross sectional, and case-control studies.* Emergency medicine journal, 2003. **20**(1): p. 54-60.

197. Grimes, D.A. and K.F. Schulz, *Cohort studies: marching towards outcomes.* The Lancet, 2002. **359**(9303): p. 341-345.

198. Sjölander, A. and S. Greenland, *Ignoring the matching variables in cohort studies - when is it valid and why?* Stat Med, 2013. **32**(27): p. 4696-708.

199. Eriksson, M., et al., *A comprehensive tool for measuring mammographic density changes over time.* Breast Cancer Res Treat, 2018. **169**(2): p. 371-379.

200. Decker, B., et al., *Rare, protein-truncating variants in ATM, CHEK2 and PALB2, but not XRCC2, are associated with increased breast cancer risks.* J Med Genet, 2017. **54**(11): p. 732-741.

201. Ellis, I., et al., *Pathology reporting of breast disease is surgical excision specimens incorporating the dataset for histological reporting of breast cancer. London: The Royal College of Pathologists; 2016.* Published by The Royal College of Pathologists, 2019.

202. Swedish Society of Pathology. *SNOMED codes in numerical order.* . Jan 2, 2021]; Available from: http://www.svfp.se/snomed_numerisk_ordning.

203. Hosmer Jr, D.W., S. Lemeshow, and R.X. Sturdivant, *Applied logistic regression*. Vol. 398. 2013: John Wiley & Sons.

204. Christensen, E., *Multivariate survival analysis using Cox's regression model.* Hepatology, 1987. **7**(6): p. 1346-1358.

205. Royston, P. and P.C. Lambert, *Flexible parametric survival analysis using Stata: beyond the Cox model*. Vol. 347. 2011: Stata press College Station, TX.

206. Lambert, P.C. and P. Royston, *Further development of flexible parametric models for survival analysis.* The Stata Journal, 2009. **9**(2): p. 265-290.

207. Nelson, H.D., et al., *Factors Associated With Rates of False-Positive and False-Negative Results From Digital Mammography Screening: An Analysis of Registry Data.* Ann Intern Med, 2016. **164**(4): p. 226-35.

208. Cole, E.B., et al., *Impact of computer-aided detection systems on radiologist accuracy with digital mammography.* AJR Am J Roentgenol, 2014. **203**(4): p. 909-16.

209. Khoo, L.A., P. Taylor, and R.M. Given-Wilson, *Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study.* Radiology, 2005. **237**(2): p. 444-9.

210. Lee, A.J., et al., *BOADICEA breast cancer risk prediction model: updates to cancer incidences, tumour pathology and web interface.* British Journal of Cancer, 2014. **110**(2): p. 535-545.

211. von Euler-Chelpin, M., et al., *Risk of breast cancer after false-positive test results in screening mammography.* Journal of the National Cancer Institute, 2012. **104**(9): p. 682-689.

212. Ciatto, S., et al., *Minority report - false negative breast assessment in women recalled for suspicious screening mammography: imaging and pathological features, and associated delay in diagnosis.* Breast Cancer Res Treat, 2007. **105**(1): p. 37-43.

213. Stone, J., *Should breast cancer screening programs routinely measure mammographic density?* J Med Imaging Radiat Oncol, 2018. **62**(2): p. 151-158.

214. Monticciolo, D.L., et al., *Breast Cancer Screening in Women at Higher-Than-Average Risk: Recommendations From the ACR.* J Am Coll Radiol, 2018. **15**(3 Pt A): p. 408-414.

215. Easton, D.F., D. Ford, and D.T. Bishop, *Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium.* Am J Hum Genet, 1995. **56**(1): p. 265-71.

216. Mao, X., et al., *Factors Associated With False-Positive Recalls in Mammography Screening.* J Natl Compr Canc Netw, 2023. **21**(2): p. 143-152.e4.

217. Henderson, L.M., et al., *Increased Risk of Developing Breast Cancer after a False-Positive Screening Mammogram.* Cancer Epidemiol Biomarkers Prev, 2015. **24**(12): p. 1882-9.

218. Rodríguez-Ruiz, A., et al., *Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System.* Radiology, 2019. **290**(2): p. 305-314.

219. Dembrower, K., et al., *Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study.* Lancet Digit Health, 2023.

220. Cazzaniga, M. and B. Bonanni, *Breast cancer chemoprevention: old and new approaches.* Journal of Biomedicine and Biotechnology, 2012. **2012**.

221. Smith, S.G., et al., *Factors affecting uptake and adherence to breast cancer chemoprevention: a systematic review and meta-analysis.* Annals of Oncology, 2016. **27**(4): p. 575-590.

222. Eriksson, M., et al., *Low-Dose Tamoxifen for Mammographic Density Reduction: A Randomized Controlled Trial.* J Clin Oncol, 2021. **39**(17): p. 1899-1908.